

Universidad Cenfotec

Maestría en Tecnología de Bases de Datos

Documento final de Proyecto de Investigación Aplicada 2

Implementación de una solución de inteligencia de negocios para el proceso de matrícula de la Universidad Cenfotec

Estudiantes:

Obando Barrantes Esteban.

Webb Calderón Gerald.

Febrero 2018.

Derechos de autor

Se prohíbe la reproducción total o parcial de este documento por cualquier medio, sin el previo consentimiento otorgado por escrito de la Universidad Cenfotec a cualquier persona y/o actividad que sean ajenas a la misma. Los lectores puedan consultar el contenido de este trabajo para fines exclusivos académicos.

Agradecimientos

A Dios.

Por darnos la sabiduría y fuerza para culminar esta etapa académica.

A nuestros padres.

Por el apoyo y valores brindados a lo largo de nuestro proceso educativo.

A nuestro tutor de proyecto, el profesor José Jaikel Cabezas.

Por su guía, comprensión, paciencia, entrega y valiosos consejos.

Al personal de la Universidad Cenfotec.

Por brindarnos su apoyo y por ofrecernos la información requerida este proyecto.

Resumen y Palabras Clave

Resumen

El presente proyecto abarca el diseño y la implementación de un almacén de datos que sirve como herramienta de apoyo para la toma de decisiones de la Universidad Cenfotec. Este centro educativo necesitaba de un sistema que colabore con el análisis de datos relacionados con la matrícula de estudiantes cada cuatrimestre, debido a que era un proceso manual el cual resultaba de gran dificultad para los usuarios finales. Debido a esto, surgió la necesidad de un sistema automatizado, fácil de utilizar que les permita el análisis de datos de una forma eficiente y segura. El sistema necesitaba contemplar la información de los alumnos, de cursos, de las carreras, de las convalidaciones y retiros de los estudiantes en la universidad, para la toma de decisiones.

La solución desarrollada a este problema es un Depósito de Datos y una Base de Datos multidimensional que contemplan todas las áreas requeridas por la empresa; con este proyecto se buscó el hacer la información fácil de analizar por los usuarios finales, y fácil de mantener por el departamento de tecnología de la universidad.

Palabras Clave

Almacén de datos, Data Warehouse, Class, ETL, Consultas, SQL Server, Matricula, SQL, T-SQL, SSIS, Herramienta de Inteligencia de negocios universitaria.

Tabla de Contenido

Derechos de autor 2

Agradecimientos 3

Resumen y Palabras Clave 4

Resumen 4

Palabras Clave 4

1. Introducción 14

1.1. Generalidades 14

1.2. Antecedentes del Problema 15

1.3. Definición y Descripción del Problema 16

1.4. Justificación 16

1.5. Viabilidad 17

1.5.1. Punto de Vista Técnico 17

1.5.2. Punto de Vista Operativo 18

1.5.3. Punto de Vista Económico 18

1.6. Objetivos 19

1.6.1. Objetivo General 20

1.6.2. Objetivos Específicos 20

1.7. Alcances y Limitaciones 20

1.7.1. Alcances 20

1.7.2. Limitaciones 21
1.8. Marco de Referencia Organizacional y Socioeconómico 22
1.8.1. Historia 22
1.8.2. Tipo de Negocio y Mercado Meta 22
1.8.3. Misión, Visión y Valores 23
1.8.4. Políticas Institucionales 24
1.9. Estado de la Cuestión 24
1.9.1. Repositorios de Información 24
1.9.2. Artículos o Libros Útiles para el Proyecto 25
2. Marco Conceptual 27
2.1. Sistema Fuente 27
2.2. Data Warehouse 28
2.2.1. Validar la Tabla de Hechos 28
2.2.2. Proceso de Granularidad 28
2.2.3. Seleccionar Dimensiones 29
2.2.4. Definir los Atributos de cada Dimensión 29
2.2.5. Identificar el Tipo de Campo 30
2.2.6. Validar restricciones de integridad 30
2.2.7. Volumen de los datos 30
2.3. Acceso de Datos Usuario Final 31

3. Marco Metodológico 32
3.1. Tipo de Investigación 32
3.2. Alcance Investigativo 32
3.3. Enfoque 32
3.4. Diseño 35
3.5. Instrumentos de Recolección de Datos 35
3.6. Técnicas de Análisis de la Información 36
3.7. Estrategia de Desarrollo de la Propuesta 36
4. Análisis del Diagnóstico 38
5. Propuesta de Solución 40
5.1. Definición de Recursos 40
5.2. Análisis de Fuente de Datos 41
5.3. Herramientas Por Utilizar 41
5.3.1. Sistema Gestor de Base de Datos 41
5.3.2. Microsoft SQL Server Integration Services 2012 41
5.3.3. Microsoft SQL Server Analysis Services 2012 41
5.4. Arquitectura de la Solución 42
5.4.1. Arquitectura de Hardware 42
5.4.2. Arquitectura de Software 43
5.4.3. Estándares de Nomenclatura de Base de Datos Transaccional 44

5.4.4. Estándar de nomenclatura de Staging Area 44
5.4.5. Estándar de nomenclatura de Data Warehouse 44
5.4.6. Estándar de nomenclatura de tareas programadas 45
5.4.7. Estándar de nomenclatura de Integration Services 45
5.4.8. Estándar de nomenclatura de Analysis Services 46
5.5. Etapa de diseño y desarrollo 46
5.5.1. Diseño y desarrollo del Staging Area 46
5.5.2. Diseño de arquitectura del Staging Area 47
5.5.3. Definición de tablas del Staging Area 48
5.5.4. Estrategia de ETL del Staging Area 49
5.5.5. Consideraciones de la creación del Staging Area 56
5.5.6. Diseño y desarrollo del DataWarehouse 58
5.5.7. Diseño de arquitectura del Data Warehouse 59
5.5.8. Definición de dimensiones y tabla de hechos 60
5.5.9. Estrategia de ETL para el Data Warehouse 61
5.5.10. Consideraciones de la creación del DataWarehouse 73
5.5.11. Diseño y desarrollo de la Base de Datos Multidimensional 76
5.5.12. Diseño de arquitectura de la Base de Datos Multidimensional 77
5.5.13. Definición de jerarquías 77
5.5.14. Definición de perspectivas 78

5.5.15. Estrategia de Base de Datos Multidimensional 78

5.5.16. Consideraciones de la creación de la Base de Datos Multidimensional 87

5.5.17. Estrategia de consultas del modelo..... 88

5.5.18. Consideraciones de Seguridad 91

5.6. Diccionario de datos..... 92

5.7. Entregables..... 93

6. Conclusiones y Recomendaciones..... 94

6.1. Conclusiones 94

6.2. Recomendaciones 96

7. Reflexiones Finales 97

8. Trabajos a Futuro 98

9. Referencias 99

10. Bibliografía..... 100

Publicaciones..... 100

11. Anexos..... 101

11.1. Diccionario de datos del Staging Area 101

11.2. Diccionario de datos del Data Warehouse 114

11.3. Manual de Usuario 125

11.3.1. Herramienta Microsoft Excel..... 125

11.3.2. Microsoft Power BI 127

11.4. Manual de Instalación 130

12. Apéndice 136

12.1. Cronograma 136

Tabla de Ilustraciones

Ilustración 1 Punto de Vista Económico. 19

Ilustración 2 Proceso de Generación del Data Warehouse. 27

Ilustración 3 Dimensión Ontológica. Fuente: Elaboración Propia. 33

Ilustración 4 Criterios de evaluación de la facilidad de implementación. 35

Ilustración 5 Instrumentos de Recolección de Datos. 36

Ilustración 6 Roles y Responsabilidades 40

Ilustración 7 Arquitectura de Hardware. Elaboración Propia. 42

Ilustración 8 Arquitectura de Software. Fuente: Elaboración Propia. 43

Ilustración 9 Estándares de Nomenclatura de Base de Datos Transaccional. 44

Ilustración 10 Estándar de nomenclatura de Staging Area. 44

Ilustración 11 Estándar de nomenclatura de Data Warehouse. 45

Ilustración 12 Estándar de nomenclatura de tareas programadas. 45

Ilustración 13 Estándar de nomenclatura de Integration Services. 46

Ilustración 14 Estándar de nomenclatura de Analysis Services. 46

Ilustración 15 Diseño de arquitectura del Staging Area. 47

Ilustración 16 Estrategia del ETL del Staging Area 1. 50
Ilustración 17 Estrategia del ETL del Staging Area 2. 51
Ilustración 18 Estrategia del ETL del Staging Area 3. 51
Ilustración 19 Estrategia del ETL del Staging Area 4. 52
Ilustración 20 Estrategia del ETL del Staging Area 5. 52
Ilustración 21 Estrategia del ETL del Staging Area 6. 53
Ilustración 22 Estrategia del ETL del Staging Area 7. 53
Ilustración 23 Estrategia del ETL del Staging Area 8. 54
Ilustración 24 Estrategia del ETL del Staging Area 9. 55
Ilustración 25 Estrategia del ETL del Staging Area 10. 55
Ilustración 26 Estrategia del ETL del Staging Area 11. 57
Ilustración 27 Estrategia del ETL del Staging Area 12. 58
Ilustración 28 Diseño y desarrollo del Data Warehouse 1. 59
Ilustración 29 Diseño y desarrollo del Data Warehouse 4. 63
Ilustración 30 Diseño y desarrollo del Data Warehouse 5. 64
Ilustración 31 Diseño y desarrollo del Data Warehouse 6. 64
Ilustración 32 Diseño y desarrollo del Data Warehouse 7. 65
Ilustración 33 Diseño y desarrollo del Data Warehouse 8. 66
Ilustración 34 Diseño y desarrollo del Data Warehouse 9. 67
Ilustración 35 Diseño y desarrollo del Data Warehouse 10. 67
Ilustración 36 Diseño y desarrollo del Data Warehouse 11. 68
Ilustración 37 Diseño y desarrollo del Data Warehouse 12. 68
Ilustración 38 Diseño y desarrollo del Data Warehouse 13. 69
Ilustración 39 Diseño y desarrollo del Data Warehouse 14. 69

Ilustración 40 Diseño y desarrollo del Data Warehouse 15. 70
Ilustración 41 Diseño y desarrollo del Data Warehouse 16. 70
Ilustración 42 Diseño y desarrollo del Data Warehouse 17. 71
Ilustración 43 Diseño y desarrollo del Data Warehouse 18. 71
Ilustración 44 Diseño y desarrollo del Data Warehouse 19. 72
Ilustración 45 Diseño y desarrollo del Data Warehouse 20. 72
Ilustración 46 Diseño y desarrollo del Data Warehouse 21. 73
Ilustración 47 Diseño y desarrollo del Data Warehouse 22. 74
Ilustración 48 Diseño y desarrollo del Data Warehouse 23. 75
Ilustración 49 Diseño y desarrollo del Data Warehouse 24. 75
Ilustración 50 Diseño y desarrollo del Data Warehouse 25. 76
Ilustración 51 Diseño de arquitectura de la Base de Datos Multidimensional 1.. 77
Ilustración 52 Diseño de arquitectura de la Base de Datos Multidimensional 2.. 79
Ilustración 53 Diseño de arquitectura de la Base de Datos Multidimensional 3.. 79
Ilustración 54 Diseño de arquitectura de la Base de Datos Multidimensional 4.. 80
Ilustración 55 Diseño de arquitectura de la Base de Datos Multidimensional 5 .. 80
Ilustración 56 Diseño de arquitectura de la Base de Datos Multidimensional 6.. 81
Ilustración 57 Diseño de arquitectura de la Base de Datos Multidimensional 7.. 81
Ilustración 58 Diseño de arquitectura de la Base de Datos Multidimensional 8.. 82
Ilustración 59 Diseño de arquitectura de la Base de Datos Multidimensional 9.. 82
Ilustración 60 Diseño de arquitectura de la Base de Datos Multidimensional 10.83
Ilustración 61 Diseño de arquitectura de la Base de Datos Multidimensional 11.83
Ilustración 62 Diseño de arquitectura de la Base de Datos Multidimensional 12.84
Ilustración 63 Diseño de arquitectura de la Base de Datos Multidimensional 13.84

Ilustración 64 Diseño de arquitectura de la Base de Datos Multidimensional 14.85
Ilustración 65 Diseño de arquitectura de la Base de Datos Multidimensional 15.85
Ilustración 66 Diseño de arquitectura de la Base de Datos Multidimensional 16.86
Ilustración 67 Diseño de arquitectura de la Base de Datos Multidimensional 17 86
Ilustración 68 Diseño de arquitectura de la Base de Datos Multidimensional 18.87
Ilustración 69 Diseño de arquitectura de la Base de Datos Multidimensional 19.87
Ilustración 70 Estrategia de consultas del modelo 1. 89
Ilustración 71 Estrategia de consultas del modelo 2. 89
Ilustración 72 Estrategia de consultas del modelo 3. 90
Ilustración 73 Estrategia de consultas del modelo 4. 90
Ilustración 74 Estrategia de consultas del modelo 5. 91

1. Introducción

1.1. Generalidades

La propuesta que se expone está dirigida al área financiera y administrativa de la Universidad Cenfotec, en la sede ubicada en la provincia de San José. Esta área tiene a cargo el diseño curricular y ejecución de programas para la formación y actualización de profesionales, así como la realización de programas de investigación en el campo de las tecnologías digitales, en este mundo globalizado donde las tecnologías ocupan un lugar de gran relevancia en la sociedad actual.

Las organizaciones dedican de sus recursos en la búsqueda de innovaciones que les permitan no solo superar a sus competencias sino también implementar mejoras que les brinde una herramienta para realizar sus tareas con mayor eficiencia y calidad, así como mejorar sus estructuras organizacionales. El proyecto desarrollado corresponde a un modelo de inteligencia de negocios que le permita al centro educativo desarrollar una mentalidad estratégica mediante la visualización de los datos desde una perspectiva diferente. Tanto el Data Warehouse como la base de datos multidimensional les facilitará controles, procesos y toma de decisiones dentro de la organización, brindando un claro panorama de la situación actual que esta presenta y cómo poder impulsarse y mejorar a un futuro tanto administrativo como organizacional.

Mediante la elaboración de este Data Warehouse la universidad cuenta con la capacidad de almacenar un histórico de todos los eventos ejecutados en sus sistemas transaccionales, lo que permite así agrupar los módulos de interés de estudio de la empresa y procesar los datos. Se generará una herramienta capaz de consultar a los

usuarios la información requerida para obtener informes que les genere apoyo en la toma de decisiones y cubrir necesidades del negocio para proyectarse en futuras acciones a tomar, y se genera inclusive optimización de recursos y esfuerzos a los puntos verdaderamente críticos de la empresa.

1.2. Antecedentes del Problema

Actualmente la Universidad Cenfotec cuenta con el sistema transaccional llamado Class, el cual le proporciona a la universidad los procesos necesarios para realizar la gestión académica. Mediante esta herramienta se almacena toda la información del proceso de matrícula universitario, así como información de los estudiantes, profesores, administradores y sus gestiones dentro del centro educativo. Esta herramienta únicamente les brinda datos de la situación actual de la universidad y no cuentan con una herramienta que les permita proyectar esta información en un análisis futuro.

El área financiera de la universidad generaba procesos manuales con información relacionada al comportamiento del estudiante durante su proceso de matrícula, hasta la finalización del cuatrimestre como tal, un proceso que conlleva la ejecución de ciertas tareas que al final permiten la obtención de información. No contaban con una herramienta flexible que les reflejara mediante informes la situación del estudiante durante su pertenencia al centro educativo para poder buscar mejorías en sus procesos que les permita atraer nuevos clientes, así también ver factores como el rendimiento y satisfacción personal de los clientes actuales.

El desarrollo del Data Warehouse surgió como respuesta a la problemática de extraer información a partir de los datos almacenados en la base de datos del sistema

transaccional Class, con el fin de que la información extraída por el Data Warehouse les permita una referencia para la toma de decisiones y generar un panorama con mayor claridad de la situación actual de los estudiantes pertenecientes al centro educativo.

1.3. Definición y Descripción del Problema

La Universidad Cenfotec no contaba con una herramienta que les permitiera estudiar el comportamiento de los estudiantes en el módulo de matrícula del sistema transaccional, desde su ingreso a la institución hasta la finalización del curso al que fue registrado.

1.4. Justificación

El presente proyecto nació como consecuencia de la petición realizada por la Universidad Cenfotec, encargada de brindar programas para la formación y actualización de profesionales, así como la realización de programas de investigación en el campo de las tecnologías digitales. Dicho proyecto les brinda una herramienta capaz de facilitar la toma de decisiones tanto gerenciales como financieras dentro de su organización, además permite efectuar tareas como el análisis del ciclo universitario del estudiante, verificación del flujo de matrícula, ya sea desde el ingreso del estudiante hasta la finalización del curso respectivo, y su rendimiento durante la gestión. Al final poder validar el historial del estudiante durante su vida estudiantil.

Toda organización requiere depositar confianza en la toma de decisiones sobre sus negocios, y para realizar dichas decisiones es requerido tener a su disposición datos y cifras que les permitan evaluar la actualidad y futuro de su compañía. Hoy en día la utilización y calidad de la información es de gran importancia en la toma de

decisiones. Las empresas que desean ser competitivas deben basar sus decisiones en información concreta y fiable, para realizar esta tarea es requerido tener a su disposición una información adecuada y de gran relevancia que les permita obtener un claro panorama de lo sucedido en la empresa, y con esto poder tomar las acciones necesarias para mejorar sus labores.

Definimos un Data Warehouse como “Colección de datos o repositorio orientado a temas, integrado, no volátil, de tiempo variante, donde se consolidan todos los sistemas de la organización y permite brindar una herramienta para el proceso de toma de decisiones gerenciales” (Inmon, 1992, p.31). Mediante este proyecto se estará contribuyendo a la Universidad Cenfotec mediante la implementación de un Data Warehouse que les permita solucionar las necesidades del negocio, brindar un soporte en la toma decisiones, generar un amplio panorama del ciclo universitario de los estudiantes y que les brinde una herramienta confiable y flexible para su gestión.

1.5. Viabilidad

1.5.1. Punto de Vista Técnico

En la actualidad existe en la Universidad Cenfotec un equipo capacitado de TI, el cual se encargaría de recolectar la información suministrada por el Data Warehouse, verificar rendimiento, fiabilidad, realizar tareas de mantenimiento y productividad.

Este proyecto les permite a los técnicos de Cenfotec realizar estas tareas de inteligencia de negocios. Para poder realizar estas tareas es necesario que el equipo de los usuarios técnicos es necesario que se realice la debida instalación y acceso a los siguientes componentes:

- Sistema Operativo Windows 8.1 y Windows 10.

- Microsoft Office 2010 o superior.
- Acceso a la herramienta Class, cuya base de datos se encuentra desarrollada en Microsoft SQL Server, y su respectivo respaldo de base de datos para realizar procesos de productividad del Data Warehouse.
- Microsoft SQL Server 2012 Enterprise Edition, con Data Tools.
- Microsoft Power BI.

1.5.2. Punto de Vista Operativo

Con respecto a la factibilidad operacional, el sistema es utilizado por los usuarios de la Universidad Cenfotec, los cuales verificaron si el proyecto cumplía con las necesidades de información, y se obtienen beneficios de este. Además, la herramienta permite implementar mejoras en los métodos de trabajo y genera ganancias en tiempo y costo, así como también brinda una nueva perspectiva de la situación actual del centro educativo. Como proyecto factible les permite a los encargados de la Universidad Cenfotec una herramienta que afecta sus decisiones operativas brindando mejoras en la toma de decisiones, siempre y cuando exista un encargado de administrar la información brindada por dicha herramienta. El proyecto por desarrollar no afecta la operativa de la institución ni interrumpe los procesos que se gestionan dentro de esta.

1.5.3. Punto de Vista Económico

Al ser este un proyecto de TFG realizado por estudiantes de la Universidad Cenfotec, no se cuenta con gastos económicos de desarrollo. Actualmente la Universidad Cenfotec posee licenciamiento de herramientas Microsoft necesarias para

implementar el Data Warehouse, por lo cual no se requieren gastos de equipo. En cuanto recurso humano para desarrollar el sistema, se requiere de:

- Un encargado de TI de la Universidad Cenfotec que conozca a profundidad licenciamiento, equipos, procesos informáticos que se realizan en la universidad.
- Un usuario experto que permita determinar los requerimientos necesarios para satisfacer las necesidades del negocio para el proyecto.
- Dos desarrolladores técnicos encargados de analizar, estructurar, desarrollar e implementar el Data Warehouse.

Además, es importante como costo teórico el validar el costo de factores de tiempo y esfuerzo que conlleva el proyecto para los técnicos encargados de su desarrollo e investigación.

RESUMEN	HORAS	COSTO X HORA	COSTO
Análisis	22,4	5,500	123,200
Documentación	100	5,500	550,000
Desarrollo	96.8	5,500	532,400
Pruebas	32	5,500	176,000
Capacitación	4	5,500	22,000
Manuales	16	5,500	88,000
Implantación	16	5,500	88,000
TOTAL	287.2	5,500	1,579,600

Ilustración 1 Punto de Vista Económico.

1.6. Objetivos

Se ha seleccionado la taxonomía de Benjamin Bloom de 1956 debido a su estructura jerárquica, que asume que éxito de los niveles superiores depende de los niveles inferiores, al igual que en el desarrollo del proyecto.

1.6.1. Objetivo General

Implementar una solución de inteligencia de negocios para el proceso de matrícula de la Universidad Cenfotec.

1.6.2. Objetivos Específicos

- Definir las necesidades del negocio para el estudio de la matrícula de la Universidad Cenfotec.
- Interpretar el proceso de matrícula de la Universidad Cenfotec y su gestión en el sistema Class.
- Desarrollar una solución de inteligencia de negocios para toma decisiones de la Universidad Cenfotec del módulo de matrícula universitaria.
- Analizar un modelo de Data Warehouse que abarque el módulo de matrícula en la base de datos.
- Seleccionar un proceso que brinde de manera clara, concisa y concreta apoyo en la toma de decisiones de la situación actual y futura del módulo de matrícula de la universidad.
- Implementar una estructura flexible con seguridad para la generación de reportes.

1.7. Alcances y Limitaciones

1.7.1. Alcances

En la realización de este proyecto se procederá a la aplicación de los conceptos básicos de un Data Warehouse para construir una herramienta que permita cumplir

todos los objetivos planeados. Los alcances del proyecto se encuentran definidos de la siguiente manera:

- Entregar un modelo Data Warehouse que estará orientado a la captación de información histórica para la elaboración de informes, para la toma de decisiones de la Universidad Cenfotec.
- El Data Warehouse implementado cuenta con un manual de usuario y diccionario de datos que servirá de guía para que el lector identifique los elementos necesarios para su utilización y respectivo mantenimiento.
- Se realizarán las respectivas pruebas e implementación del Data Warehouse desarrollado con resultados satisfactorios, mediante la entrega de una matriz de pruebas que permita validar su correcto funcionamiento.
- Una vez finalizado su desarrollo, se impartirá una sesión de entrenamiento de la utilización del Data Warehouse y su funcionalidad para la generación de reportes.

1.7.2. Limitaciones

Para este proyecto se pueden encontrar los siguientes escenarios que limiten el logro de los objetivos planeados:

- La falta de actualización de los registros del respaldo de la base de datos transaccional brindada por los funcionarios de la Universidad Cenfotec.
- La herramienta por implementar no posee la misma compatibilidad que mantenía en el ambiente de desarrollo, el hardware de producción

presenta dificultades en la ejecución de los procesos del Data Warehouse.

- La información recolectada solo se está obteniendo de los procesos realizados en el sistema transaccional Class, no se cuenta con la información de matrícula de los cursos técnicos y libres que no se generan en esta herramienta.

1.8. Marco de Referencia Organizacional y Socioeconómico

1.8.1. Historia

Cenfotec, como Centro de Formación en Tecnologías de Información y Comunicación, fue creado en enero del año 2000 por un grupo de empresas de desarrollo de software e inversionistas independientes, con el fin de proveer a la incipiente y creciente industria de desarrollo de software con personal muy calificado, especializado y rápidamente operacional que le permitiese competir internacionalmente con altos estándares de calidad (Universidad Cenfotec, 2012).

En el año 2003, Cenfotec desarrolló e impartió la carrera de Ingeniería del software, la primera en Costa Rica y Centroamérica. Posteriormente se crearon programas en las áreas de soporte a la infraestructura, administración de bases de datos, ciberseguridad, telemática y desarrollo y diseño web, con excelente recepción en el sector de las tecnologías (Universidad Cenfotec, 2012).

1.8.2. Tipo de Negocio y Mercado Meta

La Universidad Cenfotec es una institución de educación superior privada con autonomía académica, administrativa y financiera, que tiene como fin primordial el diseño curricular y ejecución de programas para la formación y actualización de

profesionales, así como la realización de programas de investigación en el campo de las tecnologías digitales (Universidad Cenfotec, 2012).

1.8.2.1. Público Meta

Profesionales en informática de reciente graduación, o estudiantes avanzados de una carrera afín (Universidad Cenfotec, 2012).

1.8.3. Misión, Visión y Valores

1.8.3.1. Misión

Formar personas con una educación rigurosa y ética en tecnología como herramienta para descubrir su propio potencial, auto determinar su destino y generar conocimiento útil para impactar positivamente la calidad de vida y competitividad en la sociedad. (Universidad Cenfotec, 2012)

1.8.3.2. Visión

Ser reconocida internacionalmente entre las mejores instituciones de educación superior en tecnología, como referente de excelencia en lo académico y en lo ético. (Universidad Cenfotec, 2012)

1.8.3.3. Valores

- Espíritu de filiación y altruismo
- Integridad, respeto y cumplimiento en todas las facetas de nuestro actuar
- Sentido del logro gracias al esfuerzo

(Universidad Cenfotec, 2012)

1.8.4. Políticas Institucionales

Estar comprometidos a formar individuos motivados y altamente calificados con conocimiento técnico actualizado, valores, actitudes y habilidades, que les permitan desempeñarse integralmente como profesionales y superar las exigencias de las organizaciones donde ejercerán.

Para ello, la Universidad en su quehacer diario estimulará el trabajo en equipo el aprendizaje constante, el espíritu de servicio, el apego a valores y principios, la racionalización y el mejoramiento de los procesos, la retroalimentación de nuestros clientes y colaboradores y la calidad de los servicios académicos y administrativos. (Universidad Cenfotec, 2012)

1.9. Estado de la Cuestión

Con este proyecto tuvimos la gran oportunidad de colaborar con una institución que necesitaba de ayuda en la identificación e implementación de una solución que hiciera posible el estudio y toma de decisiones de las futuras matrículas que vendrán y que han pasado. El fin es poder afrontar un crecimiento continuo con una robusta toma de decisiones basadas en una buena concentración de datos, los cuales sean entendibles y confiables. Como apoyo para la implementación utilizamos las siguientes fuentes.

1.9.1. Repositorios de Información

- https://scholar.google.com/

Este sitio provee una excelente fuente de información, ya que contiene gran variedad de documentos y normalmente se pueden visualizar sin costo alguno.

Utilizaremos los documentos presentes en esta página web como guía en algunos pasos que realizamos para el proyecto en cuestión.

- http://www.ieee.org

Una fuente muy buena de documentos e información, aunque es un poco costosa. Contiene documentos que pueden llegar a ser indispensables a pesar de su costo y significa un gran repositorio de información. Utilizaremos esta página web cuando necesitemos respaldo de información de una fuente confiable y no encontremos información útil de otras fuentes.

1.9.2. Artículos o Libros Útiles para el Proyecto

- https://books.google.co.cr/books?hl=es&lr=&id=XoS2oy1IcB4C&oi=fnd&pg=PA1&dq=data+warehouse&ots=1BMhqCeIdA&sig=rYMbS8_o8hMPvmdrGektBY9DgKA&redir_esc=y#v=onepage&q=data%20warehouse&f=false

El libro “The Data Warehouse Toolkit: The Complete Guide to Dimensional Modeling” contiene múltiples consejos y guías de cómo diseñar e implementar un Data Warehouse correctamente, lo cual es muy útil para la implementación de este Proyecto. Basándonos en la guía planteada en el libro fue que desarrollamos y aplicamos buenas prácticas para desarrollar el depósito de datos que representa actualmente la solución para el problema planteado por la universidad.

- http://ieeexplore.ieee.org/abstract/document/688111/

El artículo propone recomendaciones de cómo modelar y construir una base de datos multidimensional. Sirve como un excelente apoyo en la creación de este tipo de

bases de datos para complementar la manipulación de datos que genera el proyecto en cuestión.

Este proyecto es un plan piloto para otros módulos que la universidad está planeando implementar en el futuro. Eso deja una base para continuar con la implementación de las siguientes modalidades del sistema Class, como el de cajas o el de administración, basándose en esta solución se pueden hacer planes futuros de desarrollo.

2. Marco Conceptual

Para poder generar un modelo multidimensional fue requerido llevar a cabo primero el análisis de los datos con base en definiciones y componentes del Data Warehouse, como lo son las dimensiones y tablas de hechos, las cuales fueron los primeros pasos para tener un modelo estable de inteligencia de negocios. Este proceso se representa en la siguiente imagen:

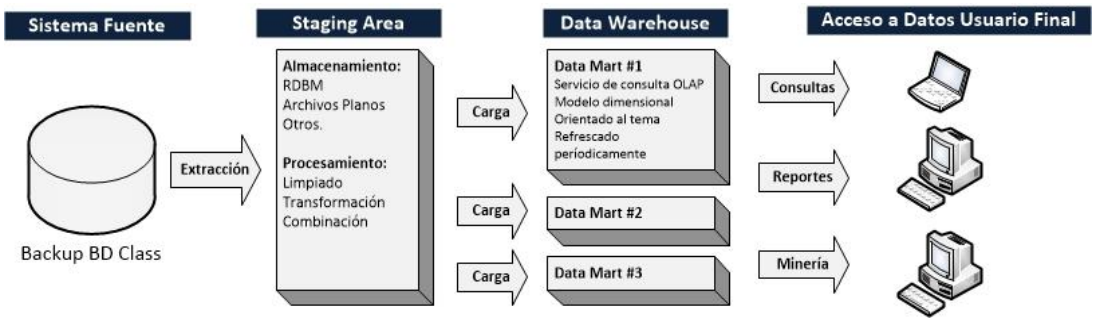


Ilustración 2 Proceso de Generación del Data Warehouse.

2.1. Sistema Fuente

Se cuenta con la base de datos transaccional del sistema llamado Class, la cual proporciona un respaldo de base de datos que permite extraer información relacionada al módulo de matrícula universitaria. Los Sistema Fuente son los sistemas operacionales de registros que tienen como función capturar las transacciones del negocio (Kimball, 2002).

Esta base de datos es un área principal de almacenamiento de datos utilizada para el procesamiento de estos durante procesos de extracción, transformación y carga, los cuales se realizarán mediante ETL utilizando la herramienta Integration Services de Microsoft SQL Server.

2.2. Data Warehouse

En este repositorio se forma el modelo del Data Warehouse, el cual es de tipo estrella, contiene todos los valores necesarios para que el centro educativo realice estudios de la información y brindar resultados de manera eficiente y clara. Según mencionaba Kimball (2002), es la unión de todos los Data Marts que la constituyen. Por su parte, un Data mart es un subgrupo lógico del Data Warehouse (Kimball, 2002). El proceso de este repositorio se encuentra establecido de la siguiente manera:

2.2.1. Validar la Tabla de Hechos

Definimos una tabla de hechos como “una colección de piezas de datos y datos de contexto. Cada hecho representa una parte del negocio, una transacción o un evento” (Kimball, 2002).

Como primer paso en la generación del modelo del Data Warehouse fue requerido identificar la tabla de hechos, la cual es el núcleo del modelo y permite identificar los procesos a nivel del negocio que se trabaja. Este debe demostrar significancia ya que constituye la parte central de un esquema dimensional y además contiene los indicadores de negocio. Su nivel de granularidad representa el nivel más atómico por el cual se definen los datos.

2.2.2. Proceso de Granularidad

En este proceso se define el grado de detalle del modelo. Esta elección depende de varios factores como las necesidades de los usuarios expertos, capacidad de los técnicos que desarrollen el modelo, siempre y cuando se trate de no excederlos detalles, ya que esto podría producir problemas de rendimiento en su generación y también que a su vez la información sea clara y concisa para su estudio y

manipulación. El nivel de granularidad adecuado facilita que el sistema no soporte cargas excesivas de información innecesaria, además permite el suficiente detalle para resolver las necesidades del negocio.

2.2.3. Seleccionar Dimensiones

Una vez que se identificó la tabla de hechos, y seleccionada la granularidad de datos por utilizar, es requerido definir las dimensiones a utilizar, las cuales son “una colección de miembros, unidades o individuos del mismo tipo” (Kimball, 2002). Son elementos que contienen atributos que se utilizan para restringir y agrupar los datos en la tabla de hechos. Contribuyen aportando información sobre los datos que se encuentran ingresados en la tabla de hechos. Una vez seleccionadas estas, se debía determinar si algún dato ha quedado fuera del análisis, además determinar si la tabla de hechos se encontraba relacionada con las dimensiones escogidas. De ser requerido, se debe implementar una dimensión con la combinación de otras, siempre y cuando aporten relevancia al modelo creando una dimensión fuertemente correlacionada, que podría ser el caso de las dimensiones de tiempo.

2.2.4. Definir los Atributos de cada Dimensión

Al definir la temática de las dimensiones que formaron parte del análisis, se empezó la tarea de identificar los atributos que conforman las dimensiones, las cuales le brindan valor al modelo. Se inicia con las dimensiones de Tiempo, validar qué fechas contengan relevancia para el negocio, qué le interesa identificar al negocio en el momento exacto en que sucedieron los hechos, y a qué nivel de detalle se requiere conocer este. Luego se parte con el resto de las dimensiones para encontrar algún tipo de código que nos identifique el registro en la dimensión, así como su descripción

siempre y cuando se mantenga al nivel de granularidad adecuado, entre otras características que facilitarán los resultados e información requerida por el negocio.

2.2.5. Identificar el Tipo de Campo

Se contemplan todos los tipos de campo con que cuentan los atributos tanto desde la base transaccional hasta de como estarán en el Data Warehouse, esto debido a procesos de conversión y transformación de datos, así en caso de requerir concatenar los campos y convertirlos en un único atributo generando modificaciones en el tamaño del tipo de dato.

2.2.6. Validar restricciones de integridad

En este paso se define las bases y restricciones necesarias. Se establecieron las relaciones que conformarán el modelo, se utilizó un atributo que permitió establecer la llave con la que identificamos el registro de nuestro Data Warehouse con la base de datos transaccional y cómo quedan determinadas nuestras dimensiones entre sí al aplicar reglas de integridad.

2.2.7. Volumen de los datos

El volumen de los datos debe estar dentro de los rangos que permite manejar el centro educativo, esto para facilitar tareas de manejo y comprensión de los datos. Los servidores deben tener la capacidad necesaria de extraer y convertir la información en un lapso aceptable por el negocio, así como el motor de base de datos debe contar con las condiciones necesarias.

2.3. Acceso de Datos Usuario Final

En esta etapa se estableció, con la herramienta Analysis Services de Microsoft SQL Server, un aplicativo que mediante generación de cubos, jerarquías y perspectivas permite acceso a los usuarios que posean los permisos respectivos al ingreso de la información. Además, genera reportes flexibles que les brinda una herramienta para la toma de decisiones.

3. Marco Metodológico

3.1. Tipo de Investigación

Esta investigación es de tipo aplicada, donde se efectuó un proceso de análisis en el que se detectaron las necesidades de los encargados de la Universidad Cenfotec, se verificó la importancia de la colaboración entre la universidad y el uso de procesos tecnológicos. Como tema a resolver, es una herramienta de inteligencia de negocios y como único beneficiario se tendría a los integrantes del centro educativo.

3.2. Alcance Investigativo

Esta investigación posee un alcance exploratorio, donde se cuenta con las bases para estructurar el Data Warehouse, pero es requerido conocer nueva información y procesos desconocidos para los analistas que se manejan en la Universidad Cenfotec. Es requerido contemplar sesiones para comprender los procesos e información que se maneja. El proceso de análisis, desarrollo, pruebas e implementación de este está establecido por el cronograma del proyecto (Apéndice 1: Cronograma).

3.3. Enfoque

Para este proyecto se utilizó un enfoque alternativo, debido a que lo cuantitativo y lo cualitativo nunca han estado separados. Para el abordaje alternativo se propusieron sus tres dimensiones de la siguiente manera:

Ontológica: Actualmente se cuenta con el personal capacitado en el sistema transaccional sobre el cual se creará el Data Warehouse, quienes manejan estos datos

a diario. Además, se brinda un diccionario de datos para validar con mayor detalle la funcionalidad de la base de datos, lo que permite realizar la investigación.

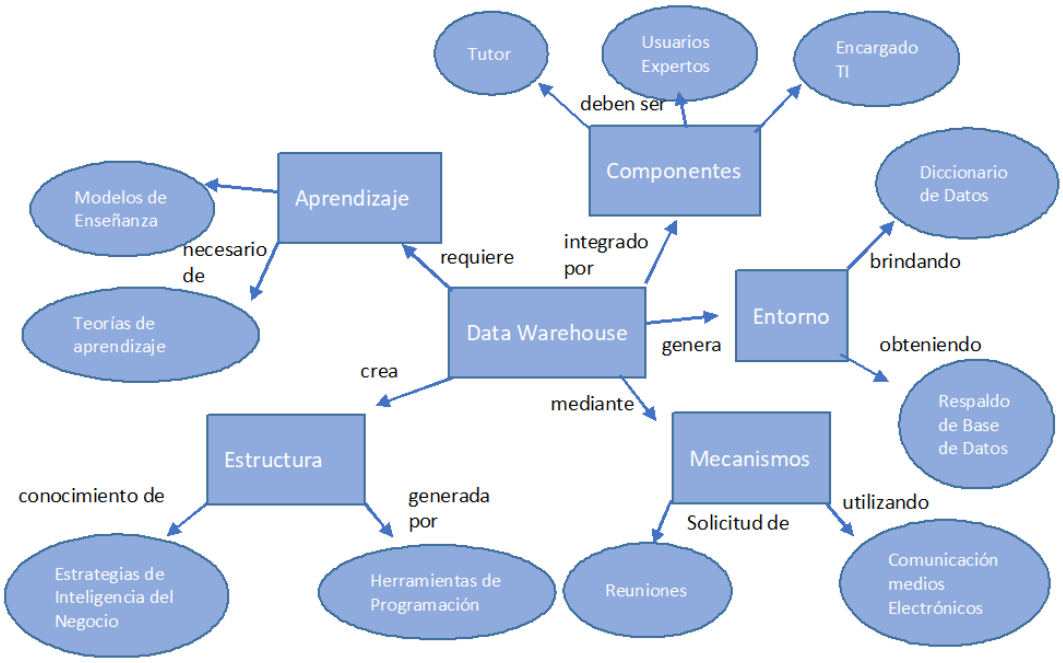


Ilustración 3 Dimensión Ontológica. Fuente: Elaboración Propia.

Epistemológica: Inicialmente, para efectuar este proyecto, los desarrolladores tomaron un rol de observadores para verificar funcionalidades y procesos que ejecuten en el transaccional, así como realizar consultas y validar necesidades del negocio; adicionalmente ver a detalle el funcionamiento del proceso de matrícula y el proceso del estudiante desde que inicia esta. Una vez finalizada la etapa de análisis los involucrados tomarán un rol más activo en el proceso de generación del modelo, así como el desarrollo e implementación del producto final.

Axiológica: Establecer los criterios de aceptación del proyecto desde la toma de requerimientos, así una vez iniciado el proceso de desarrollo y de pruebas unitarias se

podrá validar el cumplimiento de estos criterios y establecer rúbricas de aceptación para entregar el producto final con la mayor calidad posible. La facilidad de implementación, por otra parte, será evaluada de acuerdo con la escala de valores mostrada en la Tabla 1.

Nombre	**%**	**Satisfactorio 10-8**	**Satisfactorio con Recomendaciones 7-5**	**Necesita Mejorar 4-1**
Pruebas Unitarias.	15	Se validan cada uno de los componentes de una solución en un ambiente de desarrollo. Los elementos más críticos como lógica, ETL, reglas del negocio y cálculos en capa OLAP reflejan resultados satisfactorios.	Los procesos finalizan correctamente en ambiente de desarrollo, pero se encuentran procesos que podrían mejorarse en cuestiones de desarrollo y estructura.	Se presentan inconsistencias en el ambiente de desarrollo, existen procesos que no finalizan correctamente o información que no se procesa debidamente.
Pruebas del Sistema de Integración.	15	Se realiza la implementación satisfactoriamente en un ambiente de pruebas, los procesos se generan satisfactoriamente.	Los procesos finalizan correctamente en el ambiente de integración, pero se requieren ajustes para mejorar su eficiencia y procesamiento.	Se encuentran inconsistencias de compatibilidad en el sitio de pruebas, se deben realizar mejoras.
Pruebas de Validación de Datos.	15	Los datos resultantes reflejan la granularidad e integridad necesaria para su puesta en producción.	Los datos no presentan inconsistencias, pero se requiere un mayor grado de granularidad de los datos.	No se refleja la integridad de los datos requerida.
Pruebas de Aceptación de Usuario	20	Los procesos y la utilización de la herramienta se generan correctamente y	Los procesos se generan satisfactoriamente, pero la información obtenida no le	Se requieren mejoras en el modelo para poder interpretar los datos de una

		permite a los usuarios generar tareas con la mayor eficiencia.	brinda al negocio los resultados requeridos.	mejor manera.
Pruebas de Rendimiento.	15	Los procesos y extracción de información se generan con la mayor eficiencia posible.	Los procesos finalizan correctamente, pero se presentan problemas de rendimiento o exceso de uso de los recursos.	El rendimiento no concuerda con lo esperado y ocasiona lentitud anormal en los servidores de base de datos.
Pruebas de Regresión	10	Se ejecuta una última prueba para confirmar resultados satisfactorios de los procesos.	Se detectan pequeños fallos en cuanto rendimiento o información en su ejecución, pero no interrumpen la operativa.	Se presentan inconvenientes de último momento en la generación de procesos.

Ilustración 4 Criterios de evaluación de la facilidad de implementación.

3.4. Diseño

Principalmente se utilizó un diseño evaluativo al ser un proyecto práctico y ya que se cuenta con las herramientas para efectuar las respectivas pruebas. Inicialmente se formó un proceso de reuniones con los usuarios expertos y encargados del departamento de tecnologías de información del centro educativo, así como administradores del sistema transaccional y de base de datos, para comprender los principales procesos de la gestión de matrícula universitaria, así como validar las necesidades del proyecto por desarrollar.

3.5. Instrumentos de Recolección de Datos

En este caso utilizamos el método de Grupo Focal y Quanta Dimensional. Nos enfocaremos en los usuarios que necesitan el sistema y la gerencia de la universidad.

Dimensión	Ejemplo
Funcionamiento Actual	¿Cuál es el proceso actual de inteligencia de negocios sobre la plataforma de matrícula?
Respaldos de información	¿Cuentan con respaldos de la base de datos con los que podamos trabajar?
Sectores de importancia	¿Cuáles creen que son los sectores más importantes en los que deberíamos de enfocarnos?
Infraestructura de implementación	¿Cuál es la infraestructura con la que se cuenta actualmente?

Ilustración 5 Instrumentos de Recolección de Datos.

3.6. Técnicas de Análisis de la Información

Para el análisis de la información utilizaremos el modelo de marco conceptual. Mediante el marco conceptual podremos visualizar de una manera clara y concisa la composición y estructura del modelo de Data Warehouse por implementar.

3.7. Estrategia de Desarrollo de la Propuesta

Primeramente, con los requerimientos ya entendidos y depurados se procedió al desarrollo del esquema de la estrella, el cual fue la base para el producto final. Para el análisis se contó con la contribución de usuarios expertos del sistema y proceso de matrícula de la universidad, así como un encargado de la parte de tecnologías de información que permitió brindar a detalle las herramientas y características con que cuentan los equipos de la universidad para poder desarrollar este proyecto. Se procedió con la generación de la estrella, la cual cuenta con sus dimensiones y tabla de hechos, tomando en cuenta todos los factores que le permiten al negocio solucionar sus problemas y necesidades en el proceso de matrícula y estudio de los estudiantes afiliados al centro educativo.

Una vez finalizado este paso, se procedió con el desarrollo de la base de datos intermedia, también conocida como Staging Area (SA), y el desarrollo del depósito de datos en sí, el cual sirve como el Data Warehouse. Estos procesos se realizaron mediante la extracción e inserción de información en procesos que se generan de manera automática, utilizando tareas programadas que los generen.

Una vez terminados los pasos anteriores, para la visualización de datos se utilizó un Cubo (OLAP) para que los datos sean mucho más amigables con el usuario. Para el desarrollo de todos los puntos del proyecto utilizaremos herramientas Microsoft (SSIS, SSAS, MSSQL Server). Luego de estos procesos se procedió con la generación tareas programadas que ejecutan estos procesos de carga, y su debida implementación en un ambiente de producción que permite a los usuarios responsables acceder a esta información y generar sus debidos reportes.

4. Análisis del Diagnóstico

La Universidad Cenfotec cuenta con el uso de una herramienta llamada Class, que le permite realizar toda la gestión financiera, administrativa y de matrícula del centro educativo. En ella los procesos se gestionaban con el pasar de los días, no se contaba con una herramienta que les permitiera evaluar la información que manejaba en esos procesos y poder comprender de una mejor manera cómo mejorar sus decisiones administrativas. Los reportes del módulo de matrícula se manejaban de manera manual mediante la elaboración de Excel, no existía una herramienta capaz de poseer la flexibilidad de brindar estos datos y con ellos sacar conclusiones de las decisiones de la universidad. Mediante este proyecto, se puede empezar a desarrollar una cultura de datos dentro de la universidad.

En conjunto con los encargados de la Universidad Cenfotec, parte financiera y responsables de tecnologías de informacion, se generaron una serie de reuniones para plantear una solución de inteligencia del negocio, se fijó cuáles serían los objetivos de dicho proyecto, informacion de interés, la fuente de datos que es la informacion almacenada en la base de datos del sistema Class, así como el diccionario de datos de este. Nos brindaron además una serie de necesidades del negocio que debía sustentar el proyecto de inteligencia de negocios.

Una vez analizada la informacion por los técnicos, pruebas de consultas al respaldo de base de datos y descifrado de la estructura de la base de datos, se fueron estructurando modelos hasta llegar a un modelo que permitiera satisfacer las consultas y estudios necesitados por el negocio. Mediante la herramienta Microsoft SQL Server y sus herramientas se plantearon dos nuevas bases de datos conocidas como Staging

Area y Data Warehouse, la primera permite extraer únicamente las tablas que se abarcaron en el modelo y la segunda es encargada de almacenar el modelo de inteligencia de negocios, tanto las dimensiones y tabla de hechos.

Se utilizó Integration Services para idear los procesos de cargas, filtrado de datos de las dos bases de datos generadas. Para el Staging Area se eliminan las llaves primarias y se identifican únicamente tablas que se vayan a involucrar con el módulo de matrícula, el proceso de Integration Services se encarga de extraer esa información y llenar la base de datos, evitando datos duplicados, además se generó un proceso de identificación de llaves primarias ya que muchas tablas de relevancia no contaban con dicho identificador. Una vez procesados los datos en el Staging Area se forma el modelo estrella del Data Warehouse, donde todas las dimensiones del módulo de matrícula se enlazan en la tabla de hechos para así poder comunicar todo el modelo y generar relaciones entre reportes.

Por último, con la herramienta Analysis Services para el modelado de la base de datos multidimensional se establece un mecanismo de conexión para los usuarios, y poder generar así sus reportes de manera flexible y eficiente. La herramienta Microsoft Power BI se conecta directamente a la solución multidimensional y permite visualizar la información.

5. Propuesta de Solución

5.1. Definición de Recursos

Para poder llevar a cabo este proyecto se contó con dos estudiantes de ingeniería en sistemas de la Universidad Cenfotec, los cuales lo llevaron a cabo este proyecto mediante su ingenio y creatividad. Además, se contó con la colaboración de todos aquellos profesores e integrantes del centro universitario que de alguna u otra manera influenciaron con sus conocimientos y experiencias para poder llevar a cabo este proyecto.

En la siguiente tabla se reflejan los roles y responsabilidades que se estructuraron para este proyecto:

Rol	Responsabilidad
Encargado de TI	Gestionar las herramientas de desarrollo y de implementación del producto.
Encargado de Financiero	Validar los requerimientos y alcance del proyecto. Gestionar pruebas funcionales del producto.
Contraparte del Proyecto	Establecer comunicación entre las entidades de la Universidad Cenfotec y los analistas del proyecto.
Líder Técnico	Velar por que todo el proyecto vaya acorde a lo esperado por el negocio.
Analistas y Desarrolladores	Encargado del análisis y desarrollo del producto.

Ilustración 6 Roles y Responsabilidades

5.2. Análisis de Fuente de Datos

Como fuente de datos de entrada se cuenta con la base de datos del sistema transaccional CLASS, el cual es utilizado actualmente en la Universidad Cenfotec para la manipulación, consulta y transformación de los datos. Este es el sistema de administración académica que colabora con las instituciones educativas en la optimización de todos los recursos y servicios de su gestión diaria, con una integración transparente entre sus modulaciones.

5.3. Herramientas Por Utilizar

5.3.1. Sistema Gestor de Base de Datos

Como motor de base de datos se utilizó el mismo en el que se encuentra desarrollada la base de datos del sistema Class, la cual es Microsoft SQL Server en su versión 2012.

5.3.2. Microsoft SQL Server Integration Services 2012

Se desarrollaron dos proyectos mediante esta herramienta, los cuales corresponden al ETL que se encarga de la creación de la base de datos del Staging Area y los procesos de carga que realizan la generación del Data Warehouse. Estos también proporcionan las funcionalidades necesarias para procesar cargas masivas de datos de inteligencia de negocios y simplificar las tareas.

5.3.3. Microsoft SQL Server Analysis Services 2012

Mediante esta aplicación se desarrolló del proceso de generación de base de datos multidimensional, que permite realizar las conexiones necesarias a los usuarios y

así generar sus propios reportes genéricos y sacar conclusiones de lo evidenciado en la información.

5.4. Arquitectura de la Solución

5.4.1. Arquitectura de Hardware

En la siguiente imagen se refleja el diagrama de la arquitectura de hardware sobre el cual se planteó el proyecto.

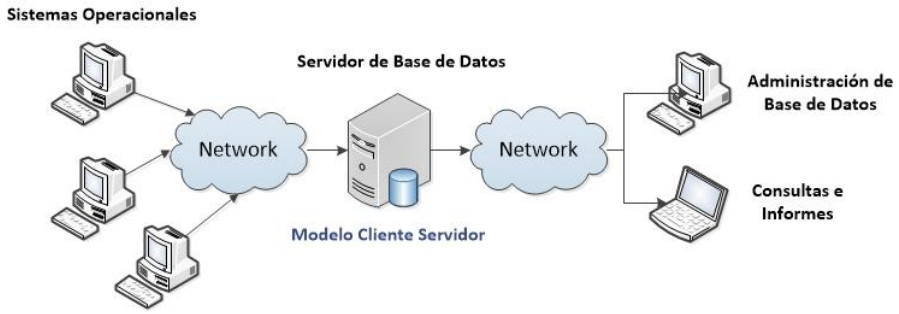


Ilustración 7 Arquitectura de Hardware. Elaboración Propia.

Los diferentes procesos realizados mediante la herramienta Class brindan una diversidad de información, la cual es almacenada en la base de datos del transaccional.

Los usuarios que utilicen el Data Warehouse podrán acceder al servidor para realizar las respectivas consultas de una información que ha pasado por un proceso de cambio y se administró de tal manera que el negocio pueda utilizarla para su gestión administrativa. Desde una computadora con roles de administrador se podrán realizar los procesos de carga y mantenimiento de todos los componentes que pertenecen a este modelo de inteligencia de negocios, así mantener la información en lo posible actualizada para su gestión.

5.4.2. Arquitectura de Software

En la siguiente imagen se visualiza la arquitectura del software utilizada y propuesta para el diseño y la explotación de este proyecto.

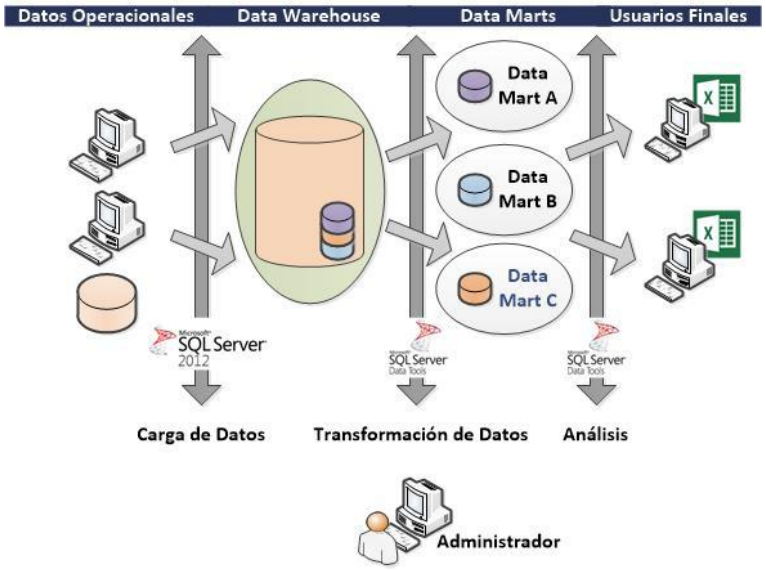


Ilustración 8 Arquitectura de Software. Fuente: Elaboración Propia.

Para la arquitectura propuesta se inicia desde los procesos realizados por los usuarios de la herramienta Class, es ahí donde un usuario administrador de los procesos de inteligencia de negocios utilizará Microsoft SQL Server 2012, específicamente la herramienta Data Tools, para que le permita a este realizar las cargas respectivas y almacenar la información en el Staging Area. Luego de eso debe ejecutar los procesos de Data Warehouse para actualizar las dimensiones, y mediante la base de datos multidimensional realizada con Analysis Service los usuarios acceden desde MS Excel o Power BI a generar los debidos reportes e informes.

5.4.3. Estándares de Nomenclatura de Base de Datos Transaccional

Objeto	Estándar	Ejemplo
Base de Datos	<NOMBREBD>	CENFOTEC
Tabla	<NOMBRETABLA>	Dbo.ADMISIONES
Columna	<PREFIJODESCRIPCION>	ADMNIV
Llave Primaria	<PKNOMBRE>	PK_ADMISION
Índice	<index_NOMBRE>	Index_ADMISIONES
Constraint	<DF_NOMBRE>	DF_NOTIFICACIONES
Vista	<ViewNombre>	ViewBecados

Ilustración 9 Estándares de Nomenclatura de Base de Datos Transaccional.

5.4.4. Estándar de nomenclatura de Staging Area

Objeto	Estándar	Ejemplo
Base de Datos	<NOMBRE_SA>	CENFOTEC_SA
Tabla	<NOMBRETABLA>	Dbo.ADMISIONES
Columna	<PREFIJODESCRIPCION>	ADMNIV

Ilustración 10 Estándar de nomenclatura de Staging Area.

5.4.5. Estándar de nomenclatura de Data Warehouse

Objeto	Estándar	Ejemplo
Base de Datos	<NOMBRE_DW>	CENFOTEC_DW
Dimensión	<TD_NOMBRE>	TD_CURSO
Tabla de Hechos	<TH_NOMBRE>	TH_MATRICULA

Columna	<NombreDescripcion>	CursoCodigo
Llave Primaria	<PK_NOMBRETABLA>	PK_TD_CURSO
Llave Foránea	<FK_NombreTabla_NombreTabla>	FK_TH_Matricula_TD_CURSO

Ilustración 11 Estándar de nomenclatura de Data Warehouse.

5.4.6. Estándar de nomenclatura de tareas programadas

Objeto	Estándar	Ejemplo
Job	<J_Nombre>	J_StagingArea

Ilustración 12 Estándar de nomenclatura de tareas programadas.

5.4.7. Estándar de nomenclatura de Integration Services

Objeto	Estándar	Ejemplo
SSIS Package	<P_Nombre>	P_StagingArea
Sequence Container	<SC_Nombre>	SC_StagingArea_Cenfotec
Connection Manager	<NOMBREBD>	CENFOTEC_SA
Execute SQL Task	<EST_NOMBRE>	EST_DEFAULT
Data Flow Task	<DFT_NOMBRE>	DFT_CURSOS
OLE DB Destination	<D_NOMBRE>	D_CURSOS
OLE DB Source	<S_NOMBRE>	S_CURSOS
Slowly Changing Dimension	<SCD_NOMBRE>	SCD_TD_CURSO
OLE DB Command	<C_NOMBRE>	C_TD_CURSO
Insert Destination	<ID_NOMBRE>	ID_TD_CURSO
Data Conversion	<DC_NOMBRE>	DC_TD_CURSO

Analysis Services Processing Task	<ASPT_NOMBRE>	ASPT_MULTIDIMENSIONAL

Ilustración 13 Estándar de nomenclatura de Integration Services.

5.4.8. Estándar de nomenclatura de Analysis Services

Objeto	Estándar	Ejemplo
Data Source View	<DSV_NOMBRE >	DSV_MATRICULA
Cube	<CB_NOMBRE>	CB_MATRICULA
Jerarquías	<J_NOMBRE>	J_FECHA

Ilustración 14 Estándar de nomenclatura de Analysis Services.

5.5. Etapa de diseño y desarrollo

5.5.1. Diseño y desarrollo del Staging Area

El Staging Area toma como fuente de datos la información registrada en los procesos realizados en la Universidad CENFOTEC, específicamente en el sistema transaccional Class, que permite realizar la gestión académica. La idea de esta solución es obtener las tablas de dicha base de datos y extraer únicamente las relacionadas con el módulo de matrícula, que por un lado permita facilitar la extracción de esta información, para así reducir considerablemente en rendimiento y eficiencia los procesos de ETL. Además, que la extracción y análisis de datos no se realice directamente de la base de datos transaccional para no afectar el rendimiento y operativa de los usuarios del sistema como tal.

5.5.2. Diseño de arquitectura del Staging Area

La imagen adjunta refleja las tablas por las que está constituido el Staging Area, este no posee relaciones entre sí, debido a que una de las intenciones de esta base de datos es obtener procesos de escritura rápidos y desconectarse pronto del transaccional. Por otro lado, a este nivel no se realizó ningún tipo de homologación en los nombres de las tablas y tampoco de los campos de estas, lo que indica que seguimos a un nivel técnico de entendimiento. No es sino hasta el Data Warehouse y la base de datos multidimensional donde se modifican los nombres debido a que es ya un nivel donde el usuario requiere entendimiento y utilización de los datos procesados.

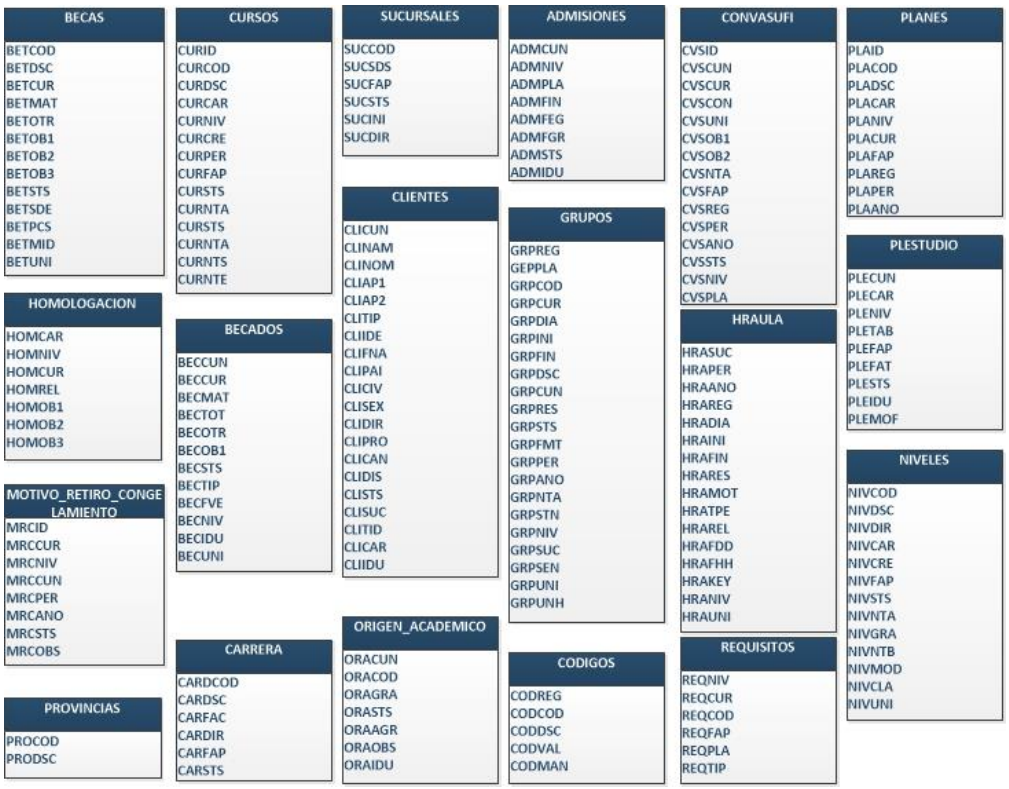


Ilustración 15 Diseño de arquitectura del Staging Area.

5.5.3. Definición de tablas del Staging Area

Para la elaboración del Staging Area se utilizó un total de 19 tablas de la base de datos transaccional, las cuales contienen la siguiente información:

- ADMISIONES: Registra los estudiantes admitidos a un plan de la universidad.
- BECADOS: Almacena los registros de estudiantes ligados a una beca.
- BECA: Contiene el catálogo de becas que ofrece el centro educativo.
- CARERRA: Registra el catálogo de carreras de la universidad.
- CLIENTES: Almacena los integrantes que se registran en la universidad.
- CÓDIGOS: Almacena el catálogo de códigos del sistema.
- CONVASUFI: Registra los estudiantes que realizaron el proceso de convalidación de materias en la universidad.
- CURSOS: Registra el catálogo de cursos que maneja la universidad.
- GRUPOS: Contiene el registro de estudiantes ligados a un curso junto con la nota final asignada.
- HOMOLOGACIÓN: Catálogo de cursos homologados, que se contemplan en más de un nivel.
- HRAULA: Registro del detalle del grupo, información general del estudiante en el curso.
- MOTIVO_RETIRO_CONGELAMIENTO: Registra los estudiantes que realizaron el retiro o congelamiento de uno o más cursos.
- NIVELES: Almacena el catálogo de niveles disponibles en la universidad.
- PLANES: Registro de planes existentes en el centro educativo.

- PLESTUDIO: Asocia los registros de clientes a uno o más planes.
- PROVINCIAS: Catálogo de provincias.
- REQUISITOS: Contiene los requisitos para necesarios para matricular un curso en específico.
- SUCURSALES: Catálogo de sucursales que posee el centro educativo.

5.5.4. Estrategia de ETL del Staging Area

A continuación, se explicará el proceso de creación del Staging Area:

Se realizó una verificación de todas las tablas de la base de datos transaccional, en conjunto con los encargados de la Universidad Cenfotec, esto para validar cuáles tablas pertenecen al módulo de matrícula universitaria.

Una vez que se delimitan las tablas se extraen los scripts de la creación de estas, y se realiza una pequeña limpieza de campos que no estarán en el modelo o no se relacionan con la matrícula universitaria, por lo cual para ahorro de espacio y rendimiento se descartan dentro del modelo.

Se efectúa la eliminación de llaves primarias, llaves foráneas de las tablas seleccionadas y se implementan los campos restantes en el Staging Area con su mismo nombre y tipo de dato. Al final del análisis el Staging Area quedó conformado por un total de 19 tablas.

```
CREATE TABLE [dbo].[HRAULA](
    [HRABNK] [nvarchar](2) NULL,
    [HRASUC] [nvarchar](3) NULL,
    [HRAPER] [nvarchar](2) NULL,
    [HRAANO] [nvarchar](4) NULL,
    [HRAREG] [nvarchar](1) NULL,
    [HRAAUL] [nvarchar](10) NULL,
    [HRAFAC] [nvarchar](80) NULL,
    [HRADIA] [nvarchar](10) NULL,
    [HRAINI] [nvarchar](13) NULL,
    [HRAFIN] [nvarchar](13) NULL,
    [HRARES] [nvarchar](20) NULL,
    [HRAMOT] [nvarchar](10) NULL,
    [HRASEC] [numeric](1, 0) NULL,
    [HRATPE] [nvarchar](15) NULL,
    [HRAREL] [nvarchar](15) NULL,
    [HRAFDD] [datetime] NULL,
    [HRAFHH] [datetime] NULL,
    [HRAFRQ] [datetime] NULL,
    [HRATIP] [nvarchar](1) NULL,
    [HRAVIR] [nvarchar](1) NULL,
    [HRAMOD] [nvarchar](15) NULL,
    [HRAUNI] [bigint] IDENTITY(1,1) NOT NULL PRIMARY KEY,
    [HRAHPR] [numeric](7, 2) NULL,
    [HRAUNR] [bigint] NULL) ON [PRIMARY]
GO
```

```
CREATE TABLE [dbo].[HRAULA](
    [HRASUC] [nvarchar](3) NULL,
    [HRAPER] [nvarchar](2) NULL,
    [HRAANO] [nvarchar](4) NULL,
    [HRAREG] [nvarchar](1) NULL,
    [HRADIA] [nvarchar](10) NULL,
    [HRAINI] [nvarchar](13) NULL,
    [HRAFIN] [nvarchar](13) NULL,
    [HRARES] [nvarchar](20) NULL,
    [HRAMOT] [nvarchar](10) NULL,
    [HRATPE] [nvarchar](15) NULL,
    [HRAREL] [nvarchar](15) NULL,
    [HRAFDD] [datetime] NULL,
    [HRAFHH] [datetime] NULL,
    [HRAKEY] [nvarchar](15) NULL,
    [HRANIV] [nvarchar](10) NULL,
    [HRAUNI] [bigint] NOT NULL,
) ON [PRIMARY]

GO
```

Ilustración 16 Estrategia del ETL del Staging Area 1.

En la imagen se logra observar cómo se redujeron las columnas de la tabla HRAULA, tomando en cuenta columnas que tuvieran gran relevancia en el modelo a futuro, y además se procedió con la eliminación de llaves primarias y foráneas de las tablas en su estructura. Esto para delimitar y extraer únicamente lo requerido por el negocio.

Se creó un nuevo proyecto con la herramienta SQL Server Integration Services 2012, el cual se encarga del proceso de generación de cargas a la base de datos CENFOTEC_SA. En el proyecto se crean 2 conexiones a la base de datos, una encargada de extraer la información de la base de datos transaccional y otra en la cual realizamos la creación del Staging Area.

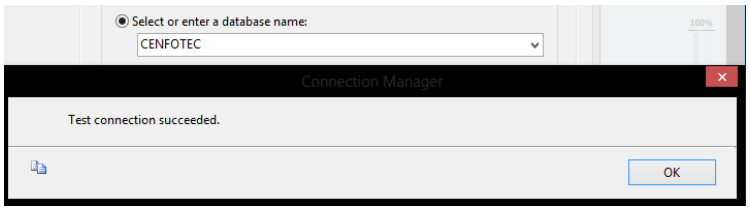


Ilustración 17 Estrategia del ETL del Staging Area 2.

Se ingresa el componente llamado SQL Task, donde como primer paso se efectúa una limpieza de las tablas de la base de datos CENFOTEC_SA para limpiar los registros e ingresar nuevamente los registros actualizados y evitar duplicidad de información.

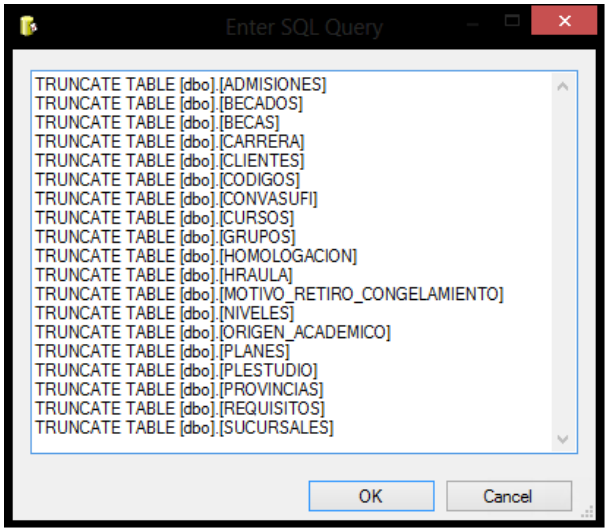


Ilustración 18 Estrategia del ETL del Staging Area 3.

El siguiente paso es referenciar cada tabla del CENFOTEC_SA mediante el componente Data Flow Task, e ir implementando el orden en que se efectuará la carga de tablas. En este paso ya tenemos contempladas las 19 tablas del Staging Area y las que serán donde almacenaremos la información relevante.



Ilustración 19 Estrategia del ETL del Staging Area 4.

Dentro de cada Data Flow Task se ingresan 2 procesos que corresponden al proceso en que se extraen todos los datos de la base de datos transaccional para una tabla en específico (OLE DB Source), y luego el proceso donde esta información extraída es insertada en la tabla del Staging Area (OLE DB Destination).



Ilustración 20 Estrategia del ETL del Staging Area 5.

Al entrar al OLE DB Source se ingresó la tabla indicada por el Data Flow Task según se realizó su nombramiento. Se ingresa como Connection manager a la base de datos transaccional, de la cual vamos a extraer la información y depositarla en el Staging Area. Como parte de buenas prácticas se especifican únicamente las columnas

que vamos a requerir en el Staging Area, por cuestiones de mapeo de columnas y rendimiento, ya que si reducimos en lo posible la cantidad de columnas estaremos mejorando en cuestiones de tiempo, rendimiento y tamaño de crecimiento del Staging Area en memoria.

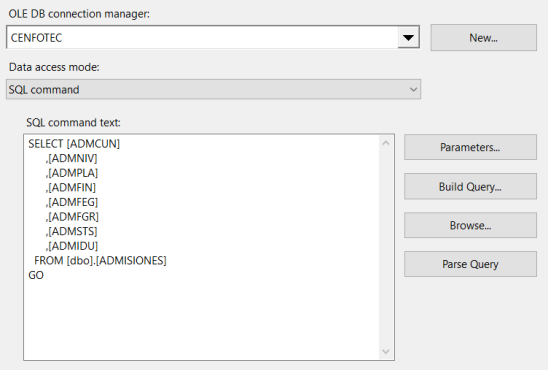


Ilustración 21 Estrategia del ETL del Staging Area 6.

Cuando se ingresó el OLE DB Destination, se selecciona la tabla que corresponda a la indicada en el Source, siempre y cuando corresponda a su misma versión, pero en la base de datos CENFOTEC_SA; esto en cuanto estructura y tipo de dato de las tablas tanto del transaccional como del Staging Area. Estamos seleccionando dónde deseamos que se ingresen los datos que extrajimos del transaccional.

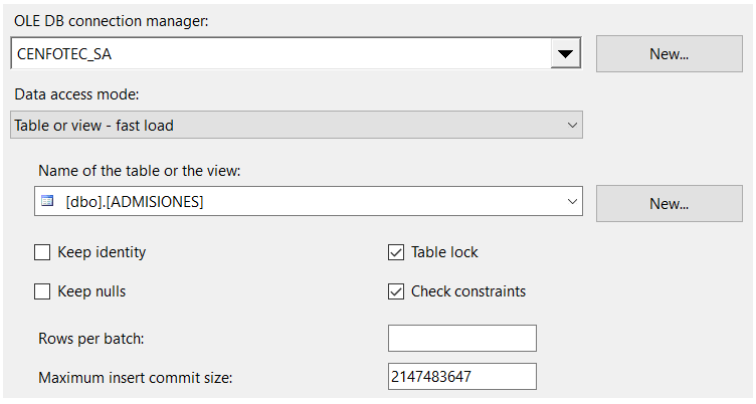
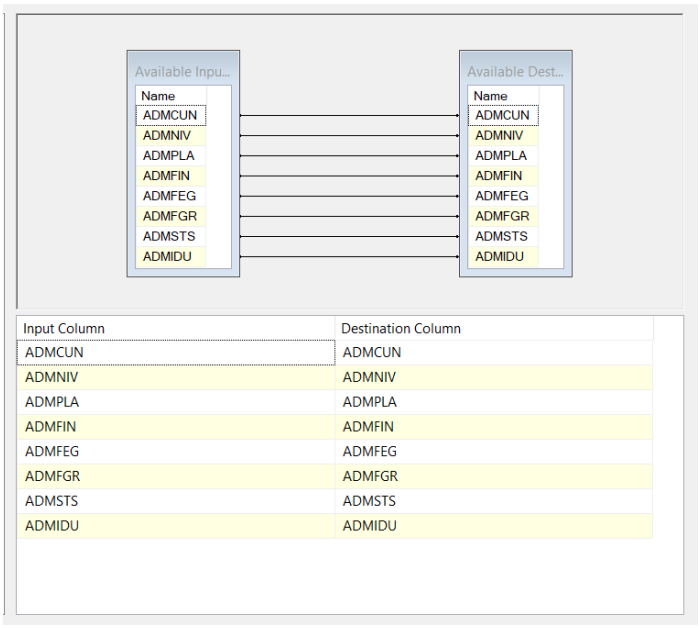


Ilustración 22 Estrategia del ETL del Staging Area 7.

Se realiza el mapeo en el destino para verificar que los campos asociados poseen compatibilidad, para el proyecto se utilizaron campos de mismo nombre y tipo de dato para pasar los datos al Staging Area.



Input Column	Destination Column
ADMCUN	ADMCUN
ADMNIV	ADMNIV
ADMPLA	ADMPLA
ADMFIN	ADMFIN
ADMFEG	ADMFEG
ADMFGR	ADMFGR
ADMSTS	ADMSTS
ADMIDU	ADMIDU

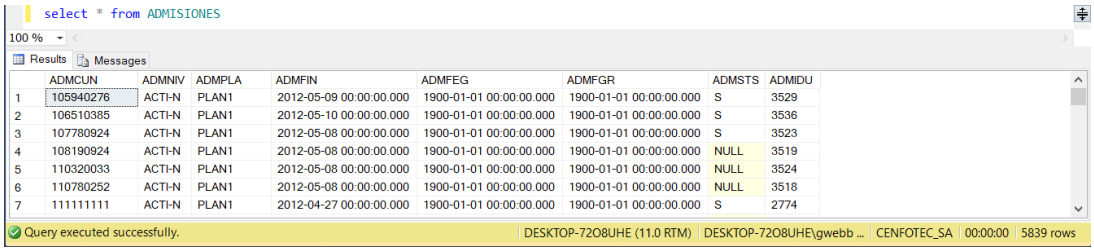
Ilustración 23 Estrategia del ETL del Staging Area 8.

El proceso se realizó para todos los Data Flow Task insertados. Una vez que se fijaron los orígenes y destinos de cada uno, se procede a ejecutar el proyecto.



Ilustración 24 Estrategia del ETL del Staging Area 9.

Al finalizar este proceso, se verifica que las tablas de la base de datos CENFOTEC_SA se encuentren con la misma información de la base de datos del transaccional.



```
select * from ADMISIONES
```

100 %

Results | Messages

	ADMCUN	ADMNIV	ADMPLA	ADMFIN	ADMFEG	ADMFGR	ADMSTS	ADMIDU
1	105940276	ACTI-N	PLAN1	2012-05-09 00:00:00.000	1900-01-01 00:00:00.000	1900-01-01 00:00:00.000	S	3529
2	106510385	ACTI-N	PLAN1	2012-05-10 00:00:00.000	1900-01-01 00:00:00.000	1900-01-01 00:00:00.000	S	3536
3	107780924	ACTI-N	PLAN1	2012-05-08 00:00:00.000	1900-01-01 00:00:00.000	1900-01-01 00:00:00.000	S	3523
4	108190924	ACTI-N	PLAN1	2012-05-08 00:00:00.000	1900-01-01 00:00:00.000	1900-01-01 00:00:00.000	NULL	3519
5	110320033	ACTI-N	PLAN1	2012-05-08 00:00:00.000	1900-01-01 00:00:00.000	1900-01-01 00:00:00.000	NULL	3524
6	110780252	ACTI-N	PLAN1	2012-05-08 00:00:00.000	1900-01-01 00:00:00.000	1900-01-01 00:00:00.000	NULL	3518
7	111111111	ACTI-N	PLAN1	2012-04-27 00:00:00.000	1900-01-01 00:00:00.000	1900-01-01 00:00:00.000	S	2774

Query executed successfully. | DESKTOP-72O8UHE (11.0 RTM) | DESKTOP-72O8UHE\gwebb ... | CENFOTEC_SA | 00:00:00 | 5839 rows

Ilustración 25 Estrategia del ETL del Staging Area 10.

El Staging Area traslada toda la información de la tabla, sin importar si un dato es nulo o no, lo único que no conserva son llaves dentro de la estructura de tablas, el tipo de dato de cada columna se mantiene. Además, se realizó una considerable

disminución de columnas para conservar únicamente las relevantes para el módulo de matrícula.

5.5.5. Consideraciones de la creación del Staging Area

Para la creación del Staging Area de la base de datos de la Universidad Cenfotec fue necesario tomar en cuenta las siguientes consideraciones:

- Para la creación del Staging Area se tomaron en cuenta únicamente tablas relacionadas al módulo de matrícula universitaria, esto para enfocar el proyecto únicamente en el módulo de interés según los objetivos fijados en el proyecto. Dichas tablas fueron seleccionadas de manera cuidadosa y luego de un profundo análisis, el cual se dividió en tres etapas: primeramente, una definición del problema, alcances y necesidades del negocio, donde se estableció que la prioridad era la matrícula del centro educativo. Una vez finalizada esta se procedió al estudio riguroso con los usuarios y encargados de la universidad para validar, conforme al diccionario de datos, cuáles eran los campos involucrados en el módulo. Por último, a nivel técnico se reestructuró el módulo de matrícula con las tablas seleccionadas anteriormente y se establecieron enlaces y se hizo una selección rigurosa de las columnas de interés.
- En la creación de tablas del Staging Area no se tomaron en cuenta algunas columnas de las tablas para disminuir el tamaño de las tablas, además mejorar el proceso de carga en cuanto tiempo y rendimiento.
- En las tablas de: CURSOS, CONVASUFI y PLANES fue requerido implementar un identificador de tabla, ya que estas tablas en la base de

datos transaccional no contaban con una llave primaria que permitiera diferenciar una columna de la otra. Como se refleja en la imagen, para crear esta llave se realizó un tipo de dato carácter, el cual se encuentra constituido de ciertos campos que le permiten a cada registro diferenciarse uno del otro.

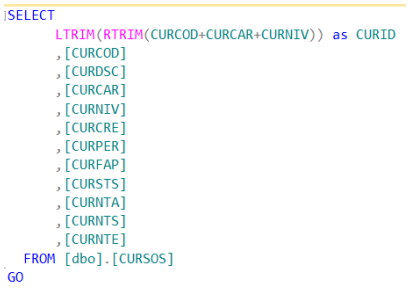


Ilustración 26 Estrategia del ETL del Staging Area 11.

- Para la tabla MOTIVO_RETIRO_CONGELAMIENTO también fue requerido realizar una llave compuesta, la diferencia es que esta aun permitía la duplicidad de informacion ya que existían registros que se encontraban más de una vez. Por ello se desarrolló un proceso en donde el sistema valide si el registro ya existe para que no lo vuelva a insertar, pero si ya existe y el primer registro no contaba con el campo de observaciones lleno, el proceso realiza la actualización del campo conforme a las observaciones del registro duplicado.

```
WHILE (@cantReg>0)
begin
--SELECCIONA UN CONSECUTIVO PARA EVALUAR
SET @seleccionadoContador=(select top 1 Consecutivo FROM @motivoRetiro1)
SET @seleccionado=( SELECT TOP 1 ID from @motivoRetiro1 WHERE Consecutivo=@seleccionadoContador)
--VERIFICA SI EXISTE YA EL ID
SET @existeTabla2=(SELECT COUNT(ID) FROM @motivoRetiro2 where ID=@seleccionado )
IF(@existeTabla2>0)
BEGIN
--ACTUALIZA SI EL ID INSERTADO NO POSEIA OBSERVACIONES
SET @existeObsTabla1=(SELECT count(OBSERVACION) FROM @motivoRetiro1 WHERE ID=@seleccionado AND OBSERVACION<>'')
--actualiza observacion para tenerla
IF(@existeObsTabla1>0)
BEGIN
SET @obsTabla1=(SELECT TOP 1 OBSERVACION FROM @motivoRetiro1 WHERE ID=@seleccionado AND OBSERVACION<>'')
 UPDATE @motivoRetiro2 SET OBSERVACION=@obsTabla1 WHERE ID=@seleccionado
END
END
ELSE
BEGIN
--INSERTA EL VALOR NO DUPLICADO
 INSERT INTO @motivoRetiro2 (ID,CODCURSO,CODNIVEL,CIFESTUDIANTE,PERIODO,ANO,ESTATUS,OBSERVACION)
 SELECT TOP 1
       ID,CODCURSO, CODNIVEL, CIFESTUDIANTE, PERIODO, ANO,ESTATUS,OBSERVACION
       FROM @motivoRetiro1  WHERE ID=@seleccionado
END
SET @cantReg=@cantReg-1
DELETE FROM @motivoRetiro1 WHERE Consecutivo=@seleccionadoContador
END
--SELECCIONA LA TABLA
SELECT
       ID AS MRCID,CODCURSO AS MRCCUR, CODNIVEL AS MRCNIV, CIFESTUDIANTE AS MRCCUN,
       PERIODO MRCPER, ANO AS MRCANO,ESTATUS AS MRCSTS,OBSERVACION AS MRCOBS
       FROM @motivoRetiro2
```

Ilustración 27 Estrategia del ETL del Staging Area 12.

5.5.6. Diseño y desarrollo del DataWarehouse

En la creación del Data Warehouse se utilizaron dos conexiones a base de datos, contamos con la conexión al Staging Area en la cual se encuentran únicamente las tablas relacionadas al módulo de matrícula universitaria y también existe la conexión a la base de datos de este Data Warehouse, que es donde se estableció el esquema principal del modelo de inteligencia de negocios, listo para ser consultado por los usuarios. Mediante un proceso de ETL se insertarán los datos en las debidas dimensiones para establecer un modelo estrella una vez conformada la tabla de hechos. Durante los procesos de ETL la información se extrae de una manera más estratégica, se enfocaron campos de suma importancia para el negocio referente a la matricula, en conjunto con los usuarios expertos y encargados del negocio se establecieron los campos que integran cada dimensión. Una vez seleccionados se realizan procesos de transformación para mejorar el rendimiento y tamaño de datos de las consultas.

Mediante el Data Warehouse los usuarios podrán acceder a información clave de los estudiantes en su vida estudiantil y facilitar la toma de decisiones de la organización. Además, permite una mayor flexibilidad y rapidez en el acceso de información al no ser esta consultada directamente de la base de datos transaccional, y la información se encuentra más robusta e íntegra que el Staging Area. Aquí entra en juego la estrategia del negocio y la información que este necesita para mejorar sus procesos y extraer únicamente lo que necesita.

5.5.7. Diseño de arquitectura del Data Warehouse

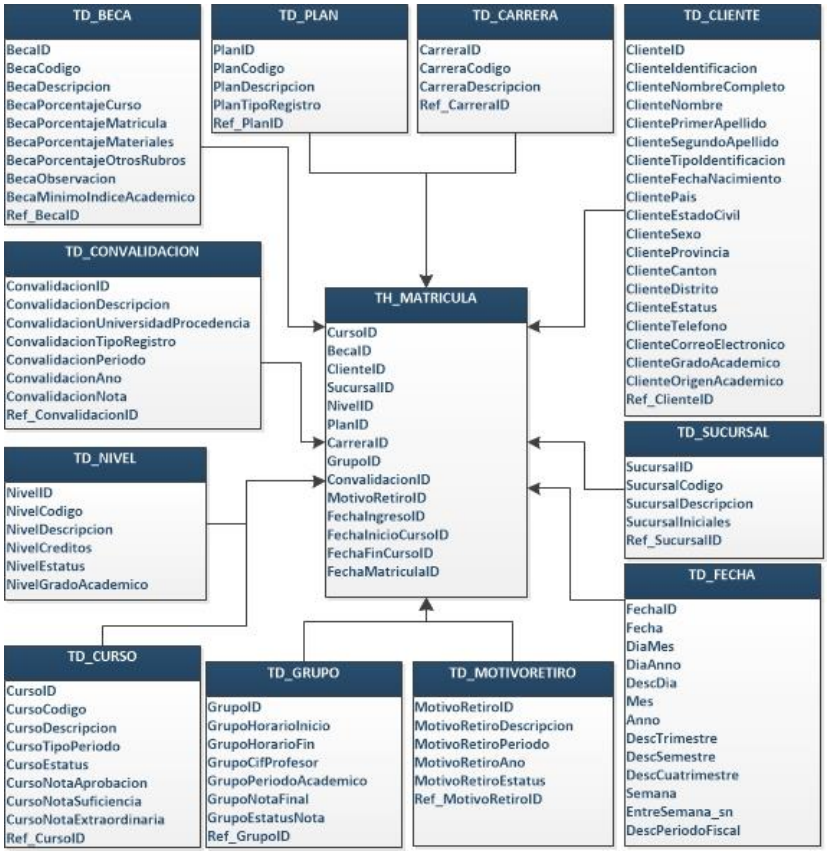


Ilustración 28 Diseño y desarrollo del Data Warehouse 1.

5.5.8. Definición de dimensiones y tabla de hechos

- TD_BECA: Dimensión que almacena el catálogo de becas de la universidad.
- TD_CARRERA: Dimensión que registra las carreras existentes en la universidad.
- TD_CLIENTE: Dimensión encargada de almacenar los registros de clientes de la universidad.
- TD_CONVALIDACION: Dimensión que registra las convalidaciones de cursos realizadas por estudiantes.
- TD_CURSO: Dimensión que almacena el catálogo de cursos del centro educativo.
- TD_FECHA: Dimensión de fechas, se relaciona con la fecha de inicio del curso, fecha fin del curso, matrícula y la fecha de ingreso del estudiante.
- TD_GRUPO: Dimensión que registra la información del grupo al que fue asociado el estudiante y registra la nota final del curso.
- TD_MOTIVORETIRO: Dimensión que almacena los motivos de retiro de los estudiantes a un curso.
- TD_NIVEL: Dimensión de catálogo de niveles de la universidad.
- TD_PLAN: Dimensión que almacena los planes que maneja la universidad.
- TD_SUCURSAL: Dimensión de sucursales existentes en la universidad.
- TH_MATRICULA: Tabla de hechos que mantiene las relaciones entre todas las dimensiones del módulo de inteligencia de negocios.

5.5.9. Estrategia de ETL para el Data Warehouse

En la creación de dimensiones se tomaron en cuenta todos los factores de relevancia para los usuarios y gestores de la Universidad Cenfotec, se realizó conforme reuniones con los encargados extracción de los puntos más relevantes de la base de datos transaccional. Se identificó qué secciones son las que necesita extraer la universidad ocupaban extraer para sacar conclusiones e información de inteligencia de negocios.

Una vez detectados los elementos de importancia para la universidad se procede con la estructura de la tabla de cada dimensión, las cuales poseen un valor tipo identidad que le brinda un identificador único al registro. Además, se incorporan los campos de referencia que establecen una relación con la llave de la tabla transaccional y por último los campos de interés para el negocio y poder realizar los respectivos análisis.

```
CREATE TABLE [dbo].[TD_CURSO](
    [CursoID] [int] IDENTITY(1,1) NOT NULL,
    [CursoCodigo] [varchar](10) NOT NULL,
    [CursoDescripcion] [varchar](200) NOT NULL,
    [CursoTipoPeriodo] [varchar](15) NOT NULL,
    [CursoEstatus] [varchar](5) NOT NULL,
    [CursoNotaAprobacion] [varchar](5)  NULL,
    [CursoNotaSuficiencia] [varchar](5)  NULL,
    [CursoNotaExtraordinaria] [varchar](5)  NULL,
    [ref_CursoID] [varchar](100) NOT NULL,
 CONSTRAINT [PK_TD_CURSO] PRIMARY KEY CLUSTERED
(
    [CursoID] ASC
)WITH (PAD_INDEX = OFF, STATISTICS_NORECOMPUTE = OFF, IGNORE_DUP_KEY = OFF, ALLOW_ROW_LOCKS = ON, ALLOW_PAGE_LOCKS = ON) ON [PRIMARY]
) ON [PRIMARY]

GO
```

Ilustración 28 Diseño y desarrollo del Data Warehouse 1.

En la validación de fechas relevantes para el modelo, se generó una única dimensión por cuestiones de espacio a nivel de la base de datos, la cual permite

extraer cualquier fecha que se encuentre registrada en el sistema. Esta dimensión se clasifica en el registro también por valores como el semestre, año, mes y cuatrimestre en que ocurrió el evento mediante un identificador. Para efectos del modelo las fechas seleccionadas son: fecha de ingreso, la cual corresponde a la fecha en que ingresó el estudiante al centro educativo, luego la fecha de inicio y fin del curso, y por último la fecha en que se realizó la matrícula de este.

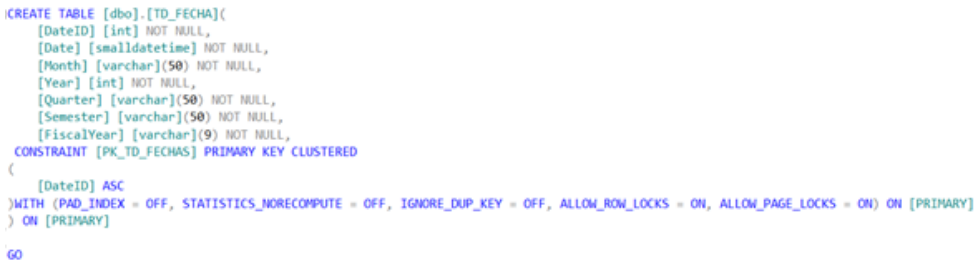
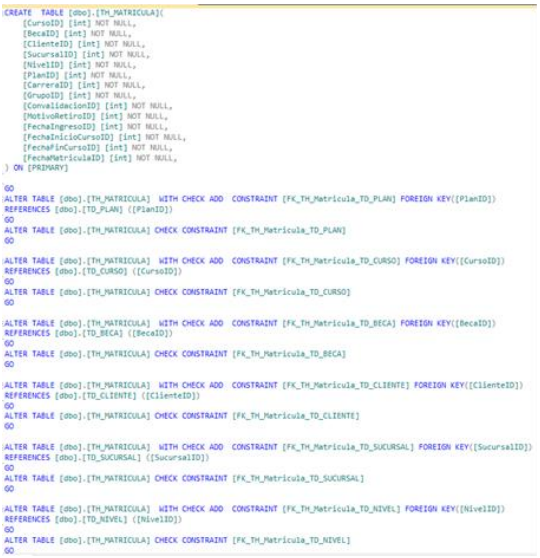


Ilustración 29 Diseño y desarrollo del Data Warehouse 2.

Una vez establecidas las dimensiones del modelo estrella, se generó el centro de este creando la tabla de hechos que hará referencia a todas las dimensiones del modelo, esto mediante el ingreso de llaves foráneas. Para el caso de la dimensión de fechas, al contar con cuatro registros de fechas, es requerido establecer 4 llaves foráneas para cada campo hacia la misma dimensión.



```
CREATE  TABLE [dbo].[TH_MATRICULA](
    [CursoID] [int] NOT NULL,
    [BecaID] [int] NOT NULL,
    [ClienteID] [int] NOT NULL,
    [SucursalID] [int] NOT NULL,
    [NivelID] [int] NOT NULL,
    [PlanID] [int] NOT NULL,
    [CarreraID] [int] NOT NULL,
    [GrupoID] [int] NOT NULL,
    [ConvalidacionID] [int] NOT NULL,
    [MotivoRetiroID] [int] NOT NULL,
    [FechaIngresoID] [int] NOT NULL,
    [FechaInicioCursoID] [int] NOT NULL,
    [FechaFinCursoID] [int] NOT NULL,
    [FechaMatriculaID] [int] NOT NULL,
) ON [PRIMARY]

GO
ALTER TABLE [dbo].[TH_MATRICULA]  WITH CHECK ADD  CONSTRAINT [FK_TH_Matricula_TD_PLAN] FOREIGN KEY([PlanID])
REFERENCES [dbo].[TD_PLAN] ([PlanID])
GO
ALTER TABLE [dbo].[TH_MATRICULA] CHECK CONSTRAINT [FK_TH_Matricula_TD_PLAN]
GO

ALTER TABLE [dbo].[TH_MATRICULA]  WITH CHECK ADD  CONSTRAINT [FK_TH_Matricula_TD_CURSO] FOREIGN KEY([CursoID])
REFERENCES [dbo].[TD_CURSO] ([CursoID])
GO
ALTER TABLE [dbo].[TH_MATRICULA] CHECK CONSTRAINT [FK_TH_Matricula_TD_CURSO]
GO

ALTER TABLE [dbo].[TH_MATRICULA]  WITH CHECK ADD  CONSTRAINT [FK_TH_Matricula_TD_BECA] FOREIGN KEY([BecaID])
REFERENCES [dbo].[TD_BECA] ([BecaID])
GO
ALTER TABLE [dbo].[TH_MATRICULA] CHECK CONSTRAINT [FK_TH_Matricula_TD_BECA]
GO

ALTER TABLE [dbo].[TH_MATRICULA]  WITH CHECK ADD  CONSTRAINT [FK_TH_Matricula_TD_CLIENTE] FOREIGN KEY([ClienteID])
REFERENCES [dbo].[TD_CLIENTE] ([ClienteID])
GO
ALTER TABLE [dbo].[TH_MATRICULA] CHECK CONSTRAINT [FK_TH_Matricula_TD_CLIENTE]
GO

ALTER TABLE [dbo].[TH_MATRICULA]  WITH CHECK ADD  CONSTRAINT [FK_TH_Matricula_TD_SUCURSAL] FOREIGN KEY([SucursalID])
REFERENCES [dbo].[TD_SUCURSAL] ([SucursalID])
GO
ALTER TABLE [dbo].[TH_MATRICULA] CHECK CONSTRAINT [FK_TH_Matricula_TD_SUCURSAL]
GO

ALTER TABLE [dbo].[TH_MATRICULA]  WITH CHECK ADD  CONSTRAINT [FK_TH_Matricula_TD_NIVEL] FOREIGN KEY([NivelID])
REFERENCES [dbo].[TD_NIVEL] ([NivelID])
GO
ALTER TABLE [dbo].[TH_MATRICULA] CHECK CONSTRAINT [FK_TH_Matricula_TD_NIVEL]
GO
```

Ilustración 30 Diseño y desarrollo del Data Warehouse 3.

Se procedió a la creación de un nuevo proyecto utilizando la herramienta SQL Server Integration Services 2012, donde se ingresan dos conexiones, una para extraer la información del Staging Area ubicada en la base de datos CENFOTEC_SA y una segunda conexión a la base de datos CENFOTEC_DW, donde se encuentra almacenado el modelo de estrella.

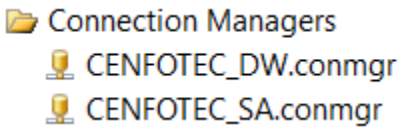


Ilustración 29 Diseño y desarrollo del Data Warehouse 4.

Las dimensiones se ingresaron en un Sequence Container, el cual permite establecer el ordenamiento de corrida del ETL. Para cada dimensión se utiliza un Data Flow Task para configurar su procedimiento.



Ilustración 30 Diseño y desarrollo del Data Warehouse 5.

Dentro de cada dimensión se ingresan los componentes OLE DB Source para establecer los datos fuente, que en este caso fueron los ubicados en el Staging Area. Se realiza un alias donde se establecen los mismos nombres de columnas que contiene la dimensión, y se utiliza una columna de referencia para hacer llamado a la llave primaria de la tabla como tal en el transaccional.

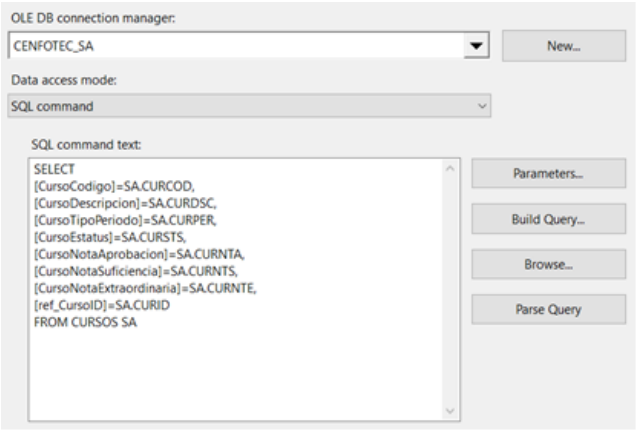


Ilustración 31 Diseño y desarrollo del Data Warehouse 6.

Para validar la correcta consulta de los datos, seleccionamos la opción de visualización y nos validará los datos que se ingresarán en la dimensión. Para algunos casos se observó cómo se toma de columna de referencia las llaves que se conformaron a partir de varias columnas concatenadas.

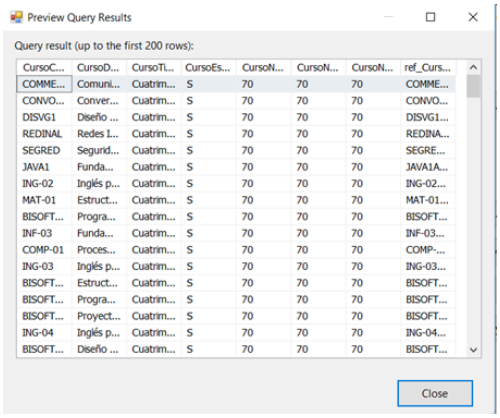


Ilustración 32 Diseño y desarrollo del Data Warehouse 7.

En la base de datos transaccional se encontraron muchos valores relevantes de tipo nvarchar, para nivel de compatibilidad y mejorar el almacenamiento del Data Warehouse se realiza una conversión del tipo de dato en las dimensiones que cuentan con valores nvarchar en sus columnas, utilizando la herramienta Data Conversión se efectúa esta tarea.

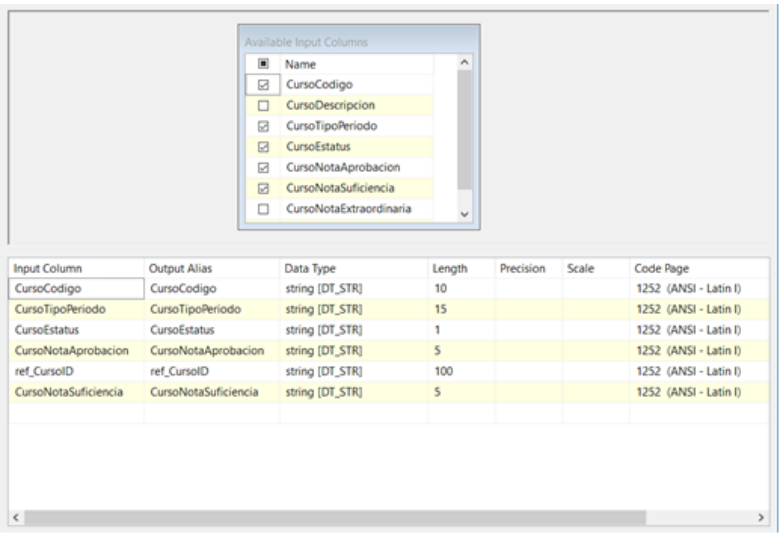


Ilustración 33 Diseño y desarrollo del Data Warehouse 8.

A la hora de realizar la inserción de datos en las dimensiones se utilizó la herramienta Slowly Changing Dimension, la cual permite evitar registros duplicados, fijar la columna de referencia del transaccional y permitir al modelo actualizarse en sus registros conforme se vaya ejecutando.

Como primer paso es requerido seleccionar lo que se conoce como Business Key, que es establecer cuál es la llave de la tabla del transaccional.

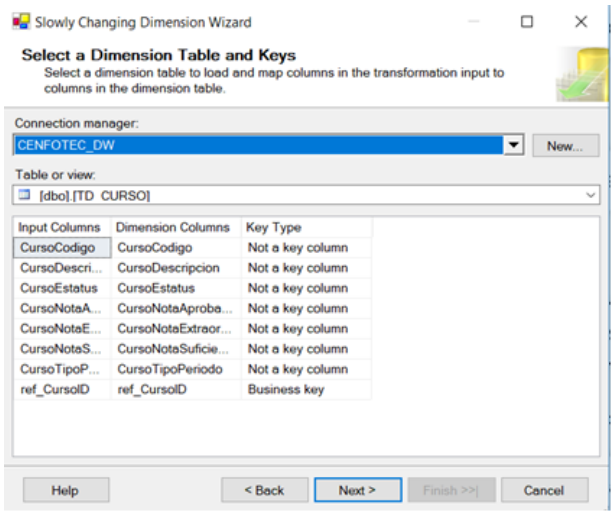


Ilustración 34 Diseño y desarrollo del Data Warehouse 9.

Se verifican los campos, y se establece la propiedad Change Type con la variable Changing attribute, la cual establece que los registros ya existentes sean borrados y actualizados.

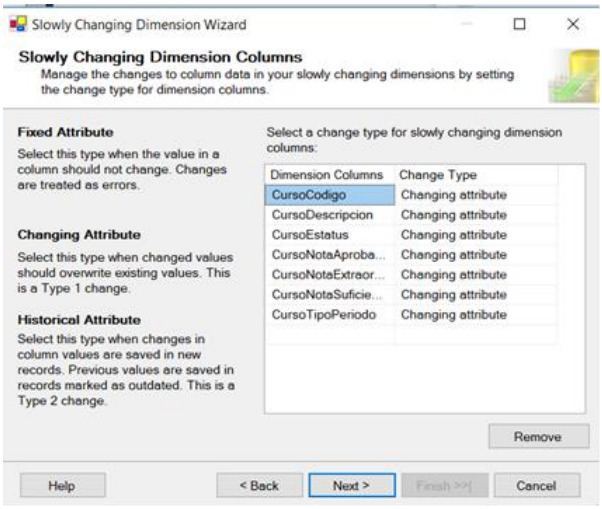


Ilustración 35 Diseño y desarrollo del Data Warehouse 10.

Al finalizar la configuración del Slowly Changing Dimension, automáticamente se crean dos archivos que validan la inserción de nuevos datos y la actualización de los existentes.

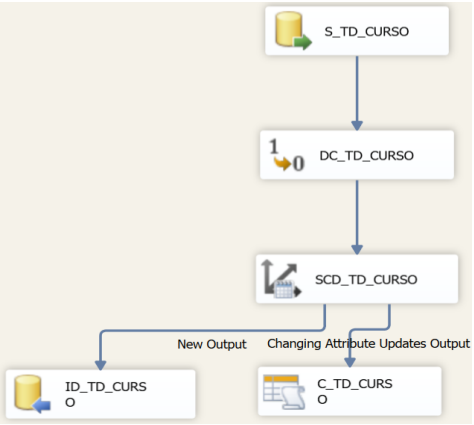


Ilustración 36 Diseño y desarrollo del Data Warehouse 11.

En la dimensión de clientes se valida si el estudiante posee un origen académico en sus registros para que le permita al negocio establecer reportes conforme a los diferentes centros de estudio de los que provienen sus estudiantes.

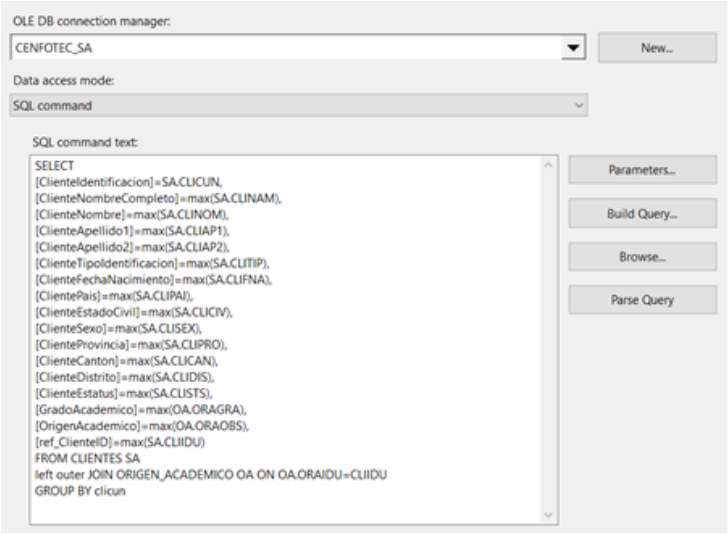


Ilustración 37 Diseño y desarrollo del Data Warehouse 12.

Para la dimensión de TD_GRUPO fue requerido validar el tipo de registro y verificar que sea únicamente de clientes, en este caso de estudiantes, ya que la tabla como tal ingresaba registros propios de los profesores.

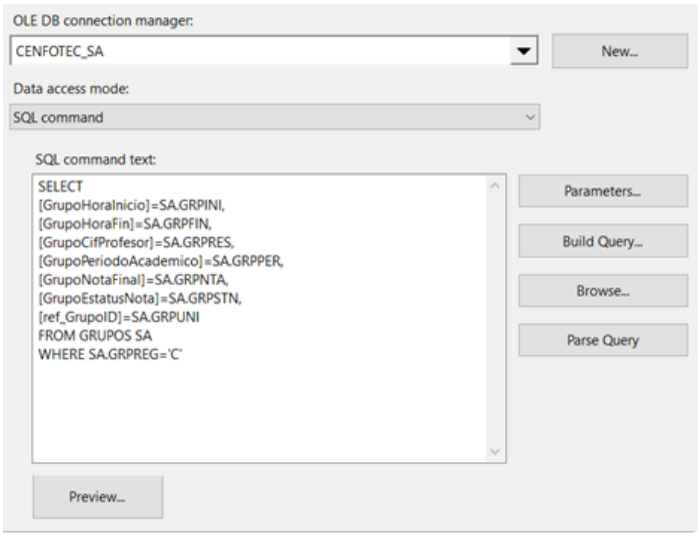


Ilustración 38 Diseño y desarrollo del Data Warehouse 13.

En el caso de la dimensión de TD_FECHA, se utiliza un Execute SQL TASK, que permita ingresar un script que sea capaz de generar un identificador para cada fecha existente del primer registro en la base de datos transaccional, hasta el día actual en que se ejecuten los procesos del modelo de inteligencia de negocios.

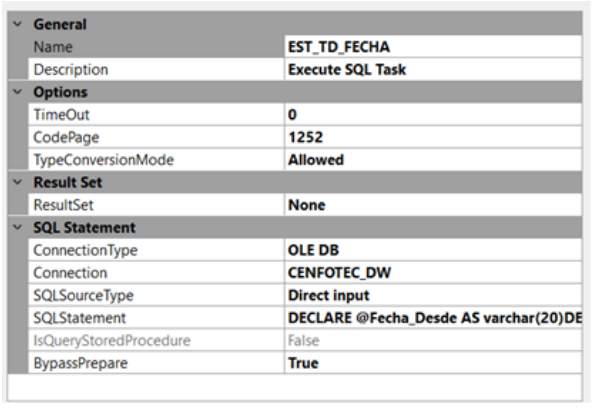


Ilustración 39 Diseño y desarrollo del Data Warehouse 14.

Al ejecutar el modelo la carga inicial se debe contemplar todos los registros ya existentes, por ello se toma la fecha del primer registro en que se asignó un grupo al estudiante. A la fecha de ejecución, en futuras ejecuciones se tomará en cuenta únicamente la fecha actual.

```
SET @Fecha_Desde = '19900101'
SET @Fecha_Hasta = '20170711'

--SET @Fecha_Desde = (SELECT CONVERT(VARCHAR,GETDATE(),112))
--SET @Fecha_Hasta = (SELECT CONVERT(VARCHAR,GETDATE(),112))
```

Ilustración 40 Diseño y desarrollo del Data Warehouse 15.

El script que genera la dimensión de fechas almacena datos como la fecha, el mes, el año, semestre en que sucedió el evento, cuatrimestre y año fiscal del suceso. Para cada fecha se establece un identificador, al insertar registros en la tabla de hechos estos identificadores serán igualados contra la dimensión de fechas para establecer la relación y poder extraer esta información.

```
DECLARE @FechaInicial AS SMALLDATETIME,
        @FechaFinal AS SMALLDATETIME,
        @Mes    AS INT,
        @Anno   AS INT


    SET NOCOUNT ON
    SELECT      @FechaInicial = convert(SMALLDATETIME,@fecha_desde,103),
                @FechaFinal = convert(SMALLDATETIME,@fecha_hasta,103)

    WHILE       @FechaInicial < @FechaFinal

        BEGIN

        SELECT  @Mes = datepart(mm,@FechaInicial), @Anno = datepart(yy,@FechaInicial)

        INSERT  [TD_FECHA]
        SELECT  FechaID =       cast(@FechaInicial AS int),
                Fecha   =       @FechaInicial,
                DiaMes  =       datepart(dd,@FechaInicial),
                DiaAnno =       datepart(dy,@FechaInicial),
                DescDia =       CASE
                                        WHEN datepart(dw,@FechaInicial) = 1 THEN 'Domingo'
                                        WHEN datepart(dw,@FechaInicial) = 2 THEN 'Lunes'
                                        WHEN datepart(dw,@FechaInicial) = 3 THEN 'Martes'
                                        WHEN datepart(dw,@FechaInicial) = 4 THEN 'Miércoles'
                                        WHEN datepart(dw,@FechaInicial) = 5 THEN 'Jueves'
                                        WHEN datepart(dw,@FechaInicial) = 6 THEN 'Viernes'
                                        ELSE 'Sábado'
                                END,
                Mes     =       @mes,
                DescMes =       CASE
                                        WHEN @mes = 1  then 'Enero'
                                        WHEN @mes = 2  then 'Febrero'
                                        WHEN @mes = 3  then 'Marzo'
                                        WHEN @mes = 4  then 'Abril'
                                        WHEN @mes = 5  then 'Mayo'
                                        WHEN @mes = 6  then 'Junio'
                                        WHEN @mes = 7  then 'Julio'
                                        WHEN @mes = 8  then 'Agosto'
                                        WHEN @mes = 9  then 'Setiembre'
                                        WHEN @mes = 10 then 'Octubre'
                                        WHEN @mes = 11 then 'Noviembre'
                                        ELSE 'Diciembre'
                                END,
                Ano     =       @anno,
                Trimestre =     CASE
                                        WHEN @mes BETWEEN 1 AND 3 THEN 'Trim 1-'+ cast(@Anno AS char(4))
                                        WHEN @mes BETWEEN 4 AND 6 THEN 'Trim 2-'+ cast(@Anno AS char(4))
                                        WHEN @mes BETWEEN 7 AND 9 THEN 'Trim 3-'+ cast(@Anno AS char(4))
                                        ELSE 'Trim 4-'+ cast(@Anno AS char(4))
```

Ilustración 41 Diseño y desarrollo del Data Warehouse 16.

Una vez establecidas las dimensiones se crea un proceso de actualización de registros, para validar que no existan campos nulos en ellos, utilizando el Execute SQL Task se utiliza un script que permita ingresar registros válidos cada vez que exista un valor nulo entre dimensiones.

```
----------[TD_CLIENTE]----------
IF NOT EXISTS(SELECT [ClienteID] FROM [dbo].[TD_CLIENTE] WHERE [ClienteID] = -1)
BEGIN

SET IDENTITY_INSERT CENFOTEC_DW.[dbo].[TD_CLIENTE] ON

INSERT INTO [dbo].[TD_CLIENTE]
           (
            [ClienteID],
            [ClienteIdentificacion],
            [ClienteNombreCompleto],
            [ClienteNombre],
            [ClienteApellido1],
            [ClienteApellido2],
            [ClienteTipoIdentificacion],
            [ClienteFechaNacimiento],
            [ClientePais],
            [ClienteEstadoCivil],
            [ClienteSexo],
            [ClienteProvincia],
            [ClienteCanton],
            [ClienteDistrito],
            [ClienteEstatus],
            [GradoAcademico],
            [OrigenAcademico],
            [ref_ClienteID]
           )
           VALUES
           (-1,'N/A','N/A','N/A','N/A','N/A','N/A','N/A','19000101','N/A','N/A','N/A','N/A','N/A','N/A','N/A','N/A','N/A',-1)

SET IDENTITY_INSERT CENFOTEC_DW.[dbo].[TD_CLIENTE] OFF

END
```

Ilustración 42 Diseño y desarrollo del Data Warehouse 17.

Se verifica que ninguna dimensión contenga valores de tipo nulo, se procede a utilizar el Data Flow Task donde se realiza el proceso de creación de la tabla de hechos TH_MATRICULA, utilizando el OLE DB Source para extraer la información y el OLE DB Destination.

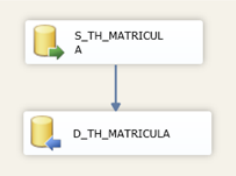


Ilustración 43 Diseño y desarrollo del Data Warehouse 18.

Dentro del OLE DB Source se establecerán las relaciones entre los campos de la tabla de hechos y las dimensiones, para ello es necesario primeramente relacionar las tablas del Staging Area con los campos de referencia de las dimensiones. Así

obtener la información a partir de la información de la tabla cliente e ir obteniendo las fechas de matrícula, ingreso a la universidad, fecha de inicio y fin del curso.

Para la relación de dimensiones se utiliza el LEFT JOIN debido a que no todos los estudiantes cuentan con un registro en todas las dimensiones, como es el caso de la beca que abarca cierta cantidad de clientes.



Ilustración 44 Diseño y desarrollo del Data Warehouse 19.

En el OLE DB Destination se establece la tabla de hechos apuntando a la conexión CENFOTEC_DW, y con ello se completa el modelo del Data Warehouse.



Ilustración 45 Diseño y desarrollo del Data Warehouse 20.

Teniendo las dimensiones, validaciones y tabla de hechos, se procedió con la corrida de los procesos con resultados satisfactorios.

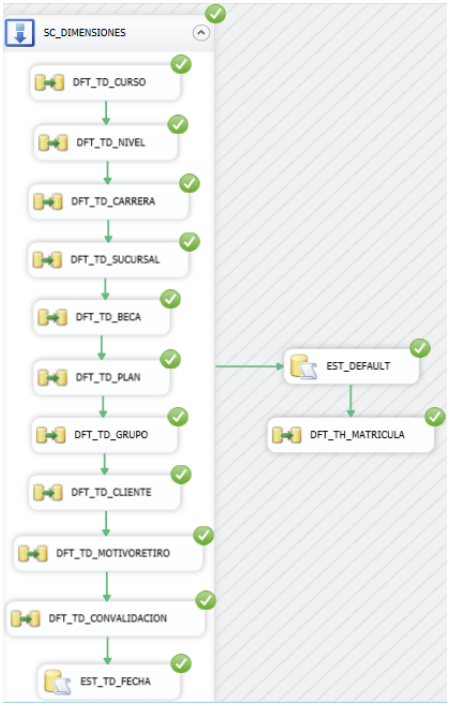


Ilustración 46 Diseño y desarrollo del Data Warehouse 21.

5.5.10.Consideraciones de la creación del DataWarehouse

Para la creación del Data Warehouse de la base de datos de la Universidad Cenfotec fue necesario tomar en cuenta las siguientes consideraciones:

- La base de datos transaccional no contaba con relaciones entre tablas, aparte existían tablas que no contaban con llaves primarias. Fue requerido un riguroso análisis para validar la generación del modelo estrella, como se mencionó anteriormente se optó por la generación de llaves compuestas en el Staging Area, las cuales ahora sirven como campo de referencia de este. Además, es importante los procesos de pruebas técnicas que permitieron conocer si la información conformada

permite validar el historial del estudiante, esto mediante las relaciones entre tablas de manera adecuada.

- Existían a nivel de la base de datos transaccional valores de tipo nvarchar, por lo cual se utilizó en las dimensiones la herramienta de Integration Services llamada Data Conversion, la cual permite realizar la conversión a varchar. Esto ahorrará espacio a nivel de la base de datos, ya que el valor del nvarchar corresponde a dos bytes, mientras que el varchar únicamente a uno.

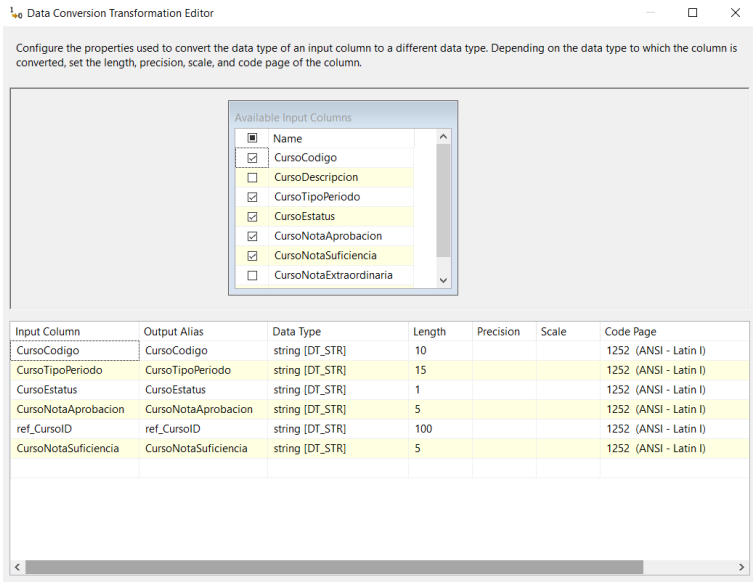


Ilustración 47 Diseño y desarrollo del Data Warehouse 22.

- Para cuestiones de rendimiento y tamaño de la base de datos se implementa una única dimensión para el manejo de fechas, esta corresponde a un script de base de datos que genera un identificador para cada fecha que se abarque dentro de los rangos especificados.

Normalmente se toma como fecha de inicio cuando se realizó la creación del primer registro de un estudiante y la fecha actual como límite máximo.

	DateID	Date	Month	Year	Quarter	Semester	FiscalYear
10...	42910	2017-06-26 00:00:00	06-June	2017	Quarter 2-2017	Semester 1-2017	2016-2017
10...	42911	2017-06-27 00:00:00	06-June	2017	Quarter 2-2017	Semester 1-2017	2016-2017
10...	42912	2017-06-28 00:00:00	06-June	2017	Quarter 2-2017	Semester 1-2017	2016-2017
10...	42913	2017-06-29 00:00:00	06-June	2017	Quarter 2-2017	Semester 1-2017	2016-2017
10...	42914	2017-06-30 00:00:00	06-June	2017	Quarter 2-2017	Semester 1-2017	2016-2017
10...	42915	2017-07-01 00:00:00	07-July	2017	Quarter 3-2017	Semester 2-2017	2016-2017

Ilustración 48 Diseño y desarrollo del Data Warehouse 23.

- Una necesidad del centro educativo correspondía a la generación de análisis con respecto al origen de procedencia del estudiante, para esto se optó por relacionar el origen académico directamente con el cliente para no generar toda una tabla para dos campos requeridos que correspondían a características propias del estudiante. Se analizó que no todos poseían un origen asignado, por lo cual se tomó únicamente donde existiera un registro válido.

```
SELECT
[ClienteIdentificacion]=SA.CLICUN,
[ClienteNombreCompleto]=max(SA.CLINAM),
[ClienteNombre]=max(SA.CLINOM),
[ClienteApellido1]=max(SA.CLIAP1),
[ClienteApellido2]=max(SA.CLIAP2),
[ClienteTipoIdentificacion]=max(SA.CLITIP),
[ClienteFechaNacimiento]=max(SA.CLIFNA),
[ClientePais]=max(SA.CLIPAI),
[ClienteEstadoCivil]=max(SA.CLICIV),
[ClienteSexo]=max(SA.CLISEX),
[ClienteProvincia]=max(SA.CLIPRO),
[ClienteCanton]=max(SA.CLICAN),
[ClienteDistrito]=max(SA.CLIDIS),
[ClienteEstatus]=max(SA.CLISTS),
[GradoAcademico]=max(OA.ORAGRA),
[OrigenAcademico]=max(OA.ORAOBS),
[ref_ClienteID]=max(SA.CLIIDU)
FROM CLIENTES SA
left outer JOIN ORIGEN_ACADEMICO OA ON OA.ORAIDU=CLIIDU
GROUP BY clicun
```

Ilustración 49 Diseño y desarrollo del Data Warehouse 24.

- Por cuestiones de depuración de datos, en la dimensión TD_GRUPO se encontraron registros no válidos propios del profesor, fue necesario analizar un campo que permitiera únicamente registrar grupos donde estuviera asignado un estudiante válido y con base en esto generar la dimensión.

```
SELECT
[GrupoHoraInicio]=SA.GRPINI,
[GrupoHoraFin]=SA.GRPFIN,
[GrupoCifProfesor]=SA.GRPRES,
[GrupoPeriodoAcademico]=SA.GRPPER,
[GrupoNotaFinal]=SA.GRPNTA,
[GrupoEstatusNota]=SA.GRPSTN,
[ref_GrupoID]=SA.GRPUNI
FROM GRUPOS SA
WHERE SA.GRPREG='C'
```

Ilustración 50 Diseño y desarrollo del Data Warehouse 25.

5.5.11. Diseño y desarrollo de la Base de Datos Multidimensional

Mediante la Base de Datos Multidimensional se le permitirá al centro educativo las funcionalidades de poder consultar el Data Warehouse desde una herramienta externa al gestor de base de datos. Esta herramienta toma el modelo de Data Warehouse, y mediante el uso de jerarquías que permite establecer las visualizaciones del modelo y perspectivas que establecen los niveles de seguridad de la información. Esto para que finalmente el usuario ingrese la conexión desde una herramienta externa y visualice toda la informacion almacenada en el Data Warehouse.

Para la base de datos Multidimensional se utilizó la conexión a la base de datos del Data Warehouse, una vez seleccionada la conexión se procede a escoger los elementos del modelo multidimensional, mediante el Data Source View. Al tener seleccionado el modelo del cual obtendremos la base de datos multidimensional, se crea el cubo donde se establecen las dimensiones, tabla de hechos y métricas de

nuestro modelo. Para cada dimensión es requerido establecer la manera en que los usuarios van a visualizar tanto el nombre de la dimensión como de sus campos, este debe ser un nombre entendible para el negocio, además de las jerarquías que permiten establecer la secuencia de visualización de los campos de las dimensiones. Por último, se procede con las perspectivas que definen a qué nivel de seguridad puede visualizar un usuario en específico el modelo.

5.5.12.Diseño de arquitectura de la Base de Datos Multidimensional

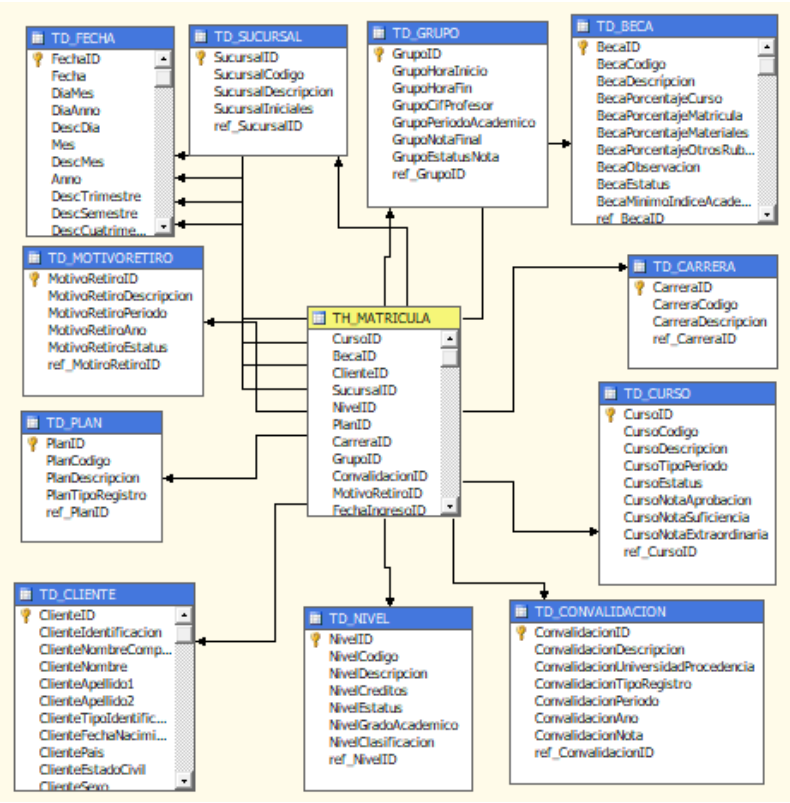


Ilustración 51 Diseño de arquitectura de la Base de Datos Multidimensional 1.

5.5.13.Definición de jerarquías

Los atributos se pueden organizar en jerarquías definidas por el usuario que proporcionan rutas de navegación para ayudar a los usuarios al examinar los datos de un cubo (Microsoft, 2016). A nivel del proyecto realizado se establecieron jerarquías

que abarquen todos los datos de cada dimensión, ya que los niveles de visualización de los datos pertenecen a un único rol de Administrador el cual podrá visualizar todo el modelo sin restricción de jerarquía. El ordenamiento de cada dimensión se basó en un código como cabecera de la jerarquía según corresponda y una descripción del registro, así como el resto de los campos de la dimensión.

5.5.14.Definición de perspectivas

Una perspectiva es un subconjunto de un cubo creado para una aplicación o grupo de usuarios específico. El cubo es la perspectiva predeterminada. Una perspectiva se muestra al cliente como un cubo. Cuando el usuario ve una perspectiva, se muestra como otro cubo. Los cambios realizados en los datos del cubo mediante la reescritura en la perspectiva se realizan en el cubo original (Microsoft, 2016). En este proyecto se utilizó únicamente la perspectiva de Administrador, quien tiene control acceso a los datos desde la herramienta externa, según lo solicitado por el negocio.

5.5.15.Estrategia de Base de Datos Multidimensional

Se crea un Data Source que haga referencia a la base de datos del Data Warehouse.

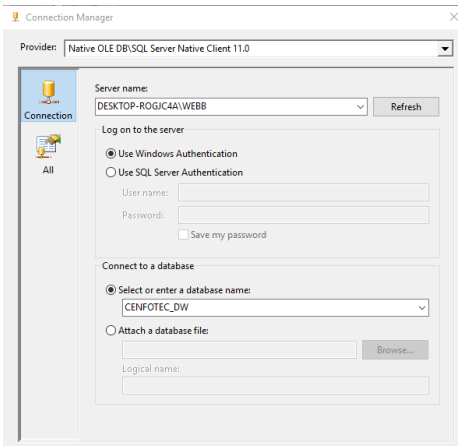


Ilustración 52 Diseño de arquitectura de la Base de Datos Multidimensional 2.

Se establece el Data Source View, para delimitar los elementos que estarán en el cubo. Primeramente, se selecciona el Data Source ingresado.

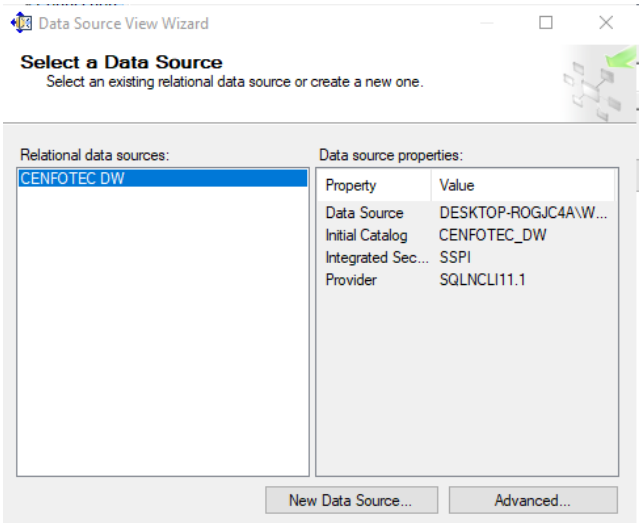


Ilustración 53 Diseño de arquitectura de la Base de Datos Multidimensional 3.

En el siguiente paso aparecen todos los elementos que existen en el Data Source seleccionado, para ello debemos de elegir todas las dimensiones y tabla de hechos del modelo.

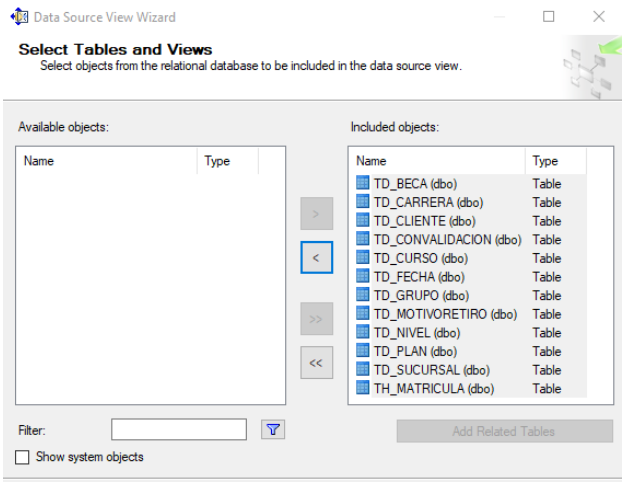


Ilustración 54 Diseño de arquitectura de la Base de Datos Multidimensional 4.

Para finalizar este paso se ingresa un nombre único al Data Source View.

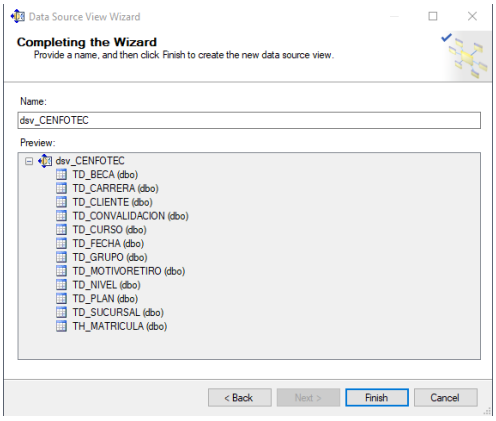


Ilustración 55 Diseño de arquitectura de la Base de Datos Multidimensional 5

Se crea el cubo respectivo seleccionado y las tablas ya seleccionadas del Data Source View.

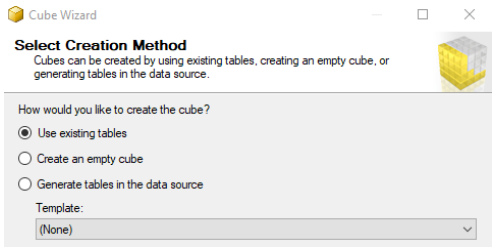


Ilustración 56 Diseño de arquitectura de la Base de Datos Multidimensional 6.

Se selecciona la tabla de hechos que posee las métricas y relaciones de dimensiones.

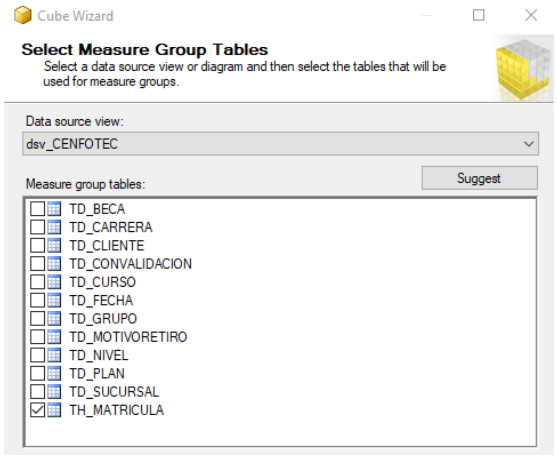


Ilustración 57 Diseño de arquitectura de la Base de Datos Multidimensional 7.

Se seleccionan todas las dimensiones que se desean visualizar en el cubo.

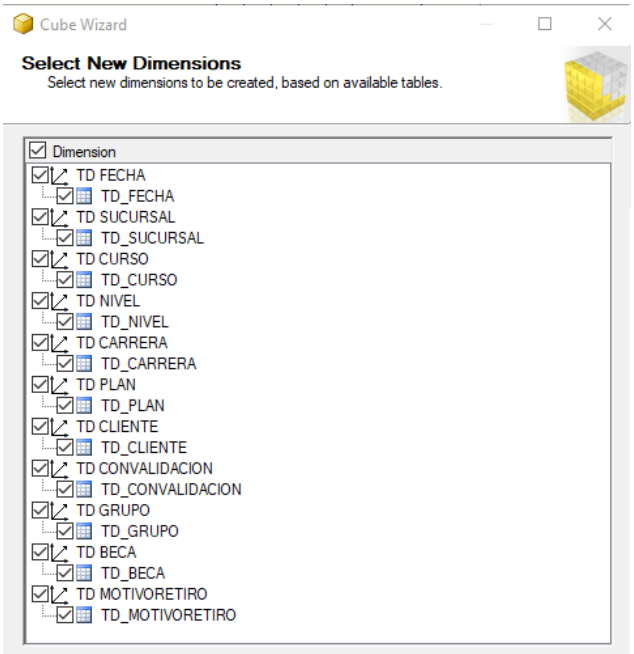


Ilustración 58 Diseño de arquitectura de la Base de Datos Multidimensional 8.

Se ingresó un nombre único al cubo creado.

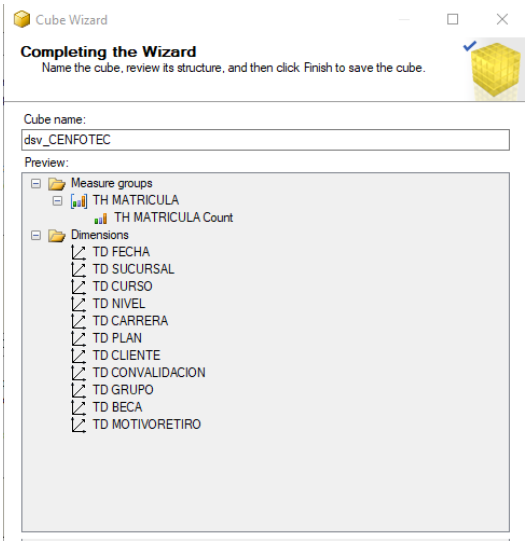


Ilustración 59 Diseño de arquitectura de la Base de Datos Multidimensional 9.

Una vez creado el cubo se genera el modelo multidimensional.

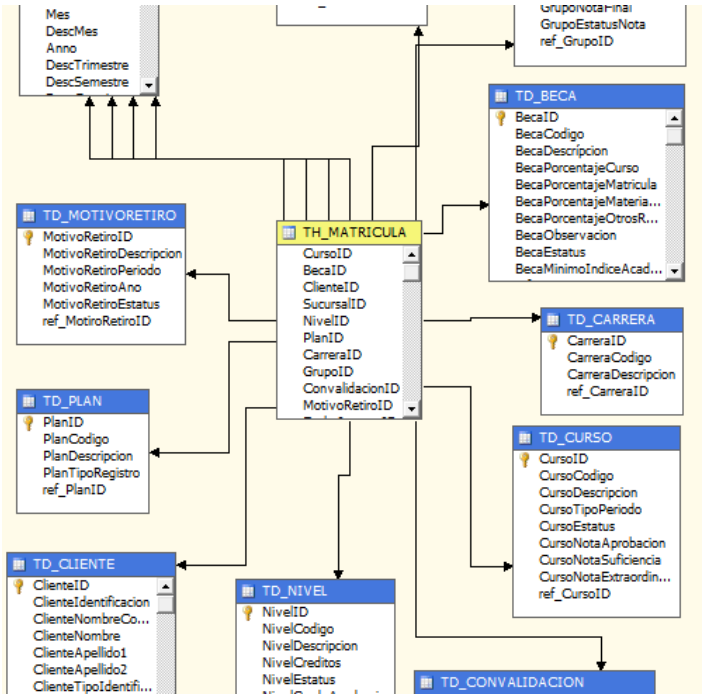


Ilustración 60 Diseño de arquitectura de la Base de Datos Multidimensional 10.

Para cada dimensión se establecen jerarquías que permiten establecer el orden de visualización. Como parte del desarrollo y análisis se estableció en las dimensiones dar como prioridad el código identificador, seguido de su descripción y el resto de los campos de cada dimensión.

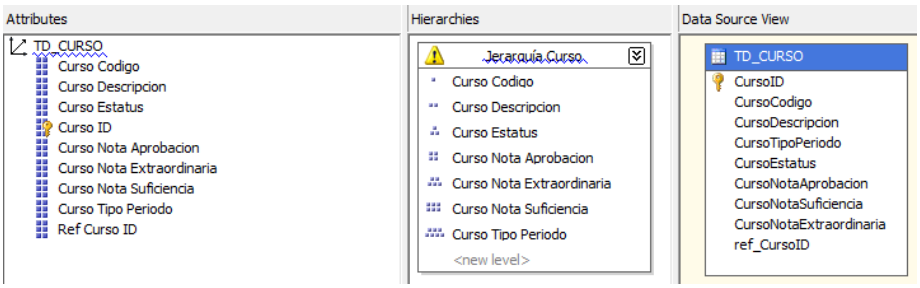
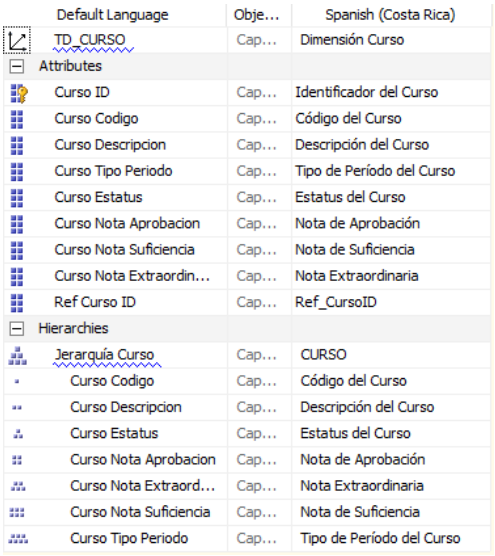


Ilustración 61 Diseño de arquitectura de la Base de Datos Multidimensional 11.

Además, en cada dimensión es requerido establecer tanto para sus campos como jerarquías la manera en que se visualizarán los nombres de cada atributo, ya que debe ser un nombre entendible por todos los usuarios del modelo y en el idioma que corresponda. Este proceso se aplica en todas las dimensiones del modelo.



Default Language	Obje...	Spanish (Costa Rica)
TD_CURSO	Cap...	Dimensión Curso
Attributes		
Curso ID	Cap...	Identificador del Curso
Curso Codigo	Cap...	Código del Curso
Curso Descripcion	Cap...	Descripción del Curso
Curso Tipo Periodo	Cap...	Tipo de Período del Curso
Curso Estatus	Cap...	Estatus del Curso
Curso Nota Aprobacion	Cap...	Nota de Aprobación
Curso Nota Suficiencia	Cap...	Nota de Suficiencia
Curso Nota Extraordin...	Cap...	Nota Extraordinaria
Ref Curso ID	Cap...	Ref_CursoID
Hierarchies		
Jerarquía Curso	Cap...	CURSO
Curso Codigo	Cap...	Código del Curso
Curso Descripcion	Cap...	Descripción del Curso
Curso Estatus	Cap...	Estatus del Curso
Curso Nota Aprobacion	Cap...	Nota de Aprobación
Curso Nota Extraord...	Cap...	Nota Extraordinaria
Curso Nota Suficiencia	Cap...	Nota de Suficiencia
Curso Tipo Periodo	Cap...	Tipo de Período del Curso

Ilustración 62 Diseño de arquitectura de la Base de Datos Multidimensional 12.

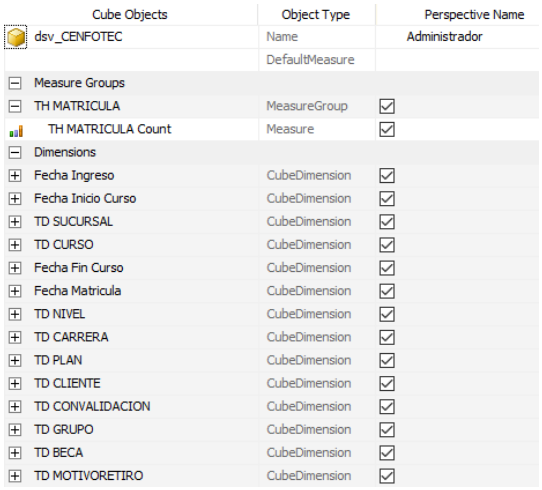
Se ingresan traducciones y cambios en la visualización de los campos a nivel de nombres de dimensiones y campos de la tabla de hechos.



Default Language	Obje...	Spanish (Costa Rica)
dsv_CENFOTEC	Cap...	
Measure Groups		
TH MATRICULA	Cap...	Matrícula
TH MATRICULA Count	Cap...	
Dimensions		
Fecha Ingreso	Cap...	Fecha de Ingreso
Fecha Inicio Curso	Cap...	Fecha de Inicio del Curso
TD SUCURSAL	Cap...	Dimensión Sucursal
TD CURSO	Cap...	Dimensión Curso
Fecha Fin Curso	Cap...	Fecha Fin del Curso
Fecha Matricula	Cap...	Fecha de Matrícula
TD NIVEL	Cap...	Dimensión Nivel
TD CARRERA	Cap...	Dimensión Carrera
TD PLAN	Cap...	Dimensión Plan
TD CLIENTE	Cap...	Dimensión Cliente
TD CONVALIDACION	Cap...	Dimensión Convalidación
TD GRUPO	Cap...	Dimensión Grupo
TD BECA	Cap...	Dimensión Beca
TD MOTIVORETIRO	Cap...	Dimensión Motivo de Retiro
Perspectives		
Administrador	Cap...	Administrador

Ilustración 63 Diseño de arquitectura de la Base de Datos Multidimensional 13.

Se procedió con la generación de perspectivas, a nivel del centro educativo se solicitó únicamente una perspectiva de un rol de administrador que permita visualizar todo el modelo.



Cube Objects	Object Type	Perspective Name
dsv_CENFOTEC	Name	Administrador
	DefaultMeasure	
Measure Groups		
TH MATRICULA	MeasureGroup	☑
TH MATRICULA Count	Measure	☑
Dimensions		
Fecha Ingreso	CubeDimension	☑
Fecha Inicio Curso	CubeDimension	☑
TD SUCURSAL	CubeDimension	☑
TD CURSO	CubeDimension	☑
Fecha Fin Curso	CubeDimension	☑
Fecha Matricula	CubeDimension	☑
TD NIVEL	CubeDimension	☑
TD CARRERA	CubeDimension	☑
TD PLAN	CubeDimension	☑
TD CLIENTE	CubeDimension	☑
TD CONVALIDACION	CubeDimension	☑
TD GRUPO	CubeDimension	☑
TD BECA	CubeDimension	☑
TD MOTIVORETIRO	CubeDimension	☑

Ilustración 64 Diseño de arquitectura de la Base de Datos Multidimensional 14.

Una vez generado el cubo, sus perspectivas, traducciones y jerarquías se procede a ejecutar el modelo multidimensional.

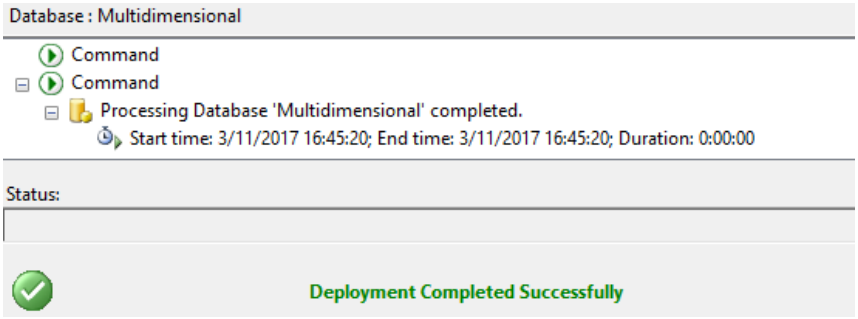


Ilustración 65 Diseño de arquitectura de la Base de Datos Multidimensional 15.

Para efectos de pase a producción se procedió con la creación de un nuevo proyecto de Integration Services, el cual permite ejecutar la actualización de la base de datos Multidimensional desde un paquete de Integration Services, y con esto generarlo

fácilmente desde el servidor de base de datos. Se ingresa el componente llamado Analysis Services Processing Task, el cual permite llamar y ejecutar una base de datos Multidimensional, se agrega la conexión a base de datos.

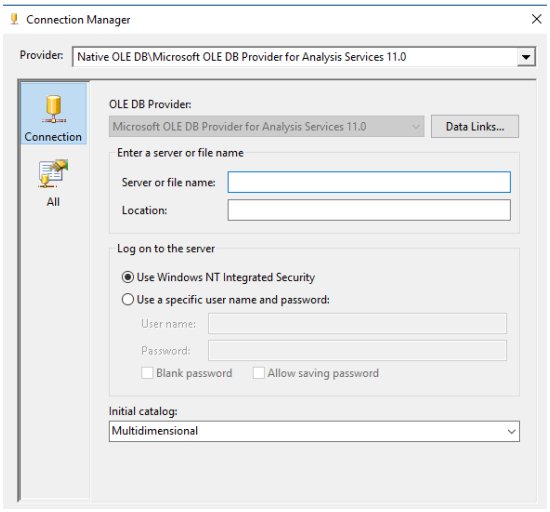


Ilustración 66 Diseño de arquitectura de la Base de Datos Multidimensional 16.

Se seleccionan los componentes que se actualizarán al ejecutar el paquete, para este escenario se seleccionó desde el padre, y realiza la actualización de todos los componentes.

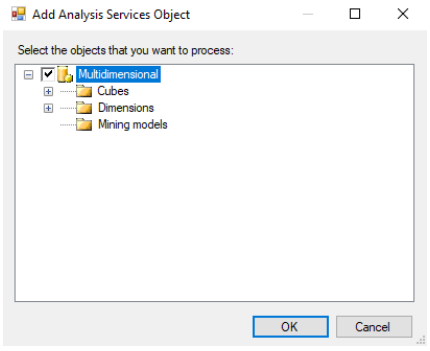


Ilustración 67 Diseño de arquitectura de la Base de Datos Multidimensional 17

Una vez realizada la configuración, se finaliza con el componente y queda listo para ejecutarse.

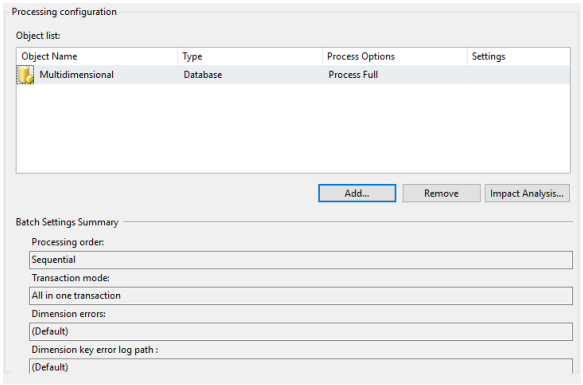


Ilustración 68 Diseño de arquitectura de la Base de Datos Multidimensional 18.

Se ejecuta con resultados satisfactorios.



Ilustración 69 Diseño de arquitectura de la Base de Datos Multidimensional 19.

5.5.16.Consideraciones de la creación de la Base de Datos Multidimensional

Para la creación de la base de datos Multidimensional de la Universidad Cenfotec fue necesario tomar en cuenta las siguientes consideraciones:

- Se requirió una modificación completa de los nombres de los siguientes elementos: nombres de dimensiones, campos de las dimensiones, nombres de las jerarquías, campos de las jerarquías, nombre de la tabla de hechos. Permite así a los usuarios establecer un lenguaje de mayor entendimiento a la hora de validar los campos del modelo y generar sus reportes sin mayor dificultad.
- A solicitud del negocio se creó una única perspectiva de administrador que permita visualizar todos los campos del modelo como tal.

- Se encontraron algunos registros que seguían conservando valores nulos; por lo cual fue necesario replantear el modelo del Data Warehouse, validando cualquier campo de tipo valor nulo.
- Se encontraron campos que no contenían la misma cantidad de caracteres de la base de datos transaccional al modelo del Data Warehouse, lo cual causó problemas en la generación de la base de datos Multidimensional. Se solucionó estableciendo estándares de tamaño en los campos.
- Se realizó la instalación de la herramienta Power BI para poder visualizar la información y generar reportes, el negocio nos consultó sobre una herramienta que permitiera generar estos con un uso flexible. Por anterior, se generaron ejemplos y pruebas con base en esta herramienta.

5.5.17. Estrategia de consultas del modelo

Para efectos de visualización del proyecto se utilizó la herramienta solicitada por el negocio, conocida como Power BI en su versión de escritorio. Power BI es un conjunto de herramientas de análisis empresarial que pone el conocimiento al alcance de toda la organización, conexión a cientos de orígenes de datos, preparación de datos simplificada, generación de análisis ad hoc. Esto permite la creación de informes que luego se publican para provecho de la organización en la Web y en dispositivos móviles (Microsoft, 2017).

Para la creación de informes simplemente se ingresa a la herramienta Power BI, se selecciona la opción Get Data, se selecciona la opción del tipo de conexión, en este caso se utiliza SQL Server Analysis Services Database.

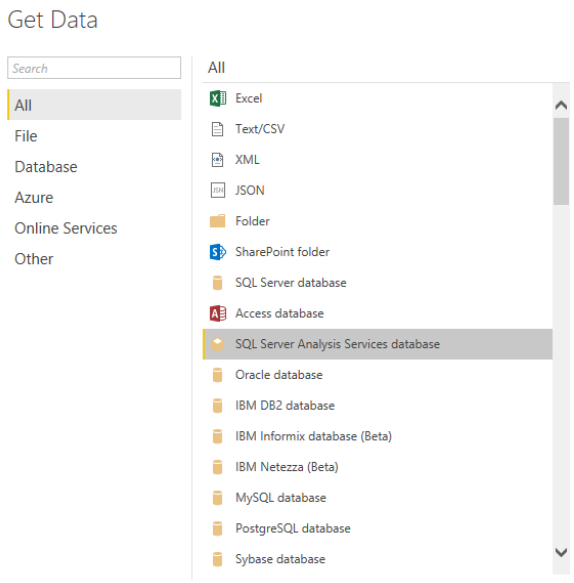


Ilustración 70 Estrategia de consultas del modelo 1.

Se crea la conexión a la base de datos, ingresando el nombre del servidor y la opción Connect live.

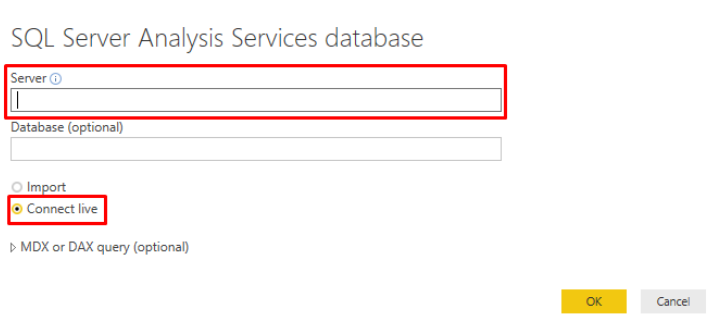


Ilustración 71 Estrategia de consultas del modelo 2.

Una vez realizada la conexión se selecciona el perfil, en este caso se cuenta con la perspectiva de Administrador.

La herramienta despliega las dimensiones del modelo multidimensional, lo que permite flexibilidad a la hora de seleccionar los campos que se desean incorporar en el

reporte. Además, permite generar filtros, así como visualizar los reportes de manera gráfica.

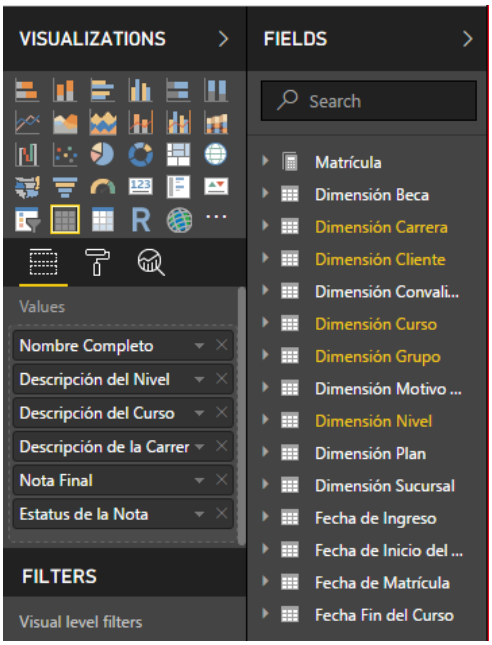


Ilustración 72 Estrategia de consultas del modelo 3.

Una vez seleccionados los campos podemos generar variedad de reportes; a continuación, unos ejemplos de estos:

Primeramente, se generó un filtro a un estudiante en específico que refleja el nivel al que pertenece, descripción de los cursos y la nota obtenida en cada uno de estos.

Nombre Completo	Descripción del Nivel	Descripción del Curso	Descripción de la Carrera	Nota Final
WEBB CALDERON GERALD CLARENCE	Maestría en Tecnología de Bases de Datos	Administración de Bases de Datos	Bases de Datos	84
WEBB CALDERON GERALD CLARENCE	Maestría en Tecnología de Bases de Datos	Afinamiento y Rendimiento de Bases de Datos	Bases de Datos	91
WEBB CALDERON GERALD CLARENCE	Maestría en Tecnología de Bases de Datos	Bases de Datos Multidimensionales	Bases de Datos	100
WEBB CALDERON GERALD CLARENCE	Maestría en Tecnología de Bases de Datos	Bases de Datos Relacionales-Objeto	Bases de Datos	88.00
WEBB CALDERON GERALD CLARENCE	Maestría en Tecnología de Bases de Datos	Depósitos de Datos	Bases de Datos	98.88
WEBB CALDERON GERALD CLARENCE	Maestría en Tecnología de Bases de Datos	Depósitos de Datos Avanzados	Bases de Datos	0.00
WEBB CALDERON GERALD CLARENCE	Maestría en Tecnología de Bases de Datos	Diseño y Calidad de Datos	Bases de Datos	93
WEBB CALDERON GERALD CLARENCE	Maestría en Tecnología de Bases de Datos	Gobernanza de la Información	Bases de Datos	96.00
WEBB CALDERON GERALD CLARENCE	Maestría en Tecnología de Bases de Datos	Programación Avanzada en SQL	Bases de Datos	86.00
WEBB CALDERON GERALD CLARENCE	Maestría en Tecnología de Bases de Datos	Seguridad de Bases de Datos	Bases de Datos	N/A

Ilustración 73 Estrategia de consultas del modelo 4.

Como segundo ejemplo se utilizó un estudiante que se encuentra asignado a dos carreras diferentes, y se muestra sin dificultad alguna la diferenciación de carreras y las notas de cada curso según corresponda.

Nombre Completo	Descripción del Nivel	Descripción de la Carrera	Descripción del Curso	Nota Final	Estatus de la Nota	Descripción de Fecha
SANCHEZ UGALDE JAVIER	Diplomado en Desarrollo y Diseño Web	Web y Medios Digitales Interactivos	Comunicación de Información en la Web	71.00	A	2015-05-11 00:00:00
SANCHEZ UGALDE JAVIER	Diplomado en Desarrollo y Diseño Web	Web y Medios Digitales Interactivos	Comunicación Social en la Web	59	R	2016-01-11 00:00:00
SANCHEZ UGALDE JAVIER	Diplomado en Desarrollo y Diseño Web	Web y Medios Digitales Interactivos	Diseño de la Experiencia del Usuario	23.00	R	2015-05-11 00:00:00
SANCHEZ UGALDE JAVIER	Diplomado en Desarrollo y Diseño Web	Web y Medios Digitales Interactivos	Diseño de la Experiencia del Usuario	80	A	2016-01-11 00:00:00
SANCHEZ UGALDE JAVIER	Diplomado en Desarrollo y Diseño Web	Web y Medios Digitales Interactivos	Diseño visual digital 1	71	A	2014-05-12 00:00:00
SANCHEZ UGALDE JAVIER	Diplomado en Desarrollo y Diseño Web	Web y Medios Digitales Interactivos	Diseño Visual Digital 2	91	A	2014-09-08 00:00:00
SANCHEZ UGALDE JAVIER	Diplomado en Desarrollo y Diseño Web	Web y Medios Digitales Interactivos	Diseño web 1	71	A	2013-09-09 00:00:00
SANCHEZ UGALDE JAVIER	Diplomado en Desarrollo y Diseño Web	Web y Medios Digitales Interactivos	Diseño Web 2	70	A	2014-05-12 00:00:00
SANCHEZ UGALDE JAVIER	Diplomado en Desarrollo y Diseño Web	Web y Medios Digitales Interactivos	Posicionamiento en la Web	N/A	N/A	2016-05-09 00:00:00
SANCHEZ UGALDE JAVIER	Diplomado en Desarrollo y Diseño Web	Web y Medios Digitales Interactivos	Programación Web Dinámica	17	R	2014-09-08 00:00:00
SANCHEZ UGALDE JAVIER	Diplomado en Desarrollo y Diseño Web	Web y Medios Digitales Interactivos	Programación Web Dinámica	79	A	2015-01-12 00:00:00
SANCHEZ UGALDE JAVIER	Diplomado en Desarrollo y Diseño Web	Web y Medios Digitales Interactivos	Programación Web Interactiva	0.00	N/A	2016-05-09 00:00:00
SANCHEZ UGALDE JAVIER	Diplomado en Desarrollo y Diseño Web	Web y Medios Digitales Interactivos	Programación Web Interactiva	51	R	2015-01-12 00:00:00
SANCHEZ UGALDE JAVIER	Diplomado en Desarrollo y Diseño Web	Web y Medios Digitales Interactivos	Programación Web Interactiva	51	R	2015-09-07 00:00:00
SANCHEZ UGALDE JAVIER	Ingeniería del Software	Desarrollo de Software	Estructuras Discretas	61	R	2013-05-13 00:00:00
SANCHEZ UGALDE JAVIER	Ingeniería del Software	Desarrollo de Software	Estructuras Discretas	70	A	2013-09-09 00:00:00
SANCHEZ UGALDE JAVIER	Ingeniería del Software	Desarrollo de Software	Ética y Profesionalismo	39	R	2014-01-13 00:00:00
SANCHEZ UGALDE JAVIER	Ingeniería del Software	Desarrollo de Software	Ética y Profesionalismo	85	A	2015-09-07 00:00:00
SANCHEZ UGALDE JAVIER	Ingeniería del Software	Desarrollo de Software	Fundamentos de Bases de Datos	68	R	2013-05-13 00:00:00
SANCHEZ UGALDE JAVIER	Ingeniería del Software	Desarrollo de Software	Fundamentos de Programación	27	R	2013-01-14 00:00:00
SANCHEZ UGALDE JAVIER	Ingeniería del Software	Desarrollo de Software	Fundamentos de Programación	70	A	2013-05-13 00:00:00
SANCHEZ UGALDE JAVIER	Ingeniería del Software	Desarrollo de Software	Inglés para Tecnologías de Información 1	77	A	2013-01-14 00:00:00
SANCHEZ UGALDE JAVIER	Ingeniería del Software	Desarrollo de Software	Inglés para Tecnologías de Información 2	75	A	2013-05-13 00:00:00
SANCHEZ UGALDE JAVIER	Ingeniería del Software	Desarrollo de Software	Inglés para Tecnologías de Información 3	70	A	2013-09-09 00:00:00
SANCHEZ UGALDE JAVIER	Ingeniería del Software	Desarrollo de Software	Inglés para Tecnologías de Información 4	70	A	2014-05-12 00:00:00

Ilustración 74 Estrategia de consultas del modelo 5.

5.5.18. Consideraciones de Seguridad

Para cuestiones de seguridad del modelo de inteligencia de negocios se tomaron las siguientes consideraciones de seguridad:

- Este proyecto cuenta con una seguridad integrada de Windows, el cliente solicito una única perspectiva a nivel de consulta de la base de datos multidimensional.
- Se recomendó la implementación de seguridad a través de la herramienta Active Directory (AD), que permita administrar los perfiles o usuarios que puedan acceder al modelo desarrollado.

- En cuando al administrador de base de datos, el cual se encarga de la gestión de cargas de los procesos del modelo, deberá contar con única máquina física donde solo este tenga los permisos respectivos tanto para el ingreso del servidor, como para la ejecución de las cargas.

- Crear nuevas perspectivas al modelo multidimensional según se vaya desarrollando el crecimiento del Data Warehouse y pensando a futuro en integrar los módulos financieros y administrativos del centro educativo.

5.6. Diccionario de datos

Como parte de los beneficios brindados en este proyecto se elaboraron dos diccionarios de datos propios de los desarrolladores del producto, uno para la elaboración del Staging Area (Ver Anexo 1) y un segundo diccionario para la elaboración del Data Warehouse (Ver Anexo 2).

Para la elaboración del Staging Area, encontramos dificultades a la hora de reconocer el diccionario de datos de la base de datos transaccional, ya que nos encontramos columnas que sí se encontraban en la base de datos, pero no se nombraban en su diccionario. Por ello, el equipo técnico tuvo que generar análisis para encontrar la definición correcta de cada campo relevante del modelo y se fijaron los nombres en este documento (ver Anexo 1) de los campos que fueron utilizados para el desarrollo del Data Warehouse. Mientras tanto, para la elaboración del Data Warehouse se creó un nuevo diccionario (ver Anexo 2), en donde se utilizaron nombres que comprendan una mayor facilidad y eficiencia tanto para el lector del documento como para los usuarios del proyecto, se pasaron a un lenguaje más estandarizado y de mayor comprensión.

5.7. Entregables

Se realiza la entrega de los siguientes entregables:

- Proyectos en Integration Services, con los paquetes del Staging Area, el Data Warehouse y Base de Datos Multidimensional.
- Generación de scripts de la creación del Staging Area, Data Warehouse y sus respectivas tablas.
- Proyecto multidimensional desarrollado en Analysis Service.
- Manual de instalación y manual de usuario.

6. Conclusiones y Recomendaciones

6.1. Conclusiones

Con el proyecto se han cubierto las necesidades de la Universidad Cenfotec con respecto al módulo de matrícula, las cuales después de varias reuniones con los usuarios finales se pudieron documentar y se realizó un plan de acción que fue implementado para mitigar dichos requerimientos. Después de una serie de investigaciones detalladas se logró la interpretación de la plataforma Class y la base de datos que maneja el módulo mencionado, se pudo entender mucho más a fondo el funcionamiento de la aplicación, y se utilizó este nuevo conocimiento para planificar la mejor solución posible que ayudara a la simplificación del proyecto y que permitiera un producto final amigable con el usuario final y que sea fácil de mantener.

Después de un análisis se pudo desarrollar una solución de inteligencia de negocios que actualmente ayuda a la universidad en la toma de decisiones gracias a un modelo de repositorio de datos, el cual moldea y transforma los datos originales en datos más fáciles de manipular y entendibles para el usuario final. Dado que este repositorio contiene información sensible, también se implementó un plan de seguridad en todas las capas del producto para garantizar que solo las personas autorizadas puedan utilizarlo.

Ya teniendo el repositorio de datos surgió la necesidad de buscar formas de cómo hacer que los datos sean aún más accesibles y que sean de rápido acceso, por lo que también se desarrolló integración con herramientas de visualización que actualmente están ayudando a los usuarios a hacer reportes fáciles de entender que facilitan la toma de decisiones y la interpretación de los datos. Con el conjunto de

entregables descritos anteriormente, la universidad estará mejor preparada para hacerse sentir aún más en el mercado y para optimizar su proceso de matrícula lo mejor posible.

Aparte de la seguridad en el depósito de datos como tal, también se implementó un sistema de jerarquías y roles integrados en las herramientas de visualización, el cual permite que solo los usuarios necesarios puedan ver los datos y es mucho más fácil de mantener la seguridad de la base de datos. Esto con el fin de hacer aún más fácil el mantenimiento de esta y minimizar los recursos humanos requeridos para realizar estas actividades.

Se finaliza con un gran producto que se logró mediante el desarrollo técnico y constante comunicación entre la universidad y los integrantes de este proyecto, se logró la entrega de un producto de calidad el cual brindará grandes beneficios a la universidad, no solo en recurso humano sino también en la mejora de las tomas de decisiones, que pueden hacer crecer aún más a la universidad. Nos parece apropiado recordar a los usuarios finales que este producto requiere atención y mantenimiento; con el refrescamiento de los datos que alimentan al producto y con el mantenimiento de este estamos seguros de que se tendrán resultados formidables y será un sistema que brindará provecho y ganancias por mucho tiempo más.

6.2. Recomendaciones

Dado que el proyecto entregado depende del refrescamiento de la base de datos de origen, se recomienda que al menos una vez al mes se realice el respaldo provisto por la compañía que mantiene la aplicación CLASS, para que así el depósito de datos siga siendo una herramienta de análisis confiable para la universidad.

También se recomienda que se respeten los usuarios de seguridad de la aplicación, para que así se pueda evitar cualquier filtración de información, ya que la base de datos contiene información sensible para la universidad. Es altamente recomendable que haya un equipo de mantenimiento que haga respaldos y velen por la integridad de la base de datos, y en caso de una falla que sean los responsables de devolverla al funcionamiento normal.

7. Reflexiones Finales

Con este proyecto tuvimos la oportunidad de observar el impacto que un depósito de datos puede tener en una organización. Los datos estructurados y fáciles de entender hacen posible la agilidad del análisis de datos y el entendimiento de estos por una organización, lo cual puede llevar a una gran mejoría en la compañía que sabe utilizarlos y que tiene una correcta cultura de datos. Esperamos que con la utilización de este producto la Universidad Cenfotec pueda verse beneficiada y pueda alcanzar nuevas fronteras. Estamos seguros de que este depósito de datos será un activo muy apreciado en la compañía.

8. Trabajos a Futuro

Como parte de la formación y crecimiento de la universidad en el manejo de tecnologías, el proyecto desarrollado abre varias puertas a la inteligencia de negocios, que le permita a la organización establecer sus decisiones conforme los resultados que evidencian los datos. Con la herramienta Class se podría brindar la apertura de nuevos modelos de Data Warehouse, enfocados en el módulo financiero del sistema, obteniendo una herramienta de mayor flexibilidad aparte de los reportes que ya ofrece la herramienta.

De igual manera es importante validar que la universidad maneje de manera manual los registros de los técnicos y cursos libres que esta ofrece mensualmente, por lo cual se podría acoplar al Data Warehouse implementado del módulo de matrícula universitario y expandir el alcance del proyecto, no solo contemplar carreras sino los cursos libres que ofrecen.

9. Referencias

Inmon, B. (1992). *Building the Data Warehouse.*

Kimball, R. (2002). *The Data Warehouse Toolkit.*

Microsoft. (2016, Noviembre). *Microsoft*. Retrieved from https://msdn.microsoft.com/es-es/library/ms174537(v=sql.120).aspx

Microsoft. (2017). *Power BI*. Retrieved from Power BI: https://powerbi.microsoft.com/es-es/

Universidad Cenfotec. (2012). Retrieved from https://www.ucenfotec.ac.cr/

10. Bibliografía

Publicaciones

a. W. H. Inmon. (2005). Building the Data Warehouse, 4th Edition. USA: The Wiley Advantage.
b. Paulraj Ponniah. (2001). Data Warehousing Fundamentals: A Comprehensive Guide for IT Professionals. New York: JOHN WILEY & SONS, INC.
c. Ralph Kimball & Margy Ross. (2002). The Data Warehouse Toolkit Second Edition. Canada: John Wiley & Sons, Inc.

11. Anexos

11.1. Diccionario de datos del Staging Area

La base de datos utilizada como Staging Area contiene un total de 19 tablas, las cuales se utilizaron para el modelo de inteligencia de negocios, con los siguientes campos:

Nombre de la Tabla: ADMISIONES.

Descripción: Contiene los estudiantes admitidos a un plan.

Total de columnas: 9.

Campo	Tipo de Dato	Longitud	Descripción
ADMCUN	Carácter	20	Cif del estudiante.
ADMNIV	Carácter	12	Código de nivel.
ADMPLA	Carácter	12	Código del plan.
ADMFIN	Tiempo	-	Fecha de ingreso.
ADMFEG	Tiempo	-	Fecha de egresado.
ADMFGR	Tiempo	-	Fecha de graduación.
ADMSTS	Carácter	1	Estatus de la admisión.
ADNIDU	Numérico	-	Identificador tabla clientes.

Nombre de la Tabla: BECAS.

Descripción: Contiene catálogo de becas de brinda la universidad.

Total de columnas: 12.

Campo	Tipo de Dato	Longitud	Descripción
BETCOD	Carácter	10	Código beca.
BETDSC	Carácter	50	Descripción tipo beca.
BETCUR	Numérico	10,2	Monto beca asignada curso.
BETMAT	Numérico	10,2	Monto beca matricula.
BETOTR	Numérico	10,2	Monto beca otros rubros.
BETOB1	Carácter	300	Observaciones 1.
BETOB2	Carácter	70	Observaciones 2.
BETOB3	Carácter	70	Observaciones 3.
BETSTS	Carácter	1	Estatus de la beca.
BETSDE	Carácter	2	Aplica beca sobre descuento.
BETPCS	Numérico	10,2	Períodos de castigo en suspensión.
BETMID	Numérico	8,2	Índice académico permitido.
BETUNI	Numérico	-	Identificador del registro.

Nombre de la Tabla: BECADOS.

Descripción: Registro de estudiantes ligados a una beca del centro educativo.

Total de columnas: 15.

Campo	Tipo de Dato	Longitud	Descripción
BECCUN	Carácter	20	Cif del estudiante.
BECCUR	Numérico	10,2	Porcentaje beca asignatura.
BECMAT	Numérico	10,2	Porcentaje beca matricula.
BECTOT	Numérico	10,2	Porcentaje beca materiales.
BECOTR	Numérico	10,2	Beca asignada otros rubros.
BECOB1	Carácter	300	Observaciones 1.
BECSTS	Carácter	1	Estatus de la beca.
BECTIP	Carácter	10	Tipo beca asignada.
BECFEV	Tiempo	-	Fecha vencimiento.
BECNIV	Carácter	10	Código de nivel.
BECIDU	Numérico	-	Identificador tabla clientes.
BECUNI	Numérico	-	Identificador de registro.

Nombre de la Tabla: CARRERA

Descripción: Catálogo de carreras que ofrece la universidad.

Total de columnas: 6.

Campo	Tipo de Dato	Longitud	Descripción
CARDCOD	Carácter	10	Código de la carrera.
CARDSC	Carácter	200	Descripción de la carrera.

CARFAC	Carácter	35	Facultar a la que pertenece.
CARDIR	Carácter	35	Director de la carrera.
CARFAP	Tiempo	-	Fecha apertura de carrera.
CARSTS	Carácter	1	Estatus de la carrera.

Nombre de la Tabla: CLIENTES.

Descripción: Registra todos los integrantes del centro educativo.

Total de columnas: 20.

Campo	Tipo de Dato	Longitud	Descripción
CLICUN	Carácter	20	Cif del estudiante.
CLINAM	Carácter	90	Nombre completo.
CLINOM	Carácter	40	Nombre.
CLIAP1	Carácter	20	Primer apellido.
CLIAP2	Carácter	20	Segundo apellido.
CLITIP	Carácter	1	Tipo CIF.
CLIIDE	Carácter	25	Número de identificación.
CLIFNA	Tiempo	-	Fecha de nacimiento.
CLIPAI	Carácter	3	Nacionalidad.
CLICIV	Carácter	1	Estado civil.
CLISEX	Carácter	1	Sexo.
CLIDIR	Carácter	200	Dirección del estudiante.
CLIPRO	Carácter	15	Provincia.
CLICAN	Carácter	30	Cantón.

CLIDIS	Carácter	30	Distrito.
CLISTS	Carácter	1	Estatus.
CLISUC	Carácter	3	Sucursal del estudiante.
CLITID	Carácter	1	Tipo de identificación.
CLICAR	Carácter	100	Carrera del estudiante.
CLIIDU	Numérico	-	Identificador de registro.
CLITE1	Numérico	-	Teléfono del estudiante.
CLIEM1	Carácter	50	Correo Electrónico de registro.

Nombre de la Tabla: CODIGOS.

Descripción: Registro de catálogo tipo código.

Total de columnas: 5.

Campo	Tipo de Dato	Longitud	Descripción
CODREG	Carácter	1	Tipo de registro.
CODCOD	Carácter	10	Código de registro.
CODDSC	Carácter	45	Descripción del registro.
CODVAL	Carácter	20	Valor asignado.
CODMAN	Carácter	1	Habilita el cambio de valor manual.

Nombre de la Tabla: CONVASUFI.

Descripción: Registro de convalidaciones.

Total de columnas: 15.

Campo	Tipo de Dato	Longitud	Descripción
CVSID	Carácter	4000	Código de la tabla
CVSCUN	Carácter	20	Cif del estudiante.
CVSCUR	Carácter	10	Curso que ha sido afectado.
CVSCON	Carácter	50	Descripción curso afectado.
CVSUNI	Carácter	50	Universidad procedencia.
CVSOB1	Carácter	60	Observaciones 1.
CVSOB2	Carácter	60	Observaciones 2.
CVSNTA	Carácter	10	Nota de aprobación.
CVSFAP	Tiempo	-	Fecha de apertura.
CVSREG	Carácter	1	Tipo de registro.
CVSPER	Carácter	2	Período en que se registra.
CVSANO	Carácter	4	Año en que se registra.
CVSSTS	Carácter	1	Estatus de la convalidación.
CVSNIV	Carácter	15	Código de nivel.
CVSPLA	Carácter	15	Código de plan.

Nombre de la Tabla: CURSOS.

Descripción: Registra el catálogo de cursos que contiene cada nivel.

Total de columnas: 12.

Campo	Tipo de Dato	Longitud	Descripción
CURID	Carácter	150	Identificador de registro.
CURCOD	Carácter	10	Código del curso.

CURDSC	Carácter	200	Descripción del curso.
CURCAR	Carácter	10	Carrera del curso.
CURNIV	Carácter	10	Nivel del curso.
CURCRE	Numérico	4,0	Créditos asignados al curso.
CURPER	Carácter	15	Tipo período asignado.
CURFAP	Tiempo	-	Fecha de apertura
CURSTS	Carácter	1	Estatus curso.
CURNTA	Carácter	5	Nota aprobación.
CURNTS	Carácter	5	Nota aprobación suficiencia.
CURNTE	Carácter	5	Nota examen extraordinario.

Nombre de la Tabla: GRUPOS

Descripción: Registra el detalle del estudiante durante un curso específico.

Total de columnas: 21.

Campo	Tipo de Dato	Longitud	Descripción
GRPREG	Carácter	1	Tipo registro.
GRPPLA	Carácter	10	Código de plan.
GRPCOD	Carácter	10	Código de aula.
GRPCUR	Carácter	10	Código de curso.
GRPDIA	Carácter	10	Día inscrito.
GRPINI	Carácter	11	Hora inicio curso.
GRPFIN	Carácter	11	Hora fin curso.
GRPDSC	Carácter	35	Descripción.

GRPCUN	Carácter	20	Cif del estudiante.
GRPRES	Carácter	20	Código de profesor.
GRPSTS	Carácter	1	Estatus de registro.
GRPFMT	Tiempo	-	Fecha de matrícula.
GRPPER	Carácter	2	Período académico.
GRPANO	Carácter	4	Año académico.
GRPNTA	Carácter	10	Nota final del curso.
GRPSTN	Carácter	1	Estatus de la nota.
GRPNIV	Carácter	10	Código de nivel.
GRPSUC	Carácter	3	Código de sucursal
GRPSEN	Carácter	1	Estatus aprobación del curso.
GRPUNI	Numérico	-	Identificador de Registro.
GRPUNH	Numérico	-	Identificador Tabla Hraula

Nombre de la Tabla: HOMOLOGACION

Descripción: Almacena los cursos que equivalen al mismo en diferentes niveles.

Total de columnas: 8.

Campo	Tipo de Dato	Longitud	Descripción
HOMCAR	Carácter	10	Código carrera.
HOMNIV	Carácter	10	Código del nivel.
HOMCUR	Carácter	10	Código del curso.
HOMREL	Carácter	10	Código Curso Relacionado.
HOMOB1	Carácter	100	Observaciones 1.

HOMOB2	Carácter	100	Observaciones 2.
HOMOB3	Carácter	100	Observaciones 3.

Nombre de la Tabla: HRAULA

Descripción: Registra los cursos abiertos con su respectivo período.

Total de columnas: 16.

Campo	Tipo de Dato	Longitud	Descripción
HRASUC	Carácter	3	Sucursal a la que pertenece.
HRAPER	Carácter	2	Período en el que está disponible.
HRAANO	Carácter	4	Año en que está disponible el curso.
HRAREG	Carácter	1	Tipo de Registro.
HRADIA	Carácter	10	Día de asistencia a Clases.
HRAINI	Carácter	13	Hora de Inicio del Curso.
HRAFIN	Carácter	13	Hora de fin del Curso
HRARES	Carácter	20	Cif del Profesor Responsable.
HRAMOT	Carácter	10	Código de curso.
HRATPE	Carácter	15	Tipo de Período.
HRAREL	Carácter	15	Letra inicial del día de asistencia.
HRAFDD	Tiempo	-	Fecha Inicio del Curso.
HRAFHH	Tiempo	-	Fecha Final del Curso.
HRAKEY	Carácter	15	Número de grupo.
HRANIV	Carácter	10	Nivel al que pertenece.
HRAUNI	Numérico	-	Identificador del registro.

Nombre de la Tabla: MOTIVO_RETIRO_CONGELAMIENTO

Descripción: Almacena el registro de los estudiantes que se retiraron o congelaron un curso.

Total de columnas: 8.

Campo	Tipo de Dato	Longitud	Descripción
MRCID	Carácter	170	Identificador del registro.
MRCCUR	Carácter	15	Código de curso.
MRCNIV	Carácter	15	Código de nivel.
MRCCUN	Carácter	20	Cif del estudiante.
MRCPER	Carácter	2	Período académico.
MRCANO	Carácter	4	Año académico.
MRCSTS	Carácter	1	Estatus de registro.
MRCOBS	Carácter	200	Observaciones.

Nombre de la Tabla: NIVELES

Descripción: Almacena registro del catálogo de niveles existentes en la universidad.

Total de columnas: 13.

Campo	Tipo de Dato	Longitud	Descripción
NIVCOD	Carácter	10	Código de nivel.
NIVDSC	Carácter	200	Descripción del nivel.
NIVDIR	Carácter	35	Director del nivel académico.

NIVCAR	Carácter	10	Código de carrera
NIVCRE	Numérico	4,2	Cantidad de créditos.
NIVFAP	Tiempo	-	Fecha de apertura.
NIVSTS	Carácter	1	Estatus de registro.
NIVNTA	Carácter	5	Nota aprobación por nivel.
NIVGRA	Carácter	70	Grado académico.
NIVNTB	Numérico	3,0	Nota de beca académica.
NIVMOD	Carácter		Módulo al que pertenece.
NIVCLA	Carácter	10	Clasificado.
NIVUNI	Numérico	-	Identificador del registro.

Nombre de la Tabla: ORIGEN_ACADEMICO

Descripción: Registra el origen académico de cada estudiante.

Total de columnas: 8.

Campo	Tipo de Dato	Longitud	Descripción
ORACUN	Carácter	20	Cif del Estudiante.
ORACOD	Carácter	10	Código Origen Académico.
ORAGRA	Carácter	10	Grado Académico.
ORASTS	Carácter	1	Estatus del Registro.
ORAAGR	Carácter	4	Año de Graduación.
ORAOBS	Carácter	200	Observación.
ORAIDU	Numérico	-	Identificador tabla clientes.

Nombre de la Tabla: PLANES

Descripción: Almacena los planes disponibles.

Total de columnas: 10.

Campo	Tipo de Dato	Longitud	Descripción
PLAID	Carácter	150	
PLACOD	Carácter	10	Código del Plan.
PLADSC	Carácter	100	Descripción del Plan.
PLACAR	Carácter	10	Carrera del Plan.
PLANIV	Carácter	10	Nivel del Plan.
PLACUR	Carácter	10	Curso del Plan
PLAFAP	Tiempo	-	Fecha de Apertura
PLAREG	Carácter	1	Tipo de Registro.
PLAPER	Numérico	5,0	Período de inicio del plan.
PLAANO	Carácter	4	Año del plan.

Nombre de la Tabla: PLEESTUDIO

Descripción: Registra el plan de estudio de cada estudiante.

Total de columnas: 9.

Campo	Tipo de Dato	Longitud	Descripción
PLECUN	Carácter	20	Cif del estudiante.
PLECAR	Carácter	10	Carrera del plan.
PLENIV	Carácter	10	Nivel del plan.
PLETAB	Carácter	10	Código del plan.

PLEFAP	Tiempo	-	Fecha Ingreso estudiante.
PLEFAT	Tiempo	-	Fecha de apertura.
PLESTS	Carácter	1	Estatus del plan.
PLEIDU	Numérico	-	Llave foránea de Clientes.
PLEMOF	Carácter	10	Estatus del registro.

Nombre de la Tabla: PROVINCIAS

Descripción: Almacena el catálogo de provincias.

Total de columnas: 2.

Campo	Tipo de Dato	Longitud	Descripción
PROCOD	Carácter	20	Código de provincia.
PRODSC	Carácter	50	Descripción de la provincia.

Nombre de la Tabla: REQUISITOS

Descripción: Registra los requisitos para los cursos respectivos.

Total de columnas: 6.

Campo	Tipo de Dato	Longitud	Descripción
REQNIV	Carácter	10	Código de nivel.
REQCUR	Carácter	10	Código del curso.
REQCOD	Carácter	10	Código del curso padre.
REQFAP	Tiempo	-	Fecha de inicio de requisito.
REQPLA	Carácter	10	Código de plan.
REQTIP	Carácter	1	Tipo de requisito.

Nombre de la Tabla: SUCURSALES

Descripción: Catálogo de sucursales.

Total de columnas: 7.

Campo	Tipo de Dato	Longitud	Descripción
SUCCOD	Carácter	3	Código de la sucursal.
SUCSDS	Carácter	50	Descripción de la sucursal.
SUCFAP	Tiempo	-	Fecha de apertura de la sucursal.
SUCSTS	Carácter	1	Estatus del registro.
SUCINI	Carácter	3	Iniciales de la sucursal.
SUCDIR	Carácter	200	Dirección de la sede.

11.2. Diccionario de datos del Data Warehouse

La base de datos utilizada como Data Warehouse contiene un total de 19 tablas, las cuales se utilizaron para el modelo de inteligencia de negocios, con los siguientes campos:

Nombre de la Tabla: TD_FECHA

Descripción: Registra el catálogo de fecha.

Total de columnas: 13.

Campo	Tipo de Dato	Longitud	Descripción
FechaID	Numérico	-	Identificador del registro.
Fecha	Tiempo	-	Fecha del registro.

DiaMes	Numérico	-	Día del Mes del registro.
DiaAnno	Numérico	-	Día del Año del registro.
Mes	Numérico	-	Mes del registro.
DescMes	Carácter	50	Descripción del mes.
Anno	Numérico	-	Año del registro.
DescTrimestre	Carácter	11	Descripción del trimestre.
DescSemestre	Carácter	10	Descripción del semestre.
DescCuatrimestre	Carácter	11	Descripción del cuatrimestre.
Semana	Numérico	-	Semana del registro.
EntreSemana_sn	Carácter	2	Indicador si el registro es entre semana o no.
DescPeriodoFiscal	Carácter	9	Descripción del período fiscal.

Nombre de la Tabla: TD_SUCURSAL

Descripción: Registra el catálogo de sucursales.

Total de columnas: 5.

Campo	Tipo de Dato	Longitud	Descripción
SucursalID	Numérico	-	Identificador del registro.
SucursalCódigo	Carácter	5	Código de sucursal.
SucursalDescripcion	Carácter	50	Descripción de sucursal.

SucursalIniciales	Carácter	5	Iniciales de sucursal.
ref_SucursalID	Carácter	5	Referencia tabla SUCURSALES.

Nombre de la Tabla: TD_GRUPO

Descripción: Registra los grupos creados.

Total de columnas: 8.

Campo	Tipo de Dato	Longitud	Descripción
GrupoID	Numérico	-	Identificador de registro.
GrupoHoraInicio	Carácter	11	Hora inicio del grupo.
GrupoHoraFin	Carácter	11	Hora Fin del grupo.
GrupoCifProfesor	Carácter	20	Cif del profesor.
GrupoPeriodoAcademico	Carácter	5	Periodo académico del grupo.
GrupoNotaFinal	Carácter	10	Nota final del estudiante.
GrupoEstatusNota	Carácter	5	Estatus de la nota.
ref_GrupoID	Numérico	-	Referencia a la tabla GRUPOS.

Nombre de la Tabla: TD_CURSO

Descripción: Registra el catálogo de cursos.

Total de columnas: 9.

Campo	Tipo de Dato	Longitud	Descripción
CursoID	Numérico.	-	Identificador del registro.
CursoCodigo	Carácter	10	Código del curso.
CursoDescripción	Carácter	200	Descripción curso.
CursoTipoPeriodo	Carácter	15	Tipo de período del curso.
CursoEstatus	Carácter	5	Estatus del curso.
CursoNotaAprobación	Carácter	5	Nota de aprobación.
CursoNotaSuficiencia	Carácter	5	Nota de Suficiencia.
CursoNotaExtraordinaria	Carácter	5	Nota de Extraordinaria.
ref_CursoID	Carácter	100	Referencia a la tabla CURSOS.

Nombre de la Tabla: TD_PLAN

Descripción: Registra el catálogo de planes.

Total de columnas: 5.

Campo	Tipo de Dato	Longitud	Descripción
PlanID	Numérico	-	Identificador del registro.
PlanCodigo	Carácter	10	Código del plan.
PlanDescripción	Carácter	100	Descripción del plan.
PlanTipoRegistro	Carácter	5	Tipo de registro del plan.
ref_PlanID	Carácter	150	Referencia a la tabla PLANES.

Nombre de la Tabla: TD_NIVEL

Descripción: Registra el catálogo de niveles.

Total de columnas: 8.

Campo	Tipo de Dato	Longitud	Descripción
NivelID	Numérico	-	Identificador del registro
NivelCodigo	Carácter	10	Código del nivel.
NivelDescripcion	Carácter	200	Descripción del nivel.
NivelCreditos	Numérico	4,0	Créditos del nivel.
NivelEstatus	Carácter	5	Estatus del nivel.
NivelGradoAcademico	Carácter	70	Grado académico del nivel.
NivelClasificacion	Carácter	10	Clasificación del nivel.
ref_NivelID	Numérico	-	Referencia a la tabla NIVELES.

Nombre de la Tabla: TD_CONVALIDACION

Descripción: Registra el catálogo de convalidaciones.

Total de columnas: 8.

Campo	Tipo de Dato	Longitud	Descripción
ConvalidacionID	Numérico	-	Identificador del registro.
ConvalidacionDescripcion	Carácter	50	Descripción de la convalidación.

ConvalidacionUniversidad Procedencia	Carácter	50	Universidad de procedencia.
ConvalidacionTipoRegistro	Carácter	5	Tipo de convalidación.
ConvalidacionPeriodo	Carácter	5	Período de la convalidación.
ConvalidacionAno	Carácter	5	Año de la convalidación.
ConvalidacionNota	Carácter	10	Nota obtenida.
ref_ConvalidacionID	Carácter	4000	Referencia a la tabla dbo. CONVASUFI.

Nombre de la Tabla: TD_CLIENTE

Descripción: Registra el catálogo de clientes.

Total de columnas: 19.

Campo	Tipo de Dato	Longitud	Descripción
ClienteID	Numérico	-	Identificador de la tabla.
ClienteIdentificacion	Carácter	20	Identificación.
ClienteNombreCompleto	Carácter	90	Nombre completo.
ClienteNombre	Carácter	40	Nombre del estudiante.
ClienteAppelido1	Carácter	20	Primer apellido.
ClienteApellido2	Carácter	20	Segundo apellido.

ClienteTipoIdentificacion	Carácter	5	Tipo de identificación.
ClienteFechaNacimiento	Tiempo	-	Fecha de nacimiento.
ClientePais	Carácter	5	País del estudiante.
ClienteEstadoCivil	Carácter	5	Estado civil del estudiante.
ClienteSexo	Carácter	5	Sexo del estudiante.
ClienteProvincia	Carácter	15	Provincia del estudiante.
ClienteCanton	Carácter	30	Cantón del estudiante.
ClienteDistrito	Carácter	30	Distrito del estudiante.
ClienteEstatus	Carácter	5	Estatus del estudiante.
ClienteTelefono	Carácter	20	Teléfono del estudiante.
ClienteCorreoElectronico	Carácter	100	Correo electrónico del estudiante.
ClienteGradoAcademico	Carácter	10	Grado académico.
ClienteOrigenAcademico	Carácter	200	Origen académico.
ref_ClienteID	Numérico	-	Referencia a la tabla dbo. CLIENTES.

Nombre de la Tabla: TD_BECA

Descripción: Registra el catálogo de Becas.

Total de columnas: 11.

Campo	Tipo de Dato	Longitud	Descripción
BecaID	Numérico	-	Identificador del registro
BecaCodigo	Carácter	10	Código de la beca.
BecaDescripcion	Carácter	50	Descripción de la beca.
BecaPorcentajeCurso	Numérico	10,2	Porcentaje Beca del curso.
BecaPorcentajeMatricula	Numérico	10,2	Porcentaje beca de matrícula.
BecaPorcentajeMateriales	Numérico	10,2	Porcentaje beca de materiales.
BecaPorcentajeOtrosRubros	Numérico	10,2	Porcentaje beca de otros rubros.
BecaObservaciones	Carácter	300	Observaciones de la beca.
BecaEstatus	Carácter	5	Estatus de la beca.
BecaMnimoIndiceAcademico	Numérico	8,2	Mínimo índice académico.
ref_BecaID	Numérico	-	Referencia a la tabla dbo. BECAS.

Nombre de la Tabla: TD_CARRERA

Descripción: Registra el catálogo de carreras.

Total de columnas: 4.

Campo	Tipo de Dato	Longitud	Descripción
CarreraID	Numérico	-	Identificador del registro.
CarreraCodigo	Carácter	10	Código de la carrera.
CarreraDescripcion	Carácter	200	Descripción de la carrera.
ref_CarreraID	Carácter	10	Referencia a la tabla dbo. CARRERAS.

Nombre de la Tabla: TD_MOTIVORETIRO

Descripción: Almacena el registro de motivos de retiro.

Total de columnas: 4.

Campo	Tipo de Dato	Longitud	Descripción
MotivoRetiroID	Numérico	-	Identificador del registro.
MotivoRetiroDescripcion	Carácter	200	Descripción del retiro.
MotivoRetiroPeriodo	Carácter	5	Período del retiro.
MotivoRetiroAno	Carácter	5	Año del retiro.
MotivoRetiroEstatus	Carácter	5	Estatus del retiro.
ref_MotivoRetiroID	Carácter	170	Referencia a la tabla dbo. MOTIVO_RETIRO_CONGELAMIENTO

Nombre de la Tabla: TH_MATRICULA

Descripción: Tabla de hechos.

Total de columnas: 14.

Campo	Tipo de Dato	Tamaño	Descripción
CarreraID	Numérico	-	Referencia a la dimensión TD_CARRERA.
NivelID	Numérico	-	Referencia a la dimensión TD_NIVEL
CursoID	Numérico	-	Referencia a la dimensión TD_CURSO.
PlanID	Numérico	-	Referencia a la dimensión TD_PLAN.
SucursalID	Numérico	-	Referencia a la dimensión TD_SUCURSAL.
BecaID	Numérico	-	Referencia a la dimensión TD_BECA.
FechaIngresoID	Numérico	-	Referencia a la dimensión TD_FECHA.
FechaInicioCursoID	Numérico	-	Referencia a la dimensión TD_ FECHA.

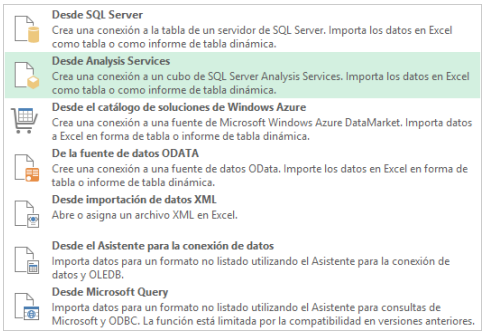
FechaFinCursoID	Numérico	-	Referencia a la dimensión TD_ FECHA.
FechaMatriculaID	Numérico	-	Referencia a la dimensión TD_ FECHA.
ConvalidacionID	Numérico	-	Referencia a la dimensión TD_CONVALIDACION.
MotivoRetiroID	Numérico	-	Referencia a la dimensión TD_MOTIVORETIRO.
GrupoID	Numérico	-	Referencia a la dimensión TD_GRUPO.
ClienteID	Numérico	-	Referencia a la dimensión TD_CLIENTE.

11.3. Manual de Usuario

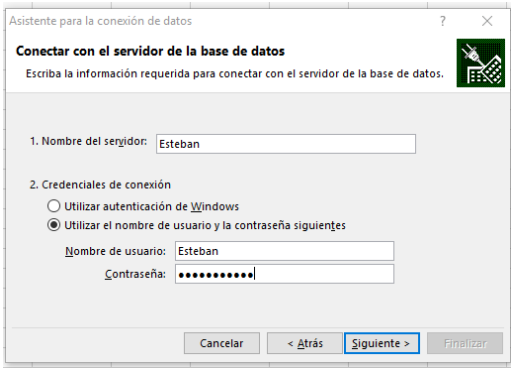
Para acceder a los datos de la base de datos multidimensional y realizar los análisis del negocio se encontraron dos métodos que se acoplan las necesidades de la universidad y se apegan a su presupuesto. Las dos opciones mencionadas son Excel y Power BI, a continuación explicaremos cómo hacer la conexión con cada uno de los métodos y cómo realizar análisis con los datos.

11.3.1. Herramienta Microsoft Excel

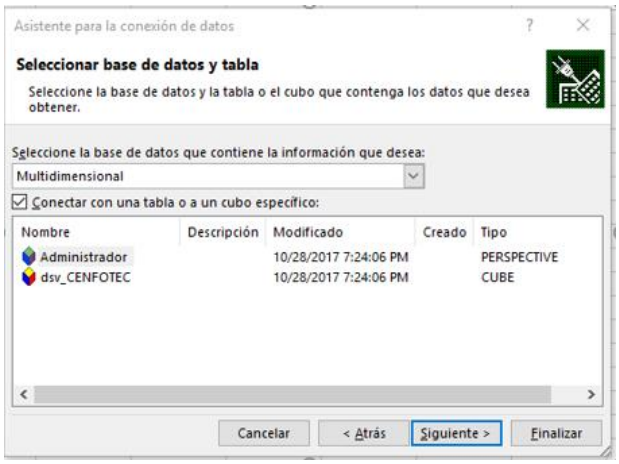
1. En una hoja nueva de Excel, nos vamos a la pestaña de datos en donde podemos ver la opción de “De otras fuentes”, bajo esa pestaña seleccionamos la opción de usar Analysis Services como fuente de datos.



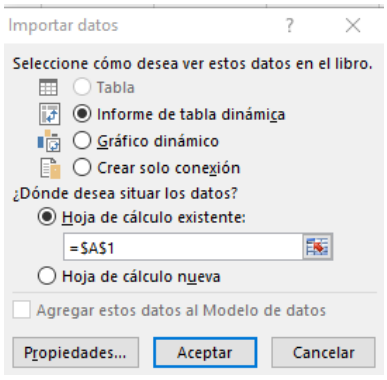
2. Tras seleccionar esa opción aparecerá una sub pantalla, la cual pedirá la información del servidor y un usuario válido del mismo.



3. Al tener la información del usuario correctamente seleccionamos siguiente y nos saldrán los roles que tenemos disponibles para nuestro usuario.



4. Seleccionamos el rol respectivo y le damos siguiente.
5. Una vez teniendo la conexión realizada y funcionando Excel nos preguntará qué tipo de gráfico o tabla se usará para el análisis de los datos.



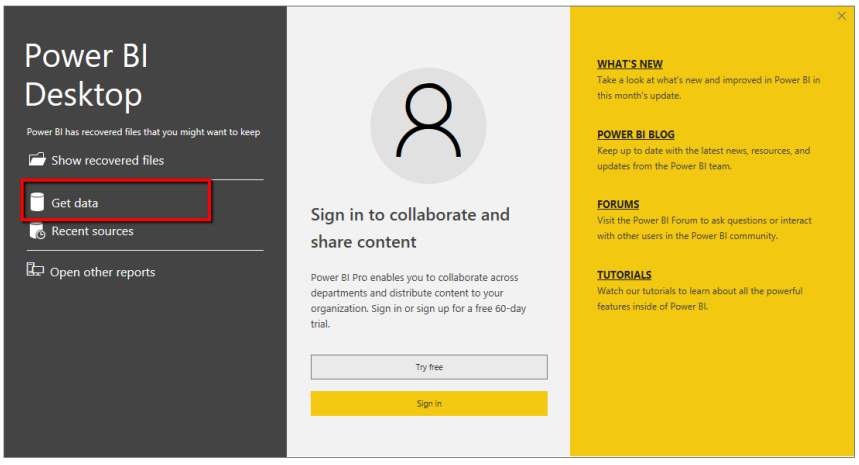
6. Tanto la tabla dinámica como el gráfico dinámico se usan de la misma forma, por lo que para este ejemplo usaremos la tabla dinámica.
7. Al seleccionar el gráfico o la tabla al lado derecho de la pantalla aparecerán todas las columnas de la base de datos multidimensional. Usar esas columnas es tan fácil como solo seleccionarlas para que aparezcan en la

tabla, también se puede elegir que columnas serán los filtros y cuáles serán las que aparecerán en la tabla como desglose.

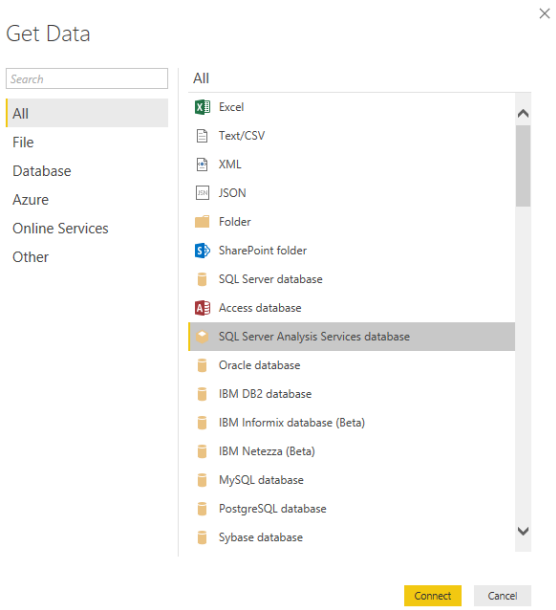


11.3.2. Microsoft Power BI

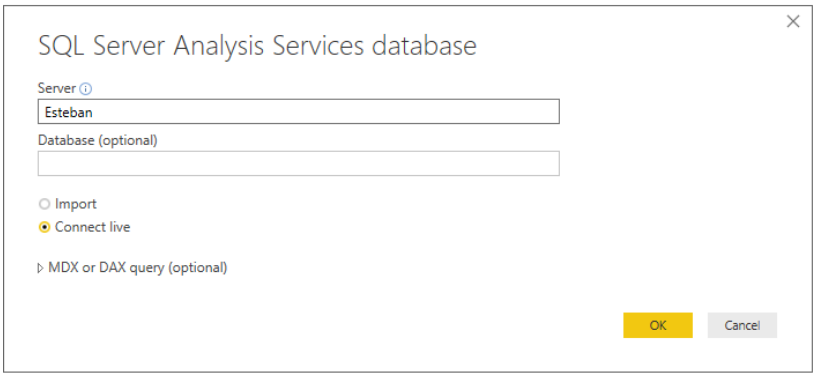
1. Al igual que Excel, Power BI necesita una conexión para poder generar lo deseado por el usuario. En la primera pantalla después de abrir el sistema seleccionamos la opción de: Obtener Datos.



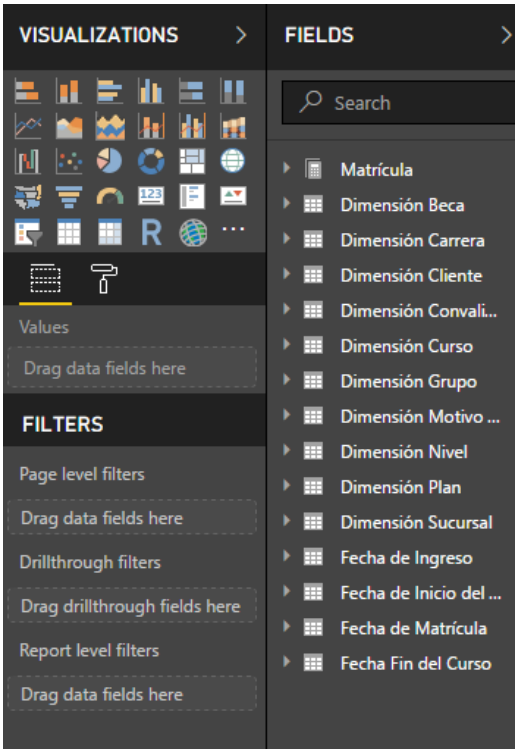
2. A continuación, nos saldrán todas las posibles conexiones que podemos hacer, como con Excel utilizaremos la conexión con Analysis Services.



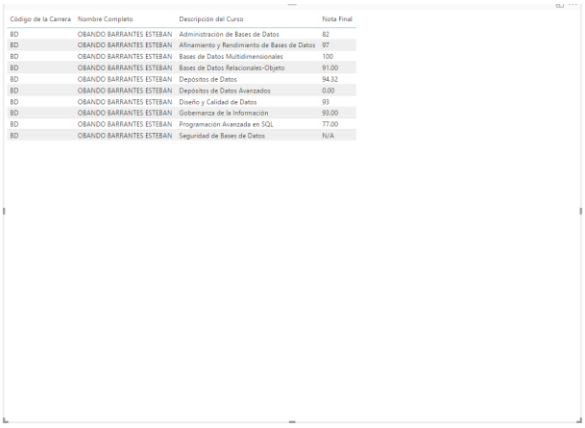
3. A continuación, seleccionamos el servidor al que queremos conectar y el tipo de conexión que queremos tener. Live es para que solo extraiga, en tiempo real, de la base de datos los datos necesarios para completar el reporte requerido, y la opción de import es para importar todos los datos de la base de datos.

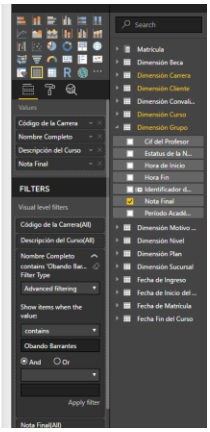


4. Tal y como vimos en Excel, Power BI también cuenta con todas las columnas disponibles de lado derecho de la pantalla, aunque cuenta con más opciones de gráficos y tablas para crear un reporte o dashboard y las diferentes operaciones que se le pueden aplicar a los datos.



5. Basta con seleccionar los campos que se requieren para el proyecto y los filtros que se desean aplicar para generar un gráfico o tabla.

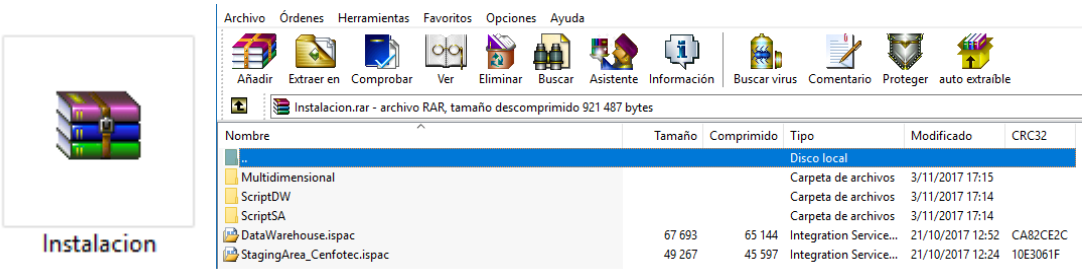




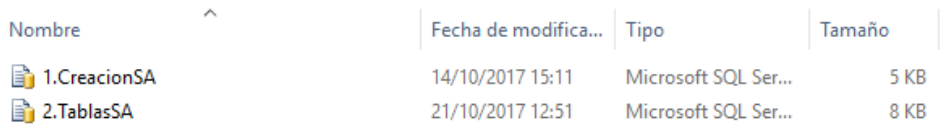
11.4. Manual de Instalación

Para realizar la instalación y proceder al correcto funcionamiento del Data Warehouse, es requerido seguir las siguientes instrucciones:

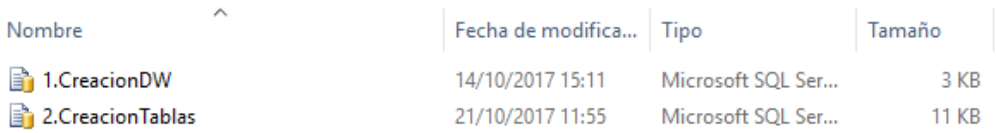
1. Descomprimir el archivo donde se encuentran almacenadas las fuentes de los proyectos y los scripts requeridos para generar las bases de datos. Este archivo incluye los procesos del Staging Area, Data Warehouse y Base de datos Multidimensional.



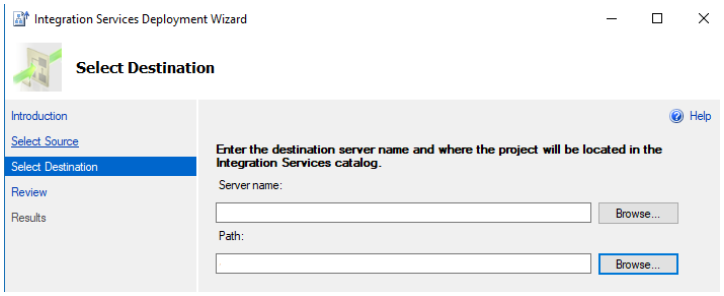
2. En el gestor de base de datos, se debe crear la base de datos del Staging Area y correr la creación de las tablas de esta, esto mediante los scripts almacenados en la carpeta ScriptSA.



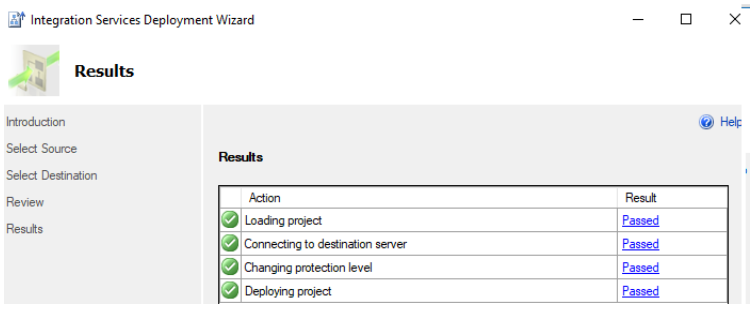
3. De igual manera, en el gestor de base de datos, vamos a correr los scripts de la carpeta ScriptDW, que permite crear la base de datos y tablas que van a almacenar la información del Data Warehouse.



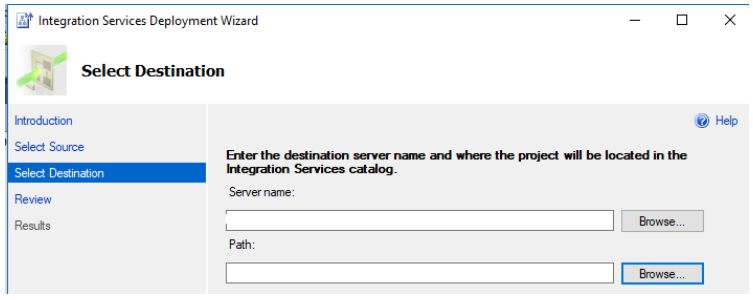
4. Se procede a ejecutar el archivo StagingArea_Cenfotec.ispac. Se selecciona el nombre del servidor de base de datos donde se encuentra el Staging Area almacenado. Este archivo nos permitirá ejecutar las cargas que llenan nuestro Staging Area.



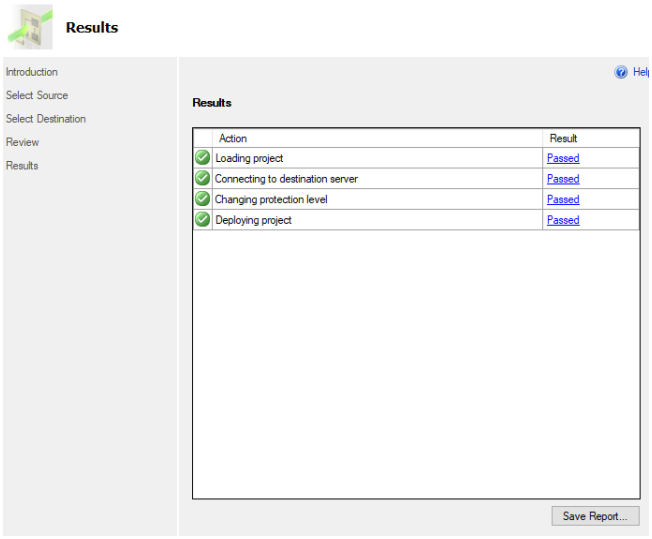
5. Una vez seguidos todos los pasos, el archivo se inserta correctamente en el motor de base de datos.



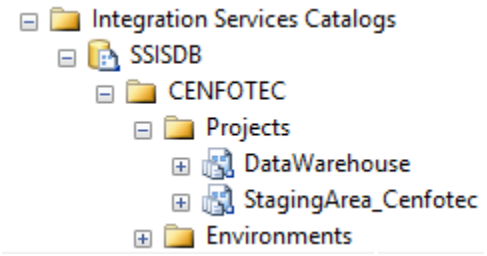
6. Se continúa ejecutando el archivo DataWarehouse.ispac. Se ingresa el nombre del servidor donde se encuentra la base de datos del Data Warehouse.



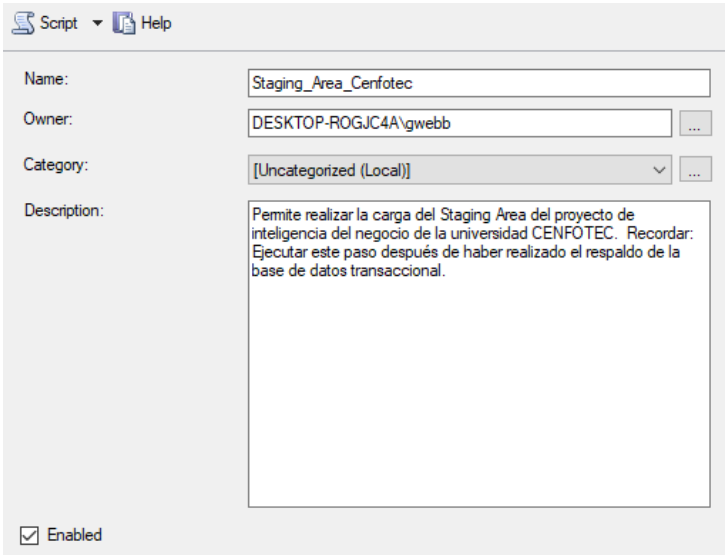
7. Una vez ingresados los datos, el archivo se inserta correctamente en el servidor indicado.



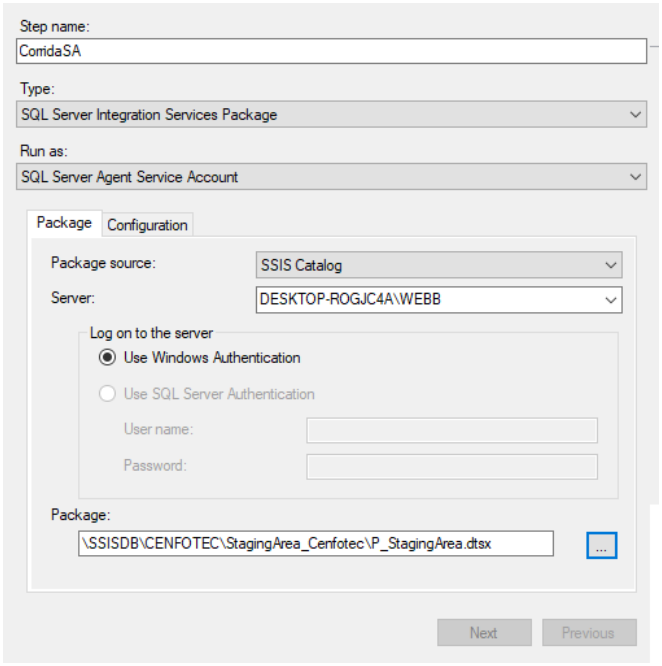
8. Los archivos se han ingresado correctamente al servidor y se visualizan de la siguiente manera.



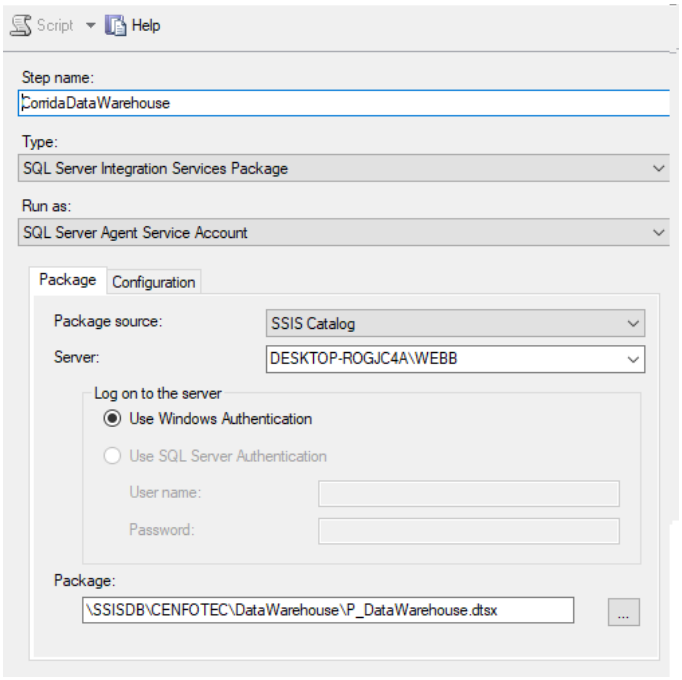
9. Se procede con la configuración del Job, este permite que las cargas del Data Warehouse y del Staging Area de ejecuten ya sea por demanda o de forma automática eligiendo fechas con hora específica para su ejecución. Para ello vamos a la herramienta SQL Agent que nos ofrece el motor SQL Server e ingresamos un nuevo Job. Al ingresar la opción del Job ingresamos un nombre, especificamos el servidor encargado y una descripción del proceso deseado.



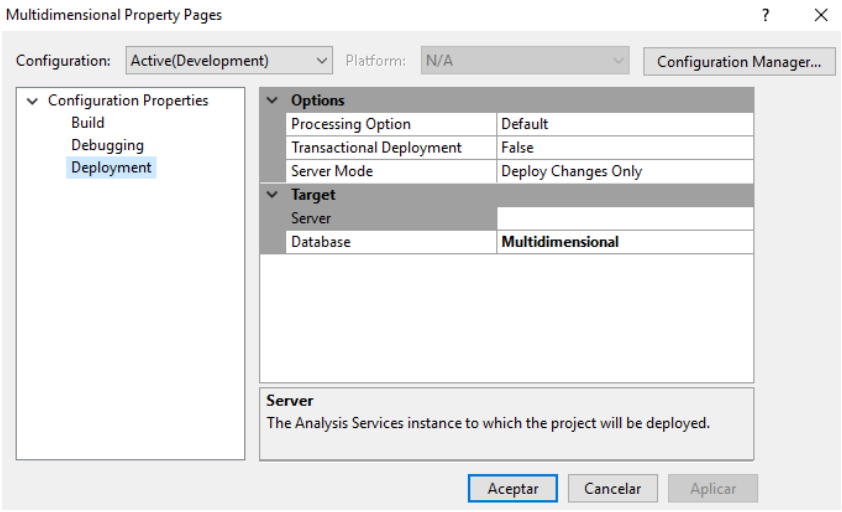
10. Se ingresa a la casilla de pasos del Job, donde se ingresa el nombre del paso, se selecciona el paquete de SQL Server Integration Services y se selecciona el paquete que va a ejecutar el Job. Se deben crear pasos tanto para el Staging Area como para el Data Warehouse.



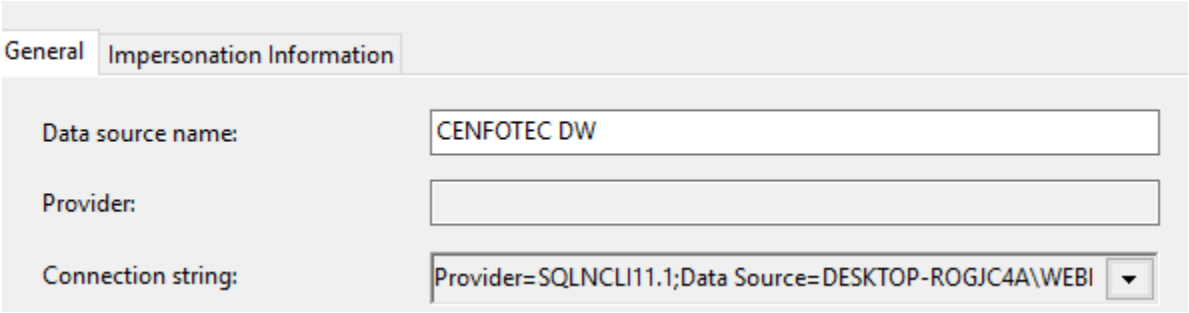
Se deben crear pasos tanto para el Staging Area como para el Data Warehouse.



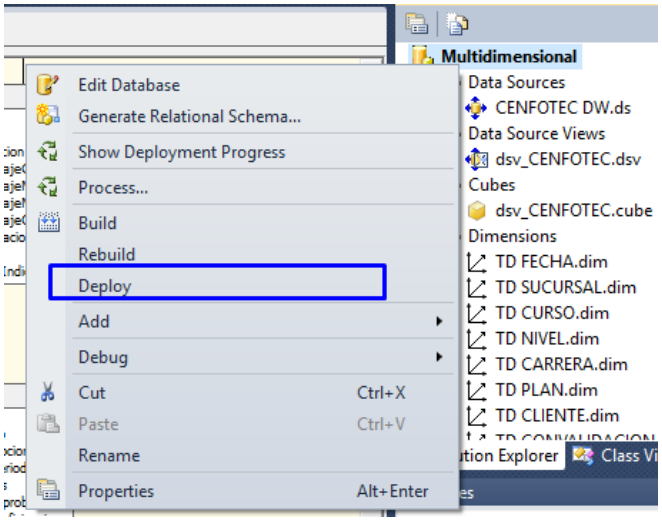
11. Para la base de datos multidimensional abrimos el proyecto respectivo, se ingresa el correcto servidor



12. Se establecen las credenciales del servidor ingresado.



13. Se ingresa la opción de Deploy para ejecutar el proyecto respectivo.



14. Por último, tomaremos el archivo. ispac llamado Correr Multidimensional, el cual nos permite actualizar y generar la multidimensional cada vez que sea requerido. De igual manera seleccionamos al servidor donde vamos a ingresarla y generamos el respectivo Job para su ejecución (Repetir pasos del 4 al 10).

12. Apéndice

12.1. Cronograma

Nombre de tarea	Duración	Comienzo	Fin	Predecesoras	Nombres de los recursos
Proyecto Inteligencia de Negocios Universidad Cenfotec	**48,14 días**	**lun 27/2/17**	**mié 20/12/17**		
Análisis	**4,07 días**	**lun 27/2/17**	**mié 22/3/17**		
Inicio de Análisis	0 días	lun 27/2/17	lun 27/2/17		
Firma de NDA	0 días	lun 27/2/17	lun 27/2/17	2	
Atención Consultas Sistema Class	2 horas	lun 27/2/17	lun 27/2/17	3	Gerald Webb; Esteban Obando
Acceso Respaldo Base de Datos	1 hora	lun 27/2/17	lun 27/2/17	4	Gerald Webb
Lectura Diccionario de Datos	10 horas	mié 1/3/17	lun 6/3/17	5	Gerald Webb
Atención Consultas Módulo de Matricula	2 horas	mié 8/3/17	mié 8/3/17	6	Esteban Obando
Elaboración Plan de Trabajo	15 horas	mié 8/3/17	sáb 18/3/17	7	Esteban Obando
Presentación Análisis de Requerimiento	5 horas	lun 20/3/17	mié 22/3/17	8	Gerald Webb
Aprobación de Requerimiento	0 horas	mié 22/3/17	mié 22/3/17	9	
Documentación	**22,67 días**	**mié 22/3/17**	**lun 14/8/17**		
Descripción del Proyecto	0,5 horas	mié 22/3/17	mié 22/3/17	10	Esteban Obando
Descripción Objetivos del Proyecto	3 horas	mié 22/3/17	sáb 25/3/17	12	Esteban Obando
Descripción Limitaciones y Alcances del Proyecto	1 hora	sáb 25/3/17	sáb 25/3/17	13	Esteban Obando
Marco de Referencia Organizacional y Socioeconómico	1 hora	sáb 25/3/17	lun 27/3/17	14	Esteban Obando
Estado de la Cuestión	3 horas	lun 27/3/17	mié 29/3/17	15	Esteban Obando
Marco Conceptual	3 horas	mié 29/3/17	sáb 1/4/17	16	Gerald Webb
Marco Metodológico	2 horas	sáb 1/4/17	sáb 1/4/17	17	Esteban Obando
Arquitectura de la Solución	2 horas	sáb 1/4/17	lun 3/4/17	18	Gerald Webb
Descripción del Negocio	3 horas	lun 3/4/17	mié 5/4/17	19	Gerald Webb
Creación de Requerimientos	26 horas	mié 5/4/17	mié 26/4/17	20	Esteban Obando
Definición de Recursos	0,5 horas	mié 26/4/17	mié 26/4/17	21	Gerald Webb
Análisis de Fuente de Datos	14 horas	sáb 29/4/17	lun 8/5/17	22	Gerald Webb
Arquitectura de Datos	3 horas	mié 10/5/17	mié 10/5/17	23	Esteban Obando
Herramientas por Utilizar	2 horas	sáb 13/5/17	sáb 13/5/17	24	Esteban Obando
Análisis de Factibilidad	3 horas	sáb 13/5/17	lun 15/5/17	25	Gerald Webb
Etapa de Diseño y Desarrollo	45 horas	lun 15/5/17	mié 14/6/17	26	Gerald Webb
Proceso de Detección de Tablas del Módulo de Matricula	30 horas	sáb 17/6/17	mié 5/7/17	27	Esteban Obando
Definición de Modelo del Módulo de Matricula	4 horas	sáb 8/7/17	sáb 8/7/17	28	Esteban Obando
Definición de Dimensiones y Tabla de Hechos	15 horas	lun 10/7/17	mié 19/7/17	29	Gerald Webb
Generación del Modelo de Data Warehouse	12 horas	mié 19/7/17	sáb 29/7/17	30	Esteban Obando
Estrategias de ETL	5 horas	sáb 29/7/17	lun 31/7/17	31	Esteban Obando
Diseño de Seguridad de la Solución	2 horas	lun 31/7/17	sáb 5/8/17	32	Esteban Obando
Diseño de Procesos	4 horas	sáb 5/8/17	lun 7/8/17	33	Gerald Webb
Diccionario de Datos	10 horas	lun 7/8/17	lun 14/8/17	34	Gerald Webb
Entregables	1 hora	lun 14/8/17	lun 14/8/17	35	Esteban Obando
Desarrollo	**12,56 días**	**lun 14/8/17**	**sáb 28/10/17**		
Inicio de Desarrollo	0 horas	lun 14/8/17	lun 14/8/17	36	
Staging Area	**2,67 días**	**lun 14/8/17**	**mié 30/8/17**		
Proceso de Restablecer Backup	0,5 horas	lun 14/8/17	lun 14/8/17	38	Esteban Obando
Definición de Base de Datos	4 horas	lun 14/8/17	sáb 19/8/17	40	Gerald Webb
Depuración de Campos	3 horas	sáb 19/8/17	sáb 19/8/17	41	Esteban Obando
Creación de Estructura	8 horas	sáb 19/8/17	sáb 26/8/17	42	Gerald Webb
Archivos de Limpieza	2,5 horas	sáb 26/8/17	sáb 26/8/17	43	Esteban Obando

Proceso de Generación de ETL	4 horas	lun 28/8/17	mié 30/8/17	44	Esteban Obando
Generación de Solución	1 hora	mié 30/8/17	mié 30/8/17	45	Gerald Webb
Data Warehouse	**6,51 días**	**mié 30/8/17**	**lun 9/10/17**		
Creación de Base de Datos	12 horas	mié 30/8/17	sáb 9/9/17	46	Esteban Obando
Creación de Proyecto	24 horas	sáb 9/9/17	lun 25/9/17	48	Gerald Webb
Generación de Dimensiones	5 horas	lun 25/9/17	mié 27/9/17	49	Esteban Obando
Validación de Dimensión de Fechas	4 horas	sáb 30/9/17	sáb 30/9/17	50	Esteban Obando
Validación de Registros	4 horas	lun 2/10/17	mié 4/10/17	51	Esteban Obando
Generación de Tabla de Hechos	6 horas	mié 4/10/17	sáb 7/10/17	52	Gerald Webb
Generación de Solución	1 hora	lun 9/10/17	lun 9/10/17	53	Gerald Webb
Base de Datos Multidimensional	**3,37 días**	**lun 9/10/17**	**sáb 28/10/17**		
Creación de Proyecto	6 horas	lun 9/10/17	sáb 14/10/17	54	Esteban Obando
Creación del Cubo	5 horas	sáb 14/10/17	lun 16/10/17	56	Esteban Obando
Generación de Jerarquías	8 horas	lun 16/10/17	sáb 21/10/17	57	Esteban Obando
Generación de Perspectivas	5 horas	lun 23/10/17	mié 25/10/17	58	Gerald Webb
Generación de Solución	1 hora	mié 25/10/17	mié 25/10/17	59	Esteban Obando
Herramienta de modelado	4 horas	sáb 28/10/17	sáb 28/10/17	60	Gerald Webb
Pruebas	**6,51 días**	**lun 30/10/17**	**mié 6/12/17**		
Definición de Ambiente de Pruebas	12 horas	lun 30/10/17	lun 6/11/17	61	Gerald Webb
Inicio de Pruebas	0 horas	lun 6/11/17	lun 6/11/17	63	
Pruebas Técnicas	**4,65 días**	**lun 6/11/17**	**lun 4/12/17**		
Elaboración de Matriz de Pruebas	20 horas	lun 6/11/17	lun 20/11/17	64	Esteban Obando
Ejecutar Matriz de Pruebas	20 horas	lun 20/11/17	lun 4/12/17	66	Gerald Webb
Pruebas de Usuario	**0,47 días**	**lun 4/12/17**	**mié 6/12/17**		
Ejecución de Pruebas	3 horas	lun 4/12/17	mié 6/12/17	67	
Acompañamiento de Pruebas	4 horas	lun 4/12/17	mié 6/12/17	69CC	Gerald Webb; Esteban Obando
Aprobación de Pruebas de Usuario	0 horas	mié 6/12/17	mié 6/12/17	70	
Capacitación	**0,23 días**	**sáb 9/12/17**	**sáb 9/12/17**		
Demostración	2 horas	sáb 9/12/17	sáb 9/12/17	71	Gerald Webb; Esteban Obando
Manuales	**0,93 días**	**sáb 9/12/17**	**mié 13/12/17**		
Creación Manual de Usuario	4 horas	sáb 9/12/17	lun 11/12/17	73	Esteban Obando
Creación Manual de Instalación	4 horas	lun 11/12/17	mié 13/12/17	75	Esteban Obando
Implantación	**1,16 días**	**sáb 16/12/17**	**mié 20/12/17**		
Elaboración Documento de Pase a Producción	3 horas	sáb 16/12/17	sáb 16/12/17	76	Gerald Webb
Ejecución Pase a Producción	4 horas	sáb 16/12/17	lun 18/12/17	78	Gerald Webb; Esteban Obando
Soporte Post	3 horas	mié 20/12/17	mié 20/12/17	79	Esteban Obando
Fin del Requerimiento	0 horas	mié 20/12/17	mié 20/12/17	80	