UNIVERSIDAD TÉCNICA ESTATAL DE QUEVEDO.

FACULTAD DE CIENCIAS DE LA INGENIERÍA.

CARRERA INGENIERÍA EN SISTEMAS.

Proyecto de Investigación previo a la obtención del título de Ingeniero en Sistemas.

Título del Proyecto de Investigación:

"DESARROLLO DE UNA METODOLOGÍA DE ALIMENTACIÓN DE DATOS PARA EL CUADRO DE MANDO INTEGRAL DE LA UTEQ"

Autor:

Marcos Diego García Coello.

Director del Trabajo de Investigación.

Ing. Washington Chiriboga Casanova.

Quevedo – Los Ríos – Ecuador.

2015.

DESARROLLO DE UNA METODOLOGÍA DE ALIMENTACIÓN DE DATOS PARA EL CUADRO DE MANDO INTEGRAL DE LA UTEQ

2015

DECLARACIÓN DE AUTORÍA Y CESIÓN DE DERECHOS.

Yo, **MARCOS DIEGO GARCIA COELLO,** declaro que el trabajo aquí descrito es de mi autoría, que no ha sido presentado previamente para ningún grado, que he consultado las referencias bibliográficas que se incluyen en este documento titulado **"Desarrollo de una metodología de alimentación de datos para el Cuadro de Mando Integral de la UTEQ",** es de mi exclusiva responsabilidad.

La Universidad Técnica Estatal de Quevedo, puede hacer uso de los derechos correspondiente a este trabajo, según lo establecido por la ley de Propiedad Intelectual, por su Reglamento y por la normativa institucional vigente.

Marcos Diego García Coello.
C.C. 1205736331

CERTIFICACIÓN DE CULMINACIÓN DEL PROYECTO DE INVESTIGACIÓN.

El suscrito, **Ing. WASHINGTON CHIRIBOGA CASANOVA**, Docente de la Universidad Técnica Estatal de Quevedo, certifica que el estudiante **MARCOS DIEGO GARCÍA COELLO**, realizó el Proyecto de Investigación de grado **"Desarrollo de una metodología de alimentación de datos para el Cuadro de Mando Integral de la UTEQ"**, previo a la obtención del título de **INGENIERÍA EN SISTEMAS**, bajo mi dirección, habiendo cumplido con las disposiciones reglamentarias establecidas para el efecto.

Ing. Washington Chiriboga.

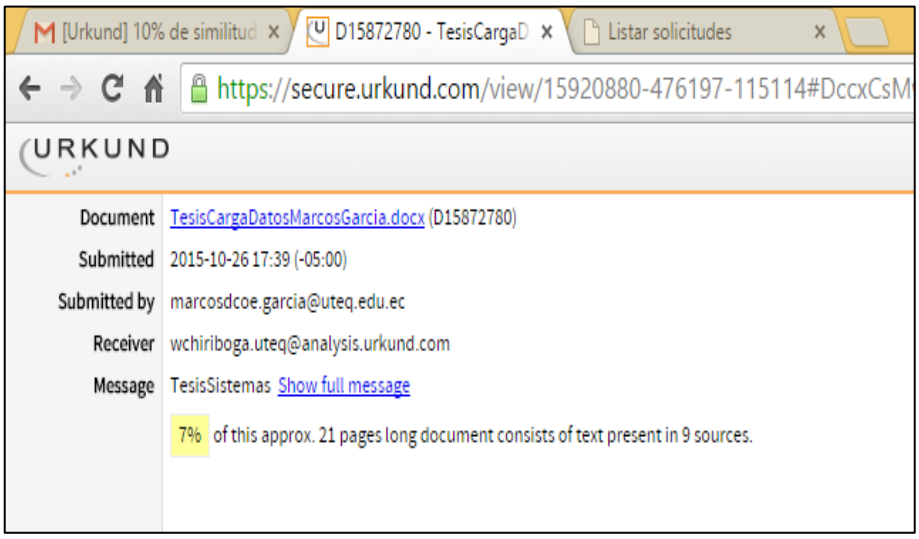
C.C. 1202869838

Director de Trabajo de Investigación

CERTIFICADO DEL REPORTE DE LA HERRAMIENTA DE PREVENCIÓN DE COINCIDENCIA Y/O PLAGIO ACADÉMICO.

En calidad de Director del Trabajo de investigación titulado: **"Desarrollo de un software de alimentación de datos para el Cuadro de Mando Integral de la UTEQ"**, me permito manifestar a usted y por intermedio del Consejo Directivo lo siguiente:

Que, el señor **MARCOS DIEGO GARCÍA COELLO,** estudiante de la carrera de Ingeniería en Sistemas, ha cumplido con las correcciones pertinentes e ingresado su trabajo de investigación a la herramienta de prevención de coincidencia y/o plagio académico (URKUND), tengo bien certificar la siguiente información sobre el informe del sistema reflejando un porcentaje del 7%, como se observa en la figura siguiente:



Ing. Washington Chiriboga

Director de Trabajo de Investigación

UNIVERSIDAD TÉCNICA ESTATAL DE QUEVEDO.

FACULTAD DE CIENCIAS DE LA INGENIERÍA.

CARRERA DE INGENIERÍA EN SISTEMAS.

PROYECTO DE INVESTIGACIÓN.

Título.

"Desarrollo de una metodología de alimentación de datos para el Cuadro de Mando Integral de la UTEQ"

Presentado a la Comisión Académica como requisito precio a la obtención del título de Ingeniero en Sistemas.

PRESIDENTE DEL TRIBUNAL DE TESIS.

Ing. Iván Jaramillo Chuqui

MIEMBRO DEL TRIBUNAL DE TESIS.

Ing. Jorge Saa Saltos

MIEMBRO DEL TRIBUNAL DE TESIS.

Ing. Kenya Guerrero

QUEVEDO – LOS RIOS – ECUADOR.

2015

AGRADECIMIENTO.

A mi padre Francisco García y a mi madre Magali Coello, por todo su apoyo incondicional, durante todo el proceso.

A mis hermanos Juan y Francisco que siempre estuvieron dando su apoyo cuando lo necesitaba.

Al director del proyecto el Ing. Washington Chiriboga por todo el apoyo recibido durante el desarrollo de este proyecto.

A mis compañeros Jim Yánez, Roselyn Troya y Ruddy Carriel por la ayuda que me brindaron

Diego.

DEDICATORIA.

A mi madre, padre y hermanos, quienes me soportaron en las buenas y en las malas siempre ofreciéndome su apoyo incondicional y cariño, permitiéndome llegar a la culminación de una etapa más de preparación como profesional.

A mis amigos que siempre estuvieron cuando más los necesité, quienes de una u otra forma supieron darme ánimos para seguir en la lucha.

Diego

RESUMEN EJECUTIVO Y PALABRAS CLAVES.

Este proyecto de investigación está orientado a desarrollar una metodología de alimentación de datos para el Cuadro de Mando Integral de la Universidad Técnica Estatal de Quevedo (UTEQ), el cual permitirá mejorar la integración de la información de los sistemas de la universidad por medio de un software.

La metodología de carga de información se fundamenta en un proceso ETL que consiste en extraer la información desde las fuentes de información, la transformación se realiza mediantes las consultas SQL y la carga de información se realiza a la base de datos del Cuadro de Mando Integral, la cual viene a actual de repositorio.

Se dispondrá de una herramienta creada para facilitar la extracción de la información de las fuentes de datos para después cárgala a las variables del Cuadro de Mando Integral, todo este proceso será supervisado por el administrador de conexiones por medio de estados. Además, se aplica un calendario de carga de cada variable, procurando de esta manera mejorar la oportunidad de la información

Las pruebas realizadas permitieron comprobar las cargas de datos desde las fuentes de información utilizando la calendarización de las variables así como las notificaciones por correo electrónico a los usuarios responsables tanto del cumplimiento como del retraso o fallo.

Palabras claves: Carga de datos, ETL, Fuentes de datos, BSC

ABSTRACT AND KEYWORDS.

This research Project is oriented to develop a methodology of information feeding for Balanced Scorecard of the Universidad Técnica Estatal de Quevedo (UTEQ), which will allow improve the integration of the information the University systems by means of a software.

The methodology of information load is based on an ETL process that consists of extracting the information from the information sources, the transformation is realized through SQL queries and the information load is realized to the database of the Balanced Scorecard, which comes to current of repository.

There will be a tool created to facilitate the extraction of information of the data sources for after charge to the variables the Balanced Scorecard, all this process it will be supervised by the connection manager by means of states. Furthermore, applies a calendar of load for every variable, trying this way to improve the opportunity of the information.

The realized tests they allowed check the data loads from the information sources using the scheduling of the variables as well as the notifications by email to the responsible users both the compliance as for the delay or fault

Keywords: Loading data, ETL, Data sources, BSC.

ÍNDICE DE CONTENIDO.

INTRODUCCIÓN. 1
CAPÍTULO I CONTEXTUALIZACIÓN DE LA INVESTIGACIÓN 2
1.1. Problema de la investigación. 3
1.1.1. Planteamiento del problema. 3
1.1.2. Formulación del problema. 4
1.1.3. Sistematización del problema. 4
1.2. Objetivos. 5
1.2.1. Objetivos General. 5
1.2.2. Objetivos Específicos. 5
1.3. Justificación. 6

CAPÍTULO II FUNDAMENTACIÓN TEÓRICA 7
2.1. Marco conceptual. 8
2.1.1. Balance ScoreCard. 8
2.1.1.1. El Balance ScoreCard. 8
2.1.1.2. Característica del Balance ScoreCard 8
2.1.1.3. Componentes esenciales del Balance ScoreCard 8
2.1.1.4. Perspectivas del Balance ScoreCard 9
2.1.1.5. Ventajas de un Balance ScoreCard 9
2.1.2. La información en el Balance ScoreCard. 9
2.1.2.1. Que información se necesita inicialmente. 9
2.1.2.2. Metadatos 11
2.1.2.2. Metadatos en la integración de la información. 11
2.1.2.3. Flujo de información en la implementación del cuadro de mando integral. 11
2.1.2.4. Almacenes de datos. 11
2.1.2.5. Almacenes de datos y procesos ETL. 12
2.1.2.6. Procesos ETL. 12
2.1.2.7. Carga Inicial. 13
2.1.3. Fuentes de información. 14
2.1.4. Tipos de fuentes de información. 14
2.1.5. Estrategias para la integración de la información heterogénea. 15
2.1.6. El entorno de archivo tradicional. 15
2.2. Marco Referencial. 16

CAPÍTULO III MÉTODOLOGIA DE LA INVESTIGACIÓN 17
3.1. Localización. 18
3.2. Tipo de investigación. 18
3.3. Método de investigación 18
3.4. Recursos Humanos y Materiales. 20
3.4.1. Recursos Humano. 20
3.4.2. Materiales. 20
3.5. Método de Software 22
3.5.1. Fases de desarrollo. 22
3.5.2. Fase de exploración. 22
3.5.3. Fase de Planificación de la entrega. 35
3.5.4. Fase de Iteraciones. 36
3.5.4.1. Primera Iteración: Elementos de desarrollo. 36
3.5.4.2. Segunda Iteración: Módulo de definición DBMS. 38
3.5.4.3. Tercera Iteración: Módulo de Fuentes de Datos 41
3.5.4.4. Cuarta Iteración: Módulo de Carga DBMS. 44
3.5.4.5. Quinta Iteración: Módulo de Calendarización. 47
3.5.5. Fase de Producción 50
3.5.6. Fase de Mantenimiento. 51
3.5.7. Fase de muerte de proyecto. 51

CAPÍTULO IV RESULTADOS Y DISCUSIÓN 52
4.1. Resultados y Discusión 53

CAPÍTULO V CONCLUSIONES Y RECOMENDACIONES 58
5.1. Conclusiones. 59
5.2. Recomendaciones. 60

CAPÍTULO VI BIBLIOGRAFÍA 61

CAPÍTULO VII ANEXOS 64

ÍNDICE DE TABLAS

Tabla 1. Hardware 20
Tabla 2. Software 21
Tabla 3. Suministros 21
Tabla 4. Personal 21
Tabla 5. Costo del Proyecto 22
Tabla 6. Historia de usuario inicio de sesión. 24
Tabla 7. Épica Gestionar versiones DBMS. 25
Tabla 8. Historia de usuario Administración de versiones DBMS 26
Tabla 9. Épica Gestionar Fuentes de datos. 26
Tabla 10. Historia de usuario Administrar las fuentes de datos. 27
Tabla 11. Épica Gestionar Cargas DBMS. 29
Tabla 12. Historia de usuario Administración de cargas DBMS. 30
Tabla 13. Historia de Usuario Gestionar cargas de datos. 31
Tabla 14. Historia de usuario Carga manual de datos. 33
Tabla 15. Historia de usuario Notificaciones al usuario. 34
Tabla 16. Elementos de la plataforma de desarrollo 36

ÍNDICE DE FIGURAS.

Figura 1 Procesos ETL 13
Figura 2. Diagrama de Caso de Uso 23
Figura 3. Diagrama de base de datos y su distribución por módulos. 37
Figura 4. Iteraciones entre módulos del sistema 38
Figura 5. Tabla del Módulo de definición DBMS. 40
Figura 6. Diagrama de clases del Módulo de definición DBMS 41
Figura 7. Tabla del Módulo de Fuentes de Datos. 43
Figura 8. Diagrama de clases del Módulo de Fuentes de Datos. 44
Figura 9. Tabla del Módulo de Carga DBMS. 46
Figura 10. Diagrama de clases del Módulo de Carga DBMS. 47
Figura 11. Tabla del Módulo de Calendarización. 49

Figura 12. Diagrama de clases del Módulo de Calendarización. 50
Figura 13. Arquitectura Data Warehouse 53
Figura 14. Metodología de Carga de Datos aplicada al software. 55

ÍNDICE DE GRÁFICOS.

Gráfico 1. Desarrollo de la Segunda Iteración. 39
Gráfico 2. Desarrollo de la Tercera Iteración. 42
Gráfico 3. Desarrollo de la Cuarta Iteración. 45
Gráfico 4. Desarrollo de la Quinta Iteración. 48

ÍNDICE DE ANEXOS.

Anexo 1. Árbol del problema 65
Anexo 2. Matriz de relación Problema, Objetivos 66
Anexo 3. Estructuras de las Historias de usuarios 67
Anexo 4. Formato de solicitud de Mantenimiento del Sistema 68
Anexo 5. Prueba realizada de conexión, extracción y carga de datos de la fuente de datos Biblioteca 69
Anexo 6. CERTIFICACIÓN DE REDACCIÓN. 72
Anexo 7. CERTIFICADO - ABSTRACT 73

Código Dublin.

Título:	Desarrollo de una metodología de alimentación de datos para el Cuadro de Mando Integral de la UTEQ			
Autor:	García Coello Marcos Diego			
Palabras clave.	Carga datos	BSC	Fuentes Datos	DBMS
Fecha Publicación:	2015			
Editorial:	Quevedo: UTEQ, 2015.			
Resumen:	Este proyecto de investigación está orientado a desarrollar una metodología de alimentación de datos para el Cuadro de Mando Integral de la Universidad Técnica Estatal de Quevedo (UTEQ), el cual permitirá mejorar la integración de la información de los sistemas de la universidad por medio de un software. La metodología de carga de información se fundamenta en un proceso ETL que consiste en extraer la información desde las fuentes de información, la transformación se realiza mediantes las consultas SQL y la carga de información se realiza a la base de datos del Cuadro de Mando Integral, la cual viene a actual de repositorio. This research Project is oriented to develop a methodology of information feeding for Balanced Scorecard of the Universidad Técnica Estatal de Quevedo (UTEQ), which will allow improve the integration of the information the University systems by means of a software. The methodology of information load is based on an ETL process that consists of extracting the information from the information sources, the transformation is realized through SQL queries and the information load is realized to the database of the Balanced Scorecard, which comes to current of repository.			
Descripción:	100 hojas, dimensiones, 29 x 21 cm + CD-ROM 6162.			
URI:				

INTRODUCCIÓN.

La información es un recurso que se debe gestionar de la mejor forma posible, resultando trascendental para la toma de decisiones, la aplicación de políticas y estrategias que deben ser monitoreadas constantemente para alcanzar los objetivos estratégicos [1]. La importancia de tener información actualizada y consistente se ha convertido en un factor importante en la todas las instituciones, incluidas la Instituciones de Educación Superior (IES)

Con las exigencias en el desarrollo institucional en las Universidades van aumentando cada vez más con el paso del tiempo. Administrar una institución, en todos sus niveles, se está volviendo más complejo por lo que se requiere la implementación de algún tipo de software que ayude a controlar todo el volumen de información que las instituciones utilizan. La mayoría de las instituciones manejan la información en forma digital [2] por lo que es de suma importancia contar con la capacidad para buscar, recolectar y preservar datos electrónicos en una forma eficiente en tiempo y costo.

El Balance ScoreCard (BSC) es un modelo basado en un mapa estratégico [3], que actualmente se maneja en empresas, instituciones, etc. Al momento de desarrollar un Balanced Scorecard, surge la necesidad de poder alimentar el modelo con los datos, sobre todo cuando dentro de la universidad se trabajan con diferentes sistemas de los cuales se pueden extraer la información que se necesita para los indicadores. Por ello hace falta una metodología que realice esa alimentación al Balance Scorecard de la manera más eficiente posible.

Este trabajo propuso desarrollar un software para conectarse a diferentes bases de datos alojadas en DBMS (Sistema Administrador de Bases de Datos) distintos, permitiendo la extracción de la información necesaria para los distintos indicadores del BSC, según los objetivos planteados. De igual forma, se desarrolló un mecanismo para ingresar datos manuales a las variables de los indicadores que no tienen esa información en las fuentes digitales. Con esto se logró que las Unidades Académicas de la Universidad logren cumplir oportunamente con el envío de la información que se necesita para evaluar y direccionar la Institución evitando también errores o atrasos.

CAPÍTULO I

CONTEXTUALIZACIÓN DE LA INVESTIGACIÓN

1.1. Problema de la investigación.

1.1.1. Planteamiento del problema.

El CEAACES plantea modelos de evaluación a la Instituciones de Educación Superior (IES) lo cuales también deben sujetarse a criterios de calidad, entre uno de los criterios está el conjunto de acciones que llevan a cabo las instituciones, con el fin de garantizar la eficiente y eficaz gestión, aplicables a las carreras, programas académicos, a las instituciones de educación [4].

La Universidad Técnica Estatal de Quevedo realiza la supervisión de su plan estratégico con un Cuadro de Mando Integral desarrollado en una hoja de cálculo (Excel) [5], esto hace que no sea muy dinámico al momento de llevar un control más estricto sobre el avance de logros de su plan estratégico. La información con la que se alimenta el Balance Scorecard proviene de varias unidades de la universidad (Financiera, Estudiantil, Docentes, Graduados y Egresados, Empleados, etc.) que en ocasiones tienen dificultades en el llenado de formularios así como para remitirlos oportunamente.

Cada unidad de la universidad gestiona su propia información almacenada en diversas bases de datos o archivos. Cuando se necesita cargar la información al Balance Scorecard desarrollado en Excel se debía digitar lo que proporcionaba cada Unidad Académica o Administrativa. Durante este proceso se podían generar inconvenientes como atrasos, errores, datos no compatible, etc. Además, se consume tiempo y recursos extrayendo información de sistemas ya en funcionamiento que luego son convertidos en datos simples que se guardan en la hoja electrónica de la Unidad de Planeamiento Estratégico (UPE). La escasa integración de datos es considerada como una debilidad en el momento de evaluar la situación estratégica de una organización.

Aunque la Unidad de Planeamiento Estratégico (UPE) envía comunicaciones para la entrega de información, se han detectado momentos en los que no se cuenta con la cifras requeridas para evaluar algunos indicadores (Hecho observado en la presentación de resultados del Plan Estratégico 2012 – 2015, 11 de Noviembre del 2105). Esto refleja la

necesidad de implementar recordatorios y acciones correctivas para que este ingreso sea cumplido con mayor efectividad.

Al centralizar la información en un solo repositorio estratégico se requiere que aquellos datos que no se pueden extraer de los sistemas de información sean digitados dentro del mismo almacén a fin de evitar los errores de digitación que ocurren al ser remitidos por cada Unidad a la UPE. Además, es necesario que quien ingrese la información asuma la responsabilidad como fuente de datos.

Para el requerimiento del plan estratégico que la Universidad Técnica Estatal de Quevedo se necesita de una herramienta más dinámica que facilite la integración de toda la información y la relacione con los indicadores alineados a cada objetivo estratégico. Se tiene información sobre la actual construcción de un software BSC, sim embargo el proceso de alimentación al mismo aún es una dificultad a superar, Por lo que se plantea una metodología de carga de datos y demostrar su aplicabilidad.

1.1.2. Formulación del problema.

¿Qué metodología es la adecuada para mejorar la carga de datos en el Cuadro de Mando Integral de la Universidad Técnica Estatal de Quevedo?

1.1.3. Sistematización del problema.

- ¿Cómo integrar la información almacenada en las diferentes fuentes de datos de la institución?
- ¿Cómo mejorar el cumplimiento de la carga oportuna de datos para el Cuadro de Mando Integral?
- ¿De qué forma se puede implementar la carga de información que no se encuentra en las fuentes de datos de la UTEQ?

1.2. Objetivos.

1.2.1. Objetivos General.

- Determinar la metodología apropiada para la carga de datos al Cuadro de Mando Integral de la Universidad Técnica Estatal de Quevedo.

1.2.2. Objetivos Específicos.

- Definir los mecanismos para extraer la información de las diferentes fuentes de datos de la institución.
- Desarrollar los métodos necesarios para calendarizar y notificar la cargar adecuada de los datos.
- Implementar las interfaces correspondiente para la carga datos que no se encuentran en las fuentes de datos de la UTEQ.

1.3. Justificación.

Este proyecto de investigación nace ante la necesidad de crear un almacén de datos que integre la información estratégica que necesita la Unidad de Planeamiento Estratégica para el Cuadro de Mando Integral de la Universidad Técnica Estatal de Quevedo, con el fin de mejorar la situación actual de cómo se lleva ese proceso disminuyendo él envió de peticiones a la Unidades Académicas para obtener los datos que se necesitan, aportando a mejorar desempeño y eficiencia de quienes hacen el seguimiento del Plan Estratégico.

Con la propuesta, se automatizan los procesos en la carga de datos para garantizar la entrega oportuna de la misma, así como para evitar los retrasos y errores que se pueden cometer cuando se realiza de forma manual, haciendo que se vuelva más eficiente el seguimiento del Plan Estratégico. Ayuda a cumplir con los requerimientos del CEAACES [4] a las Instituciones de Educación Superior, para mejorar la evaluación del Plan Estratégico dando un mayor control en las decisiones que se deben tomar.

Con la información ingresada oportunamente el personal de la Unidad de Planeamiento Estratégico (UPE) podrá desempeñar mejor su labor, aumentando su eficiencia y efectividad para la gestión estratégica, dando un soporte apropiado a la toma de decisiones.

La metodología de Carga de Datos de este proyecto de investigación utiliza un proceso ETL (Extraer, Transformar y Cargar), pero como una variante debido a que se extrae la información desde las fuentes de datos y la integra en la base de datos del Sistema de Balance ScoreCard puede servir como base para nuevos proyecto de investigación que deseen aplicar procesos de extracción y carga de información desde diversas fuentes de datos.

CAPÍTULO II

FUNDAMENTACIÓN TEÓRICA.

2.1. Marco conceptual.

2.1.1. Balance ScoreCard.

2.1.1.1. El Balance ScoreCard

El Balanced Scorecard es una de las metodologías más usadas para implementar la Administración Estratégica en las empresas y organizaciones. Esta metodología fue publicada en el año 1992 por los académicos de la Universidad de Harvard Robert Kaplan y David Norton, a través del libro "The Balanced Scorecard: Translating Strategy into Action". El Balance Scorecard permite traducir los objetivos estratégicos en resultados que ayudaran a la empresa a tomar las mejores decisiones [6].

2.1.1.2. Característica del Balance ScoreCard

El alcance y característica de un Balance Scorecard se deberán adaptar a las características, necesidades y situación que tiene la institución. Por ejemplo una Institución o empresa pequeña con un Balance Scorecard con solo las perspectivas básicas pueden ser suficientes, caso contrario tienen las empresas con mayor tamaño que sus Balance Scorecard puede tener varios niveles, cada nivel representando una perspectiva diferente. Dependiendo del tipo de actividad a que se dedique las empresas se fijaran las estrategias en el orden de importancia para dicha institución [7].

2.1.1.3. Componentes esenciales del Balance ScoreCard

El Balanced Scorecard está compuesto por una serie de elementos conectados que permiten definir la estrategia de la compañía en todos sus aspectos. Estos componentes están ligados con la razón de ser de la empresa como son la Misión, Valores, Visión, Objetivos estratégicos, Mapa estratégico, Indicadores, Metas e Iniciativas. Todos estos elementos juntos a las perspectivas permitirán plantear las estrategias que guiaran a la Institución a cumplir con todos los objetivos planteados [8].

2.1.1.4. Perspectivas del Balance ScoreCard

Robert Kaplan y David Norton definieron cuatro perspectivas que se relacionaban directamente con los indicadores, estas son las perspectiva del clientes que es “¿Cómo nos ven los clientes?”; la perspectivas interna que es “¿En qué debemos ser excelentes?”; la perspectiva de innovación y aprendizaje que es “¿Podemos seguir mejorando y creando valor?” y la perspectiva financiera que es “¿Cómo nos ven los inversionista?”. Dependiendo de la actividad en que se desempeñe la empresa se establece el orden de las perspectivas. Por ejemplo una Institución educativa puede tener como su principal perspectiva la de clientes [3].

2.1.1.5. Ventajas de un Balance ScoreCard

Mejora la efectividad de la misión de la empresa, ayudando a tener mayor productividad. Alínea las actividades de la empresa con su planificación estratégica, teniendo una retroalimentación de información para cumplir con sus objetivos planteados

Determina la eficiencia en los procesos, lo cual ayuda a realizar mejoras que deben realizarse. Basándose en la recolección de datos pasados, permite proyectar los costó más exactamente para la toma de decisiones [9].

2.1.2. La información en el Balance ScoreCard.

2.1.2.1. Que información se necesita inicialmente.

Es importante no complicarse, sobre todo al principio, con indicadores y modelos complejos: indicadores selectivos, sencillos admitidos por todos los usuarios, etc. Son una buena fórmula en las primeras etapas del BSC. Bajo el concepto general "Soporte a la toma de decisiones" se esconden varias necesidades comunes: comprobar que todo va bien, analizar diferentes aspectos de la evolución de la empresa, visualizar información de forma más intuitiva, comparar información en diferentes periodos de tiempo, comparar resultados con previsiones, identificar comportamientos y evoluciones excepcionales, confirmar o

descubrir tendencias o interrelaciones, necesidad de realizar análisis predictivos, etc. Son todas ellas necesidades parciales dentro del concepto general [10].

2.1.2.2. Metadatos

Los metadatos, Según Howe [11], el término fue acuñado por Jack Myers en la década de los 60 para describir conjuntos de datos. La primera acepción que se le dio fue la de dato sobre el dato, ya que proporcionaban la información mínima necesaria para identificar un recurso. En este mismo trabajo se afirma que puede incluir información descriptiva sobre el contexto, calidad y condición o características del dato. Teniendo en cuenta que la mayoría de sistemas de metadatos ha sido creada no sólo por profesionales de la información sino también por informáticos, diseñadores de programas, técnicos de sistemas, etc. Incidiendo sobre el aspecto básico de la definición, otros autores amplían el concepto de "dato sobre el dato" al afirmar que incluyen información sobre su contexto, contenido y control así como todo lo que tenga que ver con "el dato" [12]

Ercegovac [13], por su parte, afirma que un metadato describe los atributos de un recurso, teniendo en cuenta que el recurso puede consistir en un objeto bibliográfico, registros e inventarios archivísticos, objetos geoespaciales, recursos visuales y de museos o implementaciones de software. Aunque puedan presentar diferentes niveles de especificidad o estructura, el objetivo principal es el mismo: describir, identificar y definir un recurso para recuperar, filtrar, informar sobre condiciones de uso, autentificación y evaluación, preservación e interoperabilidad.

- Se pueden describir funciones acerca de los metadatos.
- Resumir el significado de los datos
- Permitir la búsqueda
- Determinar si el dato es el que se necesita
- Prevenir ciertos usos
- Recuperar y usar una copia del dato
- Mostrar instrucciones de cómo interpretar un dato
- Obtener información sobre las condiciones de uso (derechos de autor)
- Aportar información acerca de la vida del dato

2.1.2.3. Metadatos en la integración de la información.

Las estructuras de metadatos están adquiriendo una posición importante en lo que se refiere a la descripción de recursos electrónicos entendidos como objetos (Bases de Datos, Archivos planos). Cada vez son más numerosos los proyectos, sitios Web o sistemas de consulta que se valen de ellos para lograr mejores prestaciones a la hora de la representación, localización y recuperación de recursos electrónicos como es la información. Al mismo tiempo es posible utilizar la mayoría de conjuntos de metadatos actuales junto a cualquier lenguaje, lo que les aporta la característica de multiplataforma que los convierte en la herramienta ideal para crear un entorno de información integrada en el que el catálogo proporcione acceso tanto a los documentos tradicionales como a la información electrónica [12].

2.1.2.4. Flujo de información en la implementación del cuadro de mando integral.

Se define cuáles son las fuentes de información necesarias para el cálculo de los indicadores del cuadro de mando integral, en caso de no existir se debe incorporar al sistema un mecanismo para ingresar esta información faltante, si esto resulta muy costoso se diseña otro indicador que refleje el mismo resultado, se define además quién es el responsable de cada uno de los indicadores. En el diseño del sistema para el procesamiento de la información, se tiene en cuenta el levantamiento de los diferentes sistemas informáticos existentes en la empresa para el procesamiento y análisis de la información. Búsqueda de sistemas que permitan el acceso a la información de manera integral y desde la misma fuente de manera tal que se facilite el proceso de toma de decisiones [14].

2.1.2.5. Almacenes de datos.

Los Almacenes de Datos son "una colección de datos orientados por temas, integrados, variables en el tiempo y no volátiles para el apoyo de la toma de decisiones" (Inmon, 2005). Los Almacenes de Datos son integradores, debido a que su contenido proviene de diversas fuentes de datos como: Archivos de Textos, Base de Datos Relacionales, ERP, entre otros medios. Estos procesos gestionan la extracción de los datos desde sus fuentes

de origen, de transformarlos a la información requerida, de lograr la depuración necesaria en aquellos que lo necesiten y finalmente cargar al Almacén de Datos deseado.

Los procesos ETL cargan datos periódicamente desde las fuentes a los almacenes de datos, procurando registrar la variación que tenga el dato a lo largo del tiempo, de lo cual se puede decir que los almacenes de datos solamente sean utilizados para la lectura de la información y no como una base de datos donde se modifiquen los datos. Desde esta perspectiva, los procesos ETL son trascendentales ya que cumplen con la importante labor de recoger los datos y posicionarlos en el lugar que será la fuente del conocimiento para quienes realicen análisis sobre ellos. Las personas que construyan el ETL, y la tecnología anexa para la visualización de los datos, serán los responsables de generar toda la información necesaria para su posterior análisis. [15].

2.1.2.6. Almacenes de datos y procesos ETL.

En la construcción de un almacén de datos, una de la actividades que más consume tiempo es la elaboración de procesos ETL que tienen que ser la parte más importante en la construcción de un almacén de datos, debido a que no se le da mayor importancia a este hecho no existe un modelo formal para la representación de las actividades del Proceso ETL que permitan que los datos de las fuentes de datos de entrada estén en un formato adecuado para la carga de datos [16].

2.1.2.7. Procesos ETL.

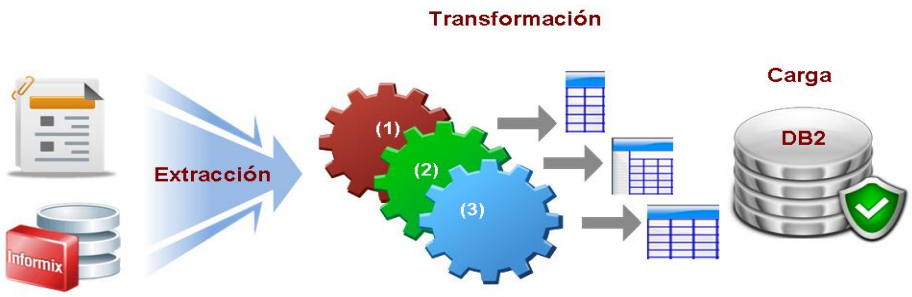
En un proceso ETL los datos son extraídos desde sistemas fuentes, los cuales pasan por una secuencia de transformaciones antes de que se carguen en una repositorio donde van a servir para ser usados posteriormente, estas fuentes de datos pueden variar desde bases de datos hasta hojas de cálculo, esta transformaciones son usualmente programadas (C Sharp, Java, etc.) y que se encuentran fuera de las bases de datos [17].

El diseño de un proceso de ETL se compone generalmente de seis tareas definidas en: [18].

a) Seleccionar los datos para la extracción.
b) Definir los datos de las fuentes

c) Transformar las fuentes
d) Unir las fuentes
e) Seleccionar el destino de carga
f) Unir los atributos de las fuentes de datos con los atributos del destino
g) Carga los datos en el repositorio

Figura 1 Procesos ETL



Fuente: https://www.ibm.com/developerworks/ssa/data/library/tipos_bases_de_datos/

2.1.2.8. Carga Inicial.

La extracción de la información desde las fuentes de origen hasta su carga en las tablas destino es el proceso más importante en este proyecto. Este tipo de actividades separan en el estudio de un conjunto de operaciones sobre dichos datos. Las actividades se clasifican según su estudio en los siguientes tipos: [19].

a) **Obtención de datos:** Extraer datos a partir de tablas origen.
b) **Chequeo de existencia:** Verificar la existencia de un elemento previamente cargado en la fuente de datos.
c) **Transformación:** Ejecutar cualquier función que sea aplicada sobre los datos para transformarlos, por ejemplo:
    - **Cambio de Formatos:** Realizar una conversión en la presentación de los datos.
    - **Generación** de códigos: Generar identificadores, aplicando funciones sobre otros datos existentes.
    - **Cálculos:** Calcular un nuevo dato a partir de otros datos base.
d) **Limpieza:** Efectuar cualquier actividad que procure identificar y corregir errores en los datos.

e) **Inserción:** Cargar los datos en las tablas destino del Data Warehouse
f) **Por simplicidad**, en adelante se menciona a las actividades como sinónimo de sus tipos de actividad.

2.1.3. Fuentes de información.

Al implementar un sistema de software que automatice los procesos manuales encargados de ingresar y gestionar la información correctamente, La información, que sirve para entender los procesos manuales de manera organizada y para evaluar los problemas que existen, en las empresa se necesitan herramienta que permitiera automatizar y disminuir las incidencias de errores detectadas en el proceso manual, de cómo ingresar incorrectamente datos del alumno, del curso, profesor, etc. No obstante, se pueden generar fallas al convencer a los usuarios en ver la conveniencia para agilizar y facilitar el proceso de carga de datos.
Los diferentes proyectos que acceden a la base de datos tienen información más confiable pues, al automatizar los procesos de carga masiva de datos, se reduce la incidencia de errores [20].

2.1.4. Tipos de fuentes de información.

Es importante caracterizar adecuadamente las diferentes fuentes de información. En una clasificación básica, atendiendo fundamentalmente a la flexibilidad del modelo de datos, entre las que tenemos a las siguientes: [21].

Fuentes no estructuradas. Son aquellas que no representan ningún tipo de esquema en la información que contienen esto pueden ser documentos de texto, Word, PDF, Páginas web, etc. Debido a que los datos no están estructurados se complican enormemente su tratamiento automatizado.

Fuentes Semiestructuradas. No representan un esquema rígido. Esto quiere decir que el esquema es implícito y está contenido en los propios datos. Esto quiere decir que estando en un esquema éste es muy vago y que permite una cierta flexibilidad. Se pueden mencionar las siguientes fuentes XML, Excel, PDF como fuentes semiestructurada

Fuentes Estructuradas. Representan un esquema rígido bien definido, diferenciado de los datos. Un ejemplo de una fuente de datos estructurados es una base de datos relacional porque representa un esquema, almacenado en un diccionario de datos que define una la organización interna de la información. El acceso automatizado a este tipo de fuentes es mediante potentes lenguajes de consultas.

2.1.5. Estrategias para la integración de la información heterogénea.

Es importante disponer de tecnologías que gestionan la complejidad en el tratamiento e integración de la información heterogénea. El desafío consiste en ser capaces de acceder y aprovechar en nuestras aplicaciones, de manera automatizada, la enorme cantidad que se tienes en los diferentes tipos de información [21].

Las diferentes soluciones que se plantean son las siguientes:

a) Homogeneización de las fuentes de datos semiestructuradas.
b) Mediante tecnologías de Data Warehouse
c) Desarrollo de un mediador que permita un tratamiento automatizado que disminuya la necesidad de modificación y colaboración de las fuentes.

Dependiendo del entorno donde se vaya a aplicar es muy importante conocer y discriminar su aplicabilidad en cada situación real.

2.1.6. El entorno de archivo tradicional.

Las instituciones crecen de manera natural, siempre que sea administrada correctamente cumpliendo con los objetivos y metas propuestas. Cuando en sus operaciones se incluyen sistemas informáticos se vuelven más dependientes de las facilidades que esto prestan. A medida que la empresa crece y madura, la organización tiene que cambiar y plantear nuevos procedimientos y estructuras para la planificación. Las necesidades de información también aumentan, pero la diferencia de la organización es que las soluciones no siempre están organizadas, creciendo de manera independiente de la planificación, donde cada departamento logra funcionar con sus propios sistemas y bases de datos [22].

2.2. Marco Referencial.

2.2.1. Herramientas para el Cuadro de mando integral.

Carlos Azabache [10], trabajo en la tesis “Sistema de información gerencial para la gestión estratégica en ventas por medio de un Cuadro de mando integral”, que muestra la información estratégica de la empresa donde se implementó. Planteó un mecanismos para no depender del envió de información de todos los departamentos de la empresa. Su Cuadro de Mando dispondrá de su propia base de datos lo cual permitirá accesos más oportunos a la información requerida ya que en él se tendrá solo información clave para el proceso de la toma de decisiones.

El Cuadro de Mando Integral: La respuesta a la implantación exitosa de la estrategia en las universidades públicas españolas [23], es una publicación sobre el empleo de un cuadro de mando integral en la instituciones de educación superior basadas en un mapa estratégico con una propuesta de objetivos en cada una de la perspectivas así como las relaciones causa causa-efecto entre ellos.

Una Aproximación Multi-Agente para el Soporte al Proceso de Extracción-Transformación-Carga en Bodegas de Datos es una publicación [24], que propone una solución de automatización de carga de datos utilizando procesos ETL en Bodegas de datos. Considera la heterogeneidad y disponibilidad de las fuentes de datos con la cual se pretende alcanzar un mayor nivel de automatización en las tareas involucradas.

Desarrollo de un sistema de información computarizado basado en datos maestros, utilizando una metodología de indicadores balanceados de desempeño (Balanced scorecard), para apoyar la estrategia corporativa de la universidad de Manizales es una tesis desarrollada por [25], donde se identifica el estado en el que se encuentran las fuentes de información de la universidad y su posible aporte al logro de los objetivos del proyecto; adicionalmente se identificaron los avances de la Universidad en la implementación de los conceptos de direccionamiento estratégico y el conocimiento que se tenía sobre el manejo de indicadores de gestión.

CAPÍTULO III

MÉTODOLOGIA DE LA INVESTIGACIÓN.

3.1. Localización.

Este proyecto se ha desarrollado para ser implementado en la Universidad Técnica Estatal de Quevedo específicamente en la Unidad de Planificación Estratégica (UPE) quienes van a estar encargados de toda la administración del sistema de carga de datos IAD-BSC.

3.2. Tipo de investigación.

El proyecto de investigación **"Desarrollo de una metodología de alimentación de datos para el Cuadro de Mando Integral de la UTEQ",** se desarrolló en la Universidad Técnica Estatal de Quevedo (UTEQ) siendo la primera propuesta que plantea una Metodología de Carga de Datos para integrar las diferentes fuentes de datos con las que trabajan en la Universidad. La revisión bibliográfica y referencial permitió obtener ciertos criterios sobre la construcción de la propuesta de solución. Sin embargo, fue necesario realizar ajustes para adoptar a la realidad de las plataformas y sistemas informáticos de la UTEQ, luego de aplicar una investigación de diagnóstico situacional, fue necesario plantear una metodología y estructurar modelos conceptuales y de datos para alcanzar los objetivos planteados en este proyecto.

Una vez desarrollado los modelos se comprobó la utilidad en un software de Balance ScoreCard en construcción, utilizando en ocasiones un proceso deductivo para aplicar las reglas generales de conectividad a sistemas específicos y en otros procesos inductivos para generalizar los procesos, librerías y evitar en lo posible utilizar la funcionalidad de los DBMS con los que trabajan los sistema de las fuente de datos.

3.3. Método de investigación

Este proyecto de investigación utilizó el método de investigación analítico debido a que realizó análisis de algunos elementos que aportaron a desarrollo del trabajo para lograr darle una solución al problema planteado en este proyecto. Los elementos que analizaron fueron los siguientes:

- Los diferentes sistemas (Fuentes de datos) con los que trabaja la universidad.

- Departamentos de la universidad que trabajan con información que la tienen almacenadas en hojas de cálculos.

- La forma de integrar este proceso de carga de información y que pueda ser utilizado con el Sistema de Balance ScoreCard.

- La forma de controlar que la información que necesita los indicadores esté actualizada siempre siguiente un proceso.

Para el desarrollo de la Investigación se requieren los siguientes pasos.

1. Recopilación de información
2. Requerimientos del BSC para carga datos
3. Análisis de sistemas BSC que utilicen la carga de datos.
4. Plataformas de trabajo en la UTEQ
5. Determinar los DBMS con los que puede trabajar la carga de datos
6. Desarrollo de la metodología de carga de datos
7. Implementación de la metodología en un Software
8. Demostrar la aplicabilidad de la metodología en un sistema BSC

La metodología de desarrollo Extreme Programming fue la empleada para el desarrollo de software del proyecto de investigación **"Desarrollo de una metodología de alimentación de datos para el Cuadro de Mando Integral de la UTEQ",** esta es una metodología de desarrollo ágil [26], que permite darle al proyecto un aspecto más técnico en desarrollo de la programación a través de la historias de usuarios. Apoya una iteración más conjunta entre el desarrollador y las personas interesadas en el sistema lo que permite realizar correcciones durante el desarrollo del proyecto.
Esta metodología aplica Historias de Usuarios (ver formato de una historia de usuario en el Anexo 3) en lugar de los casos de usos que se utilizan en otra metodología. Estas historias se trabajan conjuntamente el cliente y el desarrollador donde se plantean la horas de trabajo [27].

Se dividió la programación del software en entregas que posteriormente se convirtieron en interacciones las cuales nos permitió agruparon las historias de usuarios que tenían relación entre sí.

3.4. Recursos Humanos y Materiales.

3.4.1. Recursos Humano.

El presente proyecto fue desarrollado por Diego García Coello, quien fue el encargado de la recopilación de toda la información, documentación, análisis, diseño, implementación y las pruebas necesarias de funcionamiento.

La persona de la Unidad de Planificación Académica que ayudo a definir los requerimientos para el sistema fue el Ing. Washington Chiriboga.

El Ing. Washington Chiriboga es el director de tesis del proyecto de investigación **"Desarrollo de una metodología de alimentación de datos para el Cuadro de Mando Integral de la UTEQ".**

3.4.2. Materiales.

Los materiales que se utilizaron para el desarrollo del proyecto de investigación se detallan a continuación:

3.4.2.1. Hardware

Tabla 1. Hardware

Ítem	Características	Precio
Laptop	Procesador Inter Core i3 de segunda generación Disco Duro de 320 GB Memoria RAM de 3GB	$ 800.00
Impresora	Canon Pixma 20	$ 80.00
Total		$ 880.00

3.4.2.2. Software

Tabla 2. Software

Ítem	Características	Precio
Sistema Operativo	Windows 8 Profesional x84	$ 196.00
Motor de Base de datos	Microsoft SQL Server 2012 MySQL Postgres	$ 80.00 $ 0.00 $ 0.00
Archivos para Fuentes de Datos	Microsoft Office Access Microsoft Office Excel	$ 150.00 $ 150.00
Lenguaje de programación	Visual Studio 2012	$ 550.00
Digitación	Microsoft Office Word	$ 264.00
Diagramas de ejecución del proyecto	Microsoft Project 2010	$ 493.00
Total		$ 1883.00

3.4.2.3. Suministros

Tabla 3. Suministros

Cantidad	Ítem	Costo Unitario	Costo Total
3	Resma de papel A4	$ 6.00	$ 18.00
1	Lapiceros y lápiz	$ 2.50	$ 80.00
3	Tinta de impresora	$ 28.00	$ 78.00
Total			$ 176.00

3.4.2.4. Personal

Tabla 4. Personal

Cantidad	Ítem	Costo Unitario	Costo Total
3 meses	Sueldo de Desarrollador	$ 1000.00	$ 3000.00
Total			$ 3000.00

3.4.2.5. Costo Total del Proyecto

Tabla 5. Costo del Proyecto

Nombre del Recurso	Valor
Hardware	$ 880.00
Software	$ 1883.00
Suministros	$ 176.00
Personal	$ 3000.00
Total	$ 5939.00

3.5. Método de Software

El modelo XP (Extreme Programming) que es una metodología ágil se tomó en cuenta para desarrollar la aplicación, este modelo divide el ciclo de vida de un proyecto en fases (Bibliografía) que son muy dinámicas. El sistema gestor de accesos de datos permitirá integrar los diferentes sistemas que trabajan en la UTEQ, con el objetivo de tener la información necesaria para el Balance ScoreCard.

Las fases que se van a emplear para desarrollar el software del proyecto de investigación son las siguientes:

3.5.1. Fases de desarrollo.

Las fases de desarrollo que se van a intervenir en el desarrollo del proyecto son las siguientes:

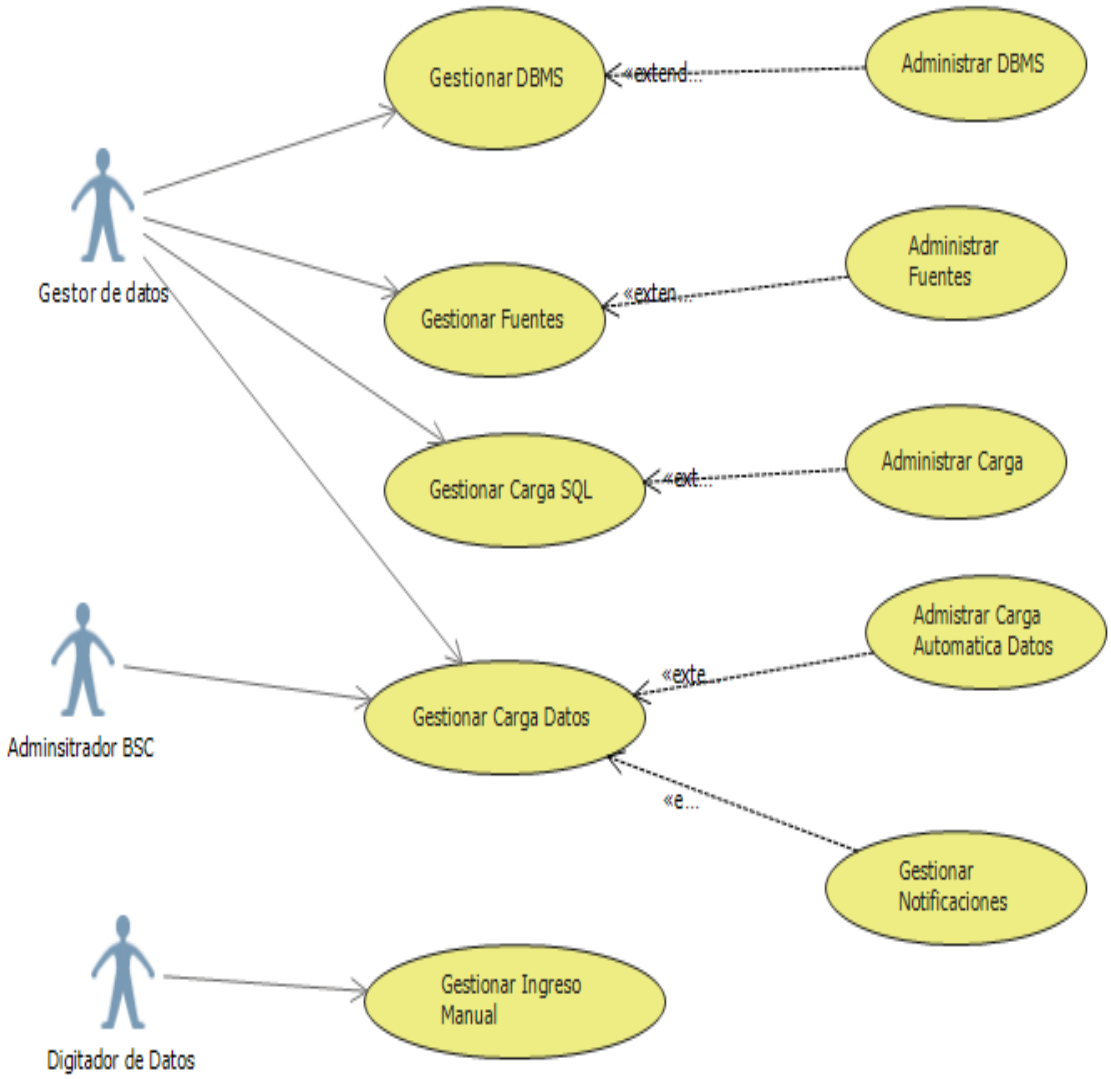
3.5.2. Fase de exploración.

En esta fase de definen los alcances y requerimientos de la personas interesadas en el proyecto. Se plantean los casos de uso y las historia de usuarios que van a dar un esquema

del funcionamiento del sistema. Las historias de usuarios se trabajaron directamente con las personas encargadas en la Unidad de Planificación Académica (UPA).

A continuación se muestra el diagrama de caso de uso que demuestra la estructura principal del proyecto de investigación.

Figura 2. Diagrama de Caso de Uso



Se plantearon las siguientes historias de usuarios en la fase de exploración, donde se presentan un bosquejo del diseño del sistema.

Tabla 6. Historia de usuario inicio de sesión.

Título:	**Historia de usuario:** Inicio sesión	Prioridad:	Baja
		Estimación:	2 horas
Descripción:	Como usuario del sistema quiero acceder al software para configurar los accesos y carga de datos de la variables del sistema BSC		
Criterio de aceptación:	Con el ingreso del usuario y contraseña correcto, se accede al sistema La contraseña no debe ser legible Un usuario o contraseña incorrecto emite el mensaje de error		
Prototipo:			

Tabla 7. Épica Gestionar versiones DBMS.

Título:	**Épica:** Gestionar Versiones DBMS	Prioridad:	Media
		Estimación:	10 horas
Descripción:	Como gestor de conexiones Quiero gestionar (agregar, editar y eliminar) las versiones de los DBMS. Para definir los DBMS a los que pueda acceder el sistema BSC		
Criterio de aceptación:	Presenta una lista de los DBMS y sus versiones Para eliminar una Versión DBMS de la lista doy un click en botón de la misma fila. Para cambiar los datos de una Versión DBMS ya ingresada doy click en el botón de la misma fila. Se debe dar click en el botón Nuevo para crear una nueva versión DBMS		
Prototipo:			

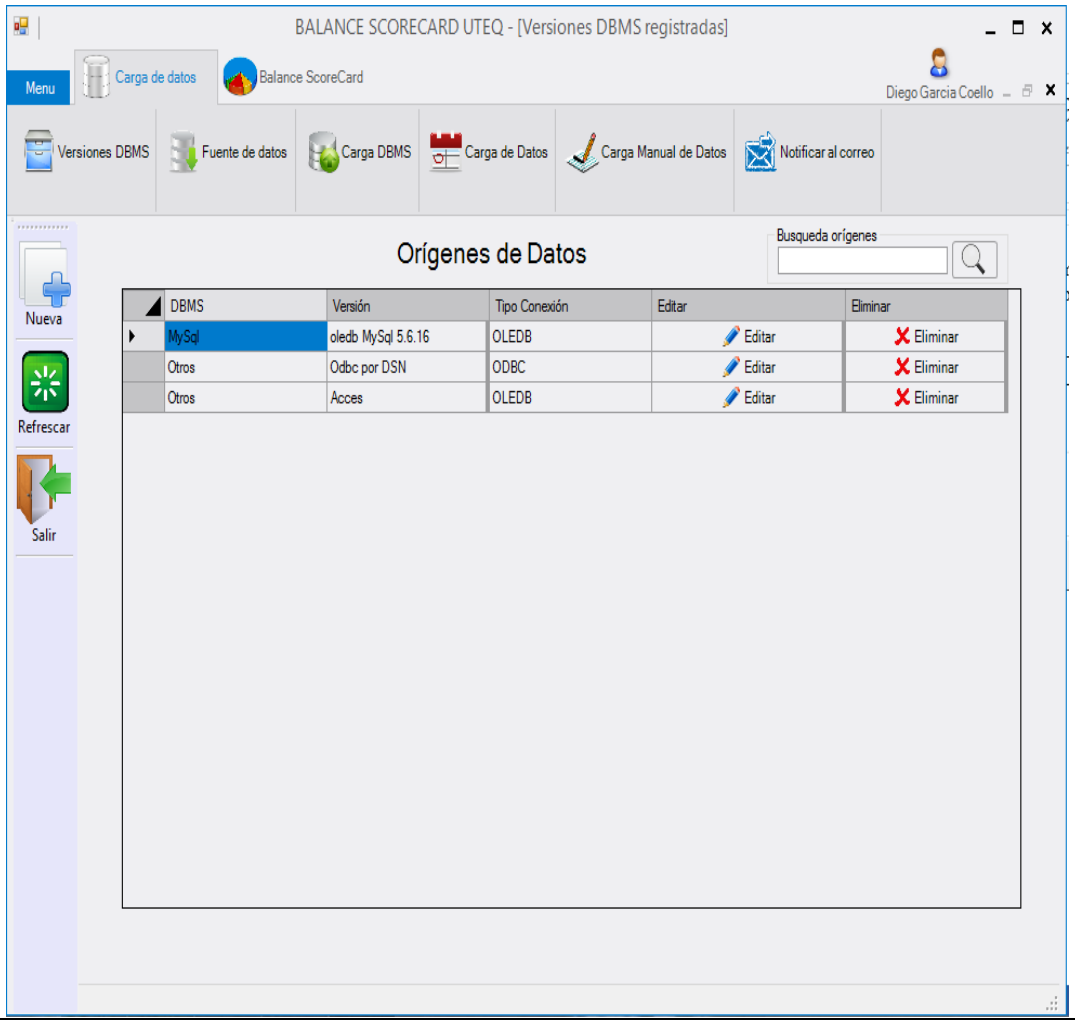


Tabla 8. Historia de usuario Administración de versiones DBMS

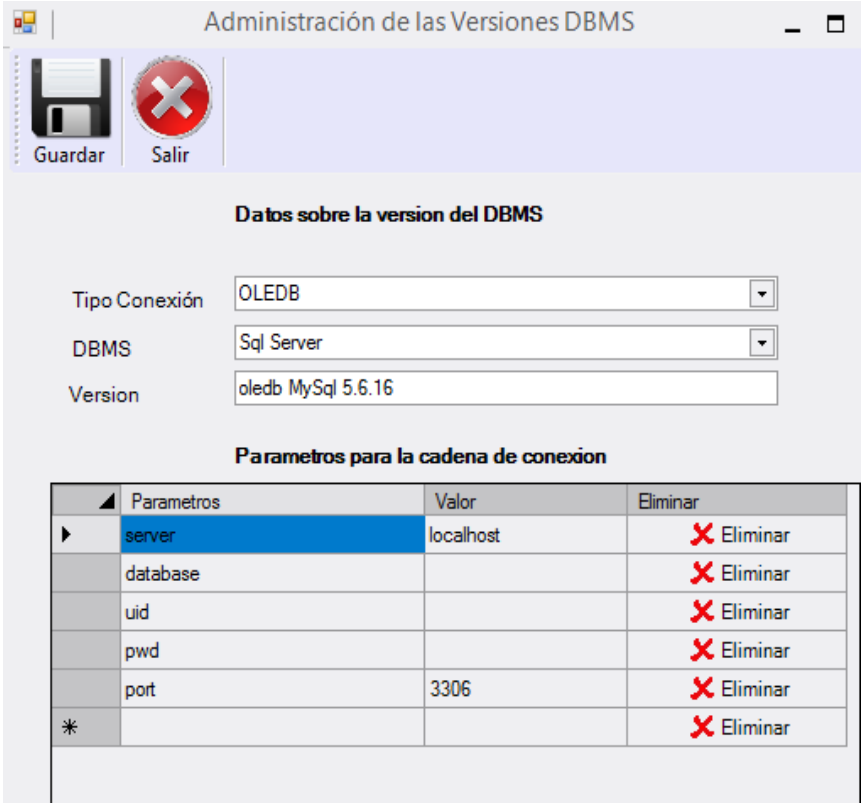
Título:	**Historia de usuario:** Administración de versiones DBMS	Prioridad:	Media
		Estimación:	15 horas
Descripción:	Como gestor de conexiones quiero administrar (agregar o editar) las versiones de los DBMS para definir los parámetros de conexión que debe llevar cada versión		
Criterio de aceptación:	Se debe elegir el DBMS desde la lista cargada. Se debe elegir el tipo de conexión que se va a necesitar. Se debe escribir la versión, que es un dato obligatorio del registro de versión. Presenta una lista de parámetros si se está modificando la versión. Se debe agregar nuevos parámetros en el Grid de parámetros. Se debe dar click en botón Eliminar de la misma para borrar el parámetro. Se debe dar click en el botón Guardar para registrar o modificar la versión. Si se da click sobre el botón Salir se cancela el proceso que se está realizando.		
Prototipo:			

Tabla 9. Épica Gestionar Fuentes de datos.

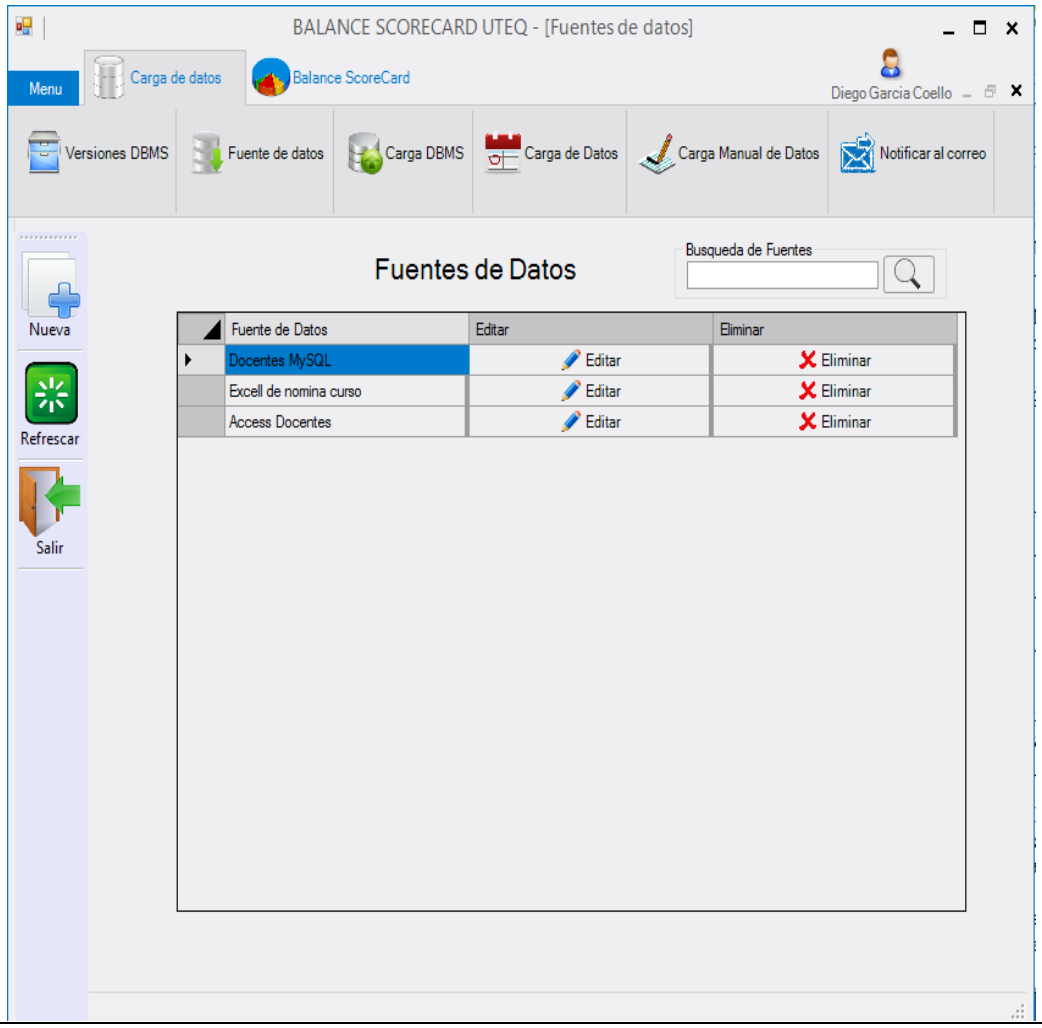
Título:	**Épica:** Gestionar Fuentes de datos	Prioridad:	Media
		Estimación:	10 horas
Descripción:	Como gestor de conexiones Quiero gestionar (agregar. Editar y eliminar) las fuentes de datos. Para definir las fuentes de datos que van a estar disponibles para la carga de datos.		
Criterio de aceptación:	Presenta una lista de las fuentes de datos Para eliminar una fuente de datos de la lista doy un click en botón de la misma fila. Para cambiar los datos de una fuente ya ingresada doy click en el botón de la misma fila. Se debe dar click en el botón Nuevo para crear una fuente de datos.		
Prototipo:			

Tabla 10. Historia de usuario Administrar las fuentes de datos.

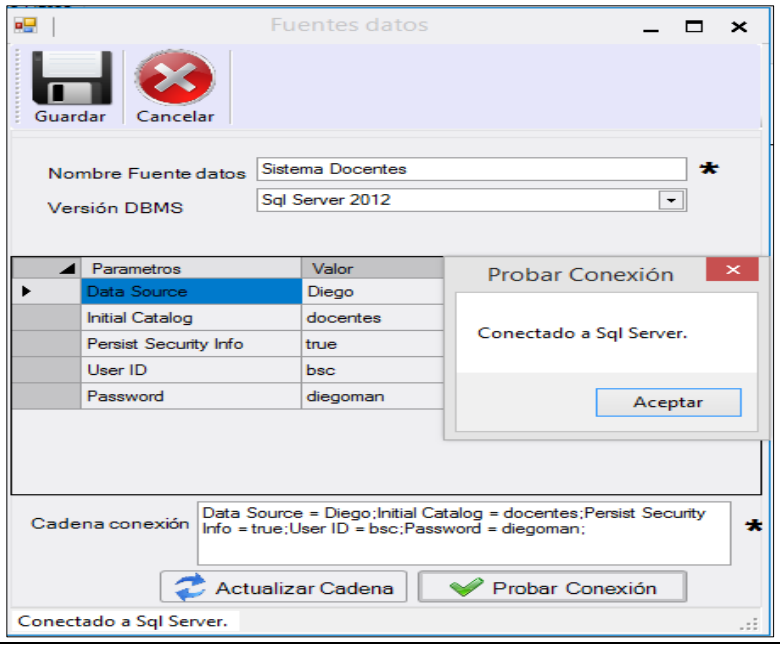
Título:	**Historia de usuario:** Administrar las fuentes de datos	Prioridad:	Media
		Estimación:	20 horas
Descripción:	Como gestor de conexiones quiero administrar (agregar o editar) las fuente de datos para validar los parámetros de conexión que debe llevar cada fuente		
Criterio de aceptación:	Se de elegir la versión DBMS de la lista del combo. Presenta una lista de parámetros según la versión DBMS. Si se está editando la fuente se cargaran los parámetros almacenados de la fuente, pero al cambiar la versión del DBMS se cargaran los parámetros por defecto de la versión seleccionada. Se debe completar la información necesaria para la cadena de conexión. La cadena de conexión de va actualizando automáticamente cada vez que llena información en el Grid parámetros. Se debe dar click para actualizar la cadena de conexión. Se debe dar click en botón Probar Conexión para comprobar si están correctos los parámetros de la cadena de conexión. Se debe llenar la información básica para guardar la fuente. Se debe dar click en el botón Guardar para registrar o modificar la versión. Si se da click sobre el botón Salir se cancela el proceso que se está realizando.		
Prototipo:			

Tabla 11. Épica Gestionar Cargas DBMS.

Título:	**Épica:** Gestionar Cargas DBMS	Prioridad:	Alta
		Estimación:	10 horas
Descripción:	Como gestor de conexiones Quiero gestionar (agregar, editar y eliminar) las cargas DBMS Para definir las cargas que se van a mostrar al momento de crear variables		
Criterio de aceptación:	Presenta una lista de las cargas registradas. Para eliminar una carga de la lista doy un click en botón de la misma fila. Para cambiar los datos de una carga ya ingresada doy click en el botón de la misma fila. Se debe dar click en el botón para crear una nueva carga.		

Prototipo:

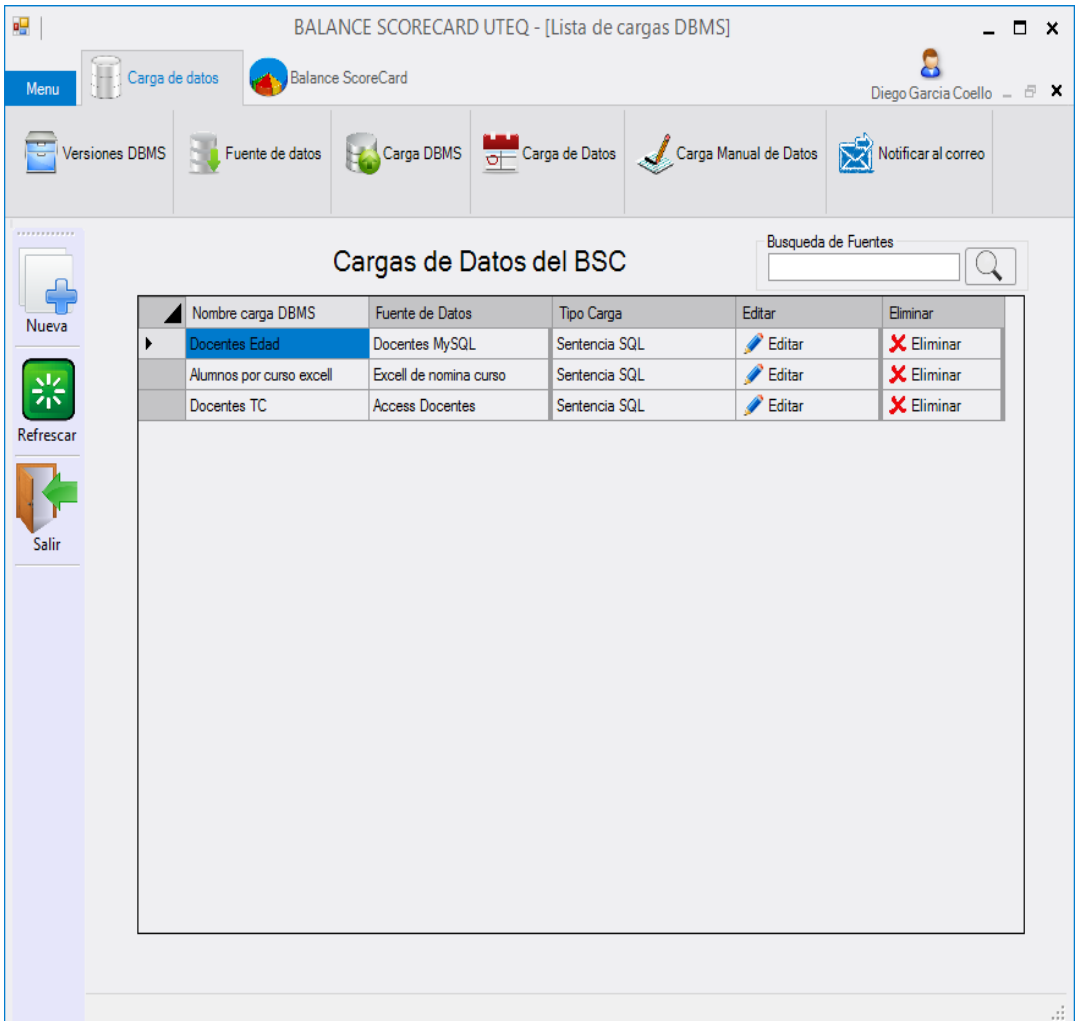


Tabla 12. Historia de usuario Administración de cargas DBMS.

Título:	**Historia de usuario:** Administración de cargas DBMS.	Prioridad:	Alta
		Estimación:	35 horas
Descripción:	Como gestor de conexiones quiero administrar (agregar o editar) las cargas Para definir los datos que se van a utilizar al momento de realizar las consultas a las fuentes de datos.		
Criterio de aceptación:	Se debe escribir el nombre de la carga que es un registro obligatorio. Se debe elegir la fuente de datos de una lista. Se debe elegir el tipo de carga que se va a utilizar para extraer la información. Se debe escribir la instrucción SQL base que es obligatoria. Si la consulta incluye parámetros se debe elegir desde el Grid, dando doble click sobre el elemento. Los parámetros se agregan en la lista, definiendo el tipo de dato. Se elimina un parámetro dando click en el botón eliminar sobre el elemento seleccionado. Se debe dar click en botón Ejecutar Carga de Prueba para comprobar si están correcta la instrucción SQL y los parámetros con su valor. Se mostrará un Grid con los datos de la carga de prueba (Gráfico 2). Se debe completar la información obligatoria para guardar. Se debe dar click en el botón Guardar para registrar o modificar. Para guardar, la instrucción SQL solo debe tener como respuesta un solo registro. Si se da click sobre el botón Salir se cancela el proceso que se está realizando.		

Prototipo:

Gráfico 1

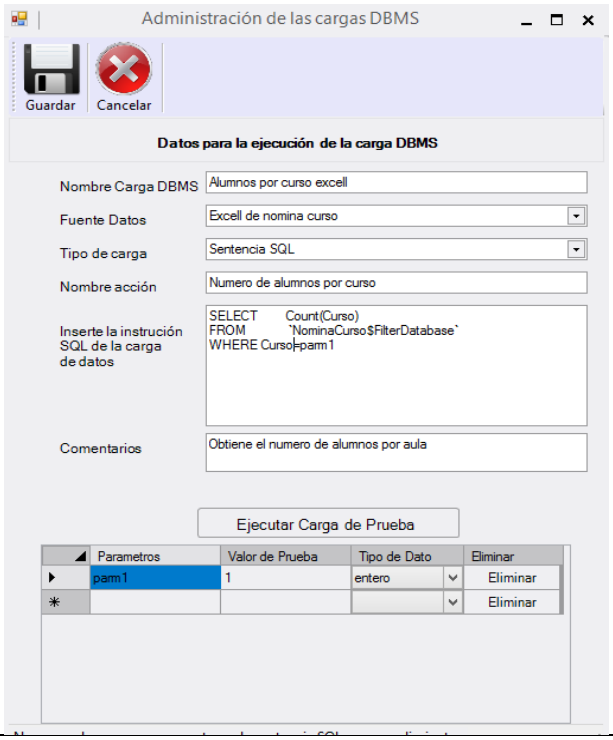


Gráfico 2

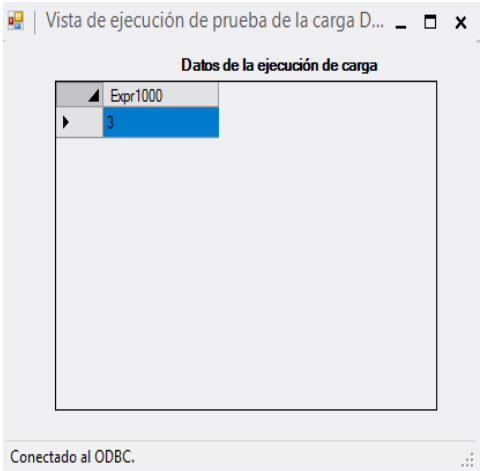


Tabla 13. Historia de Usuario Gestionar cargas de datos.

Título:	**Épica:** Gestionar Carga de datos	Prioridad:	Alta
		Estimación:	40 horas
Descripción:	Como gestor de conexiones Quiero gestionar cargas de datos. Para definir los datos que se van a almacenar en la historia variable.		
Criterio de aceptación:	Presenta una lista de carga de datos pendientes (Gráfico 1). Se debe especificar las fechas en la cuales se quiere realiza la consulta de la cargas de datos. La carga de datos se ejecutan sobre un elemento seleccionado de la lista Se debe dar click en el botón ejecutar todas las cargas para vaciar el Grid. Si esta seleccionado el Check se mostrara el historial de la variable que acaba de ser procesada su carga (Gráfico 2). Para consultar el registro de las cargas realizadas se debe seleccionar la pestaña (Gráfico 3). Se debe elegir las fechas, seleccionar una variable de la lista y seleccionar un estado de la lista para realizar la consulta que llenara el Grid.		

Prototipo:

Gráfico 1

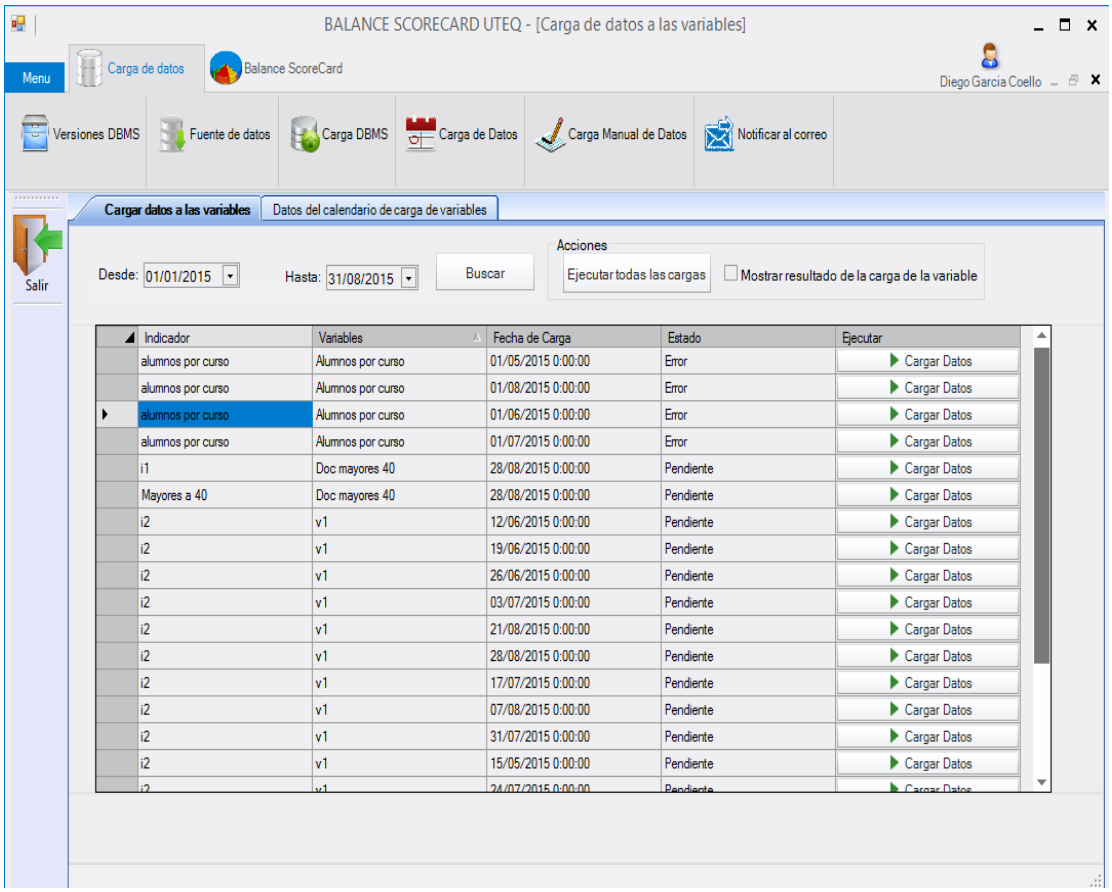


Gráfico 2

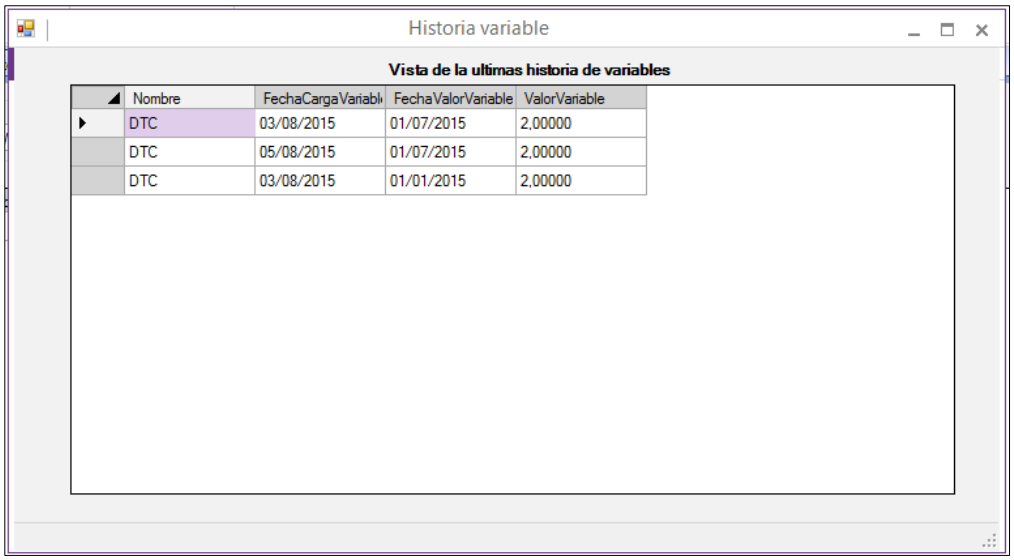


Gráfico 3

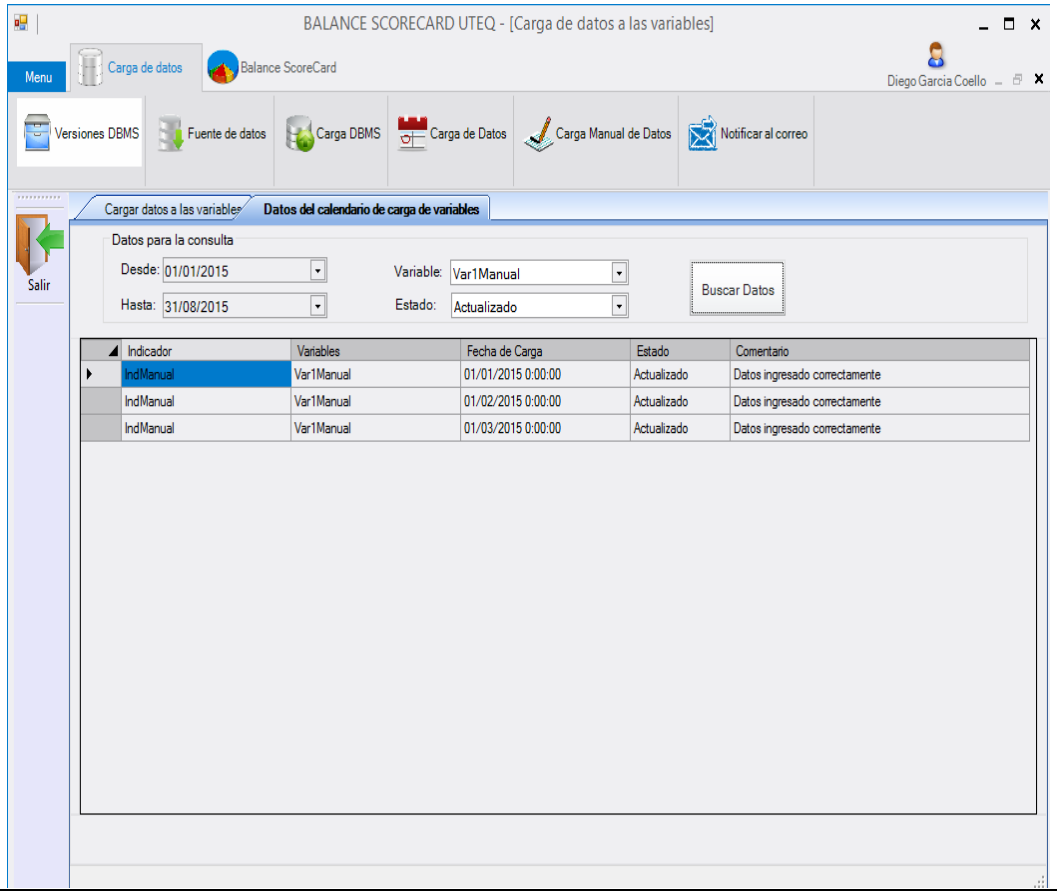


Tabla 14. Historia de usuario Carga manual de datos.

Título:	**Historia de usuario:** Carga manual de datos.	Prioridad:	Alta
		Estimación:	10 horas
Descripción:	Como gestor de conexiones quiero registrar los valores Para definir los datos que se van a utilizar en la historia de la variable.		
Criterio de aceptación:	Se debe escoger el rango de fechas en la que se va a buscar el calendario de la variable. Se debe elegir la variable de la lista. Se debe dar click en el botón Buscar para mostrar las cargas (pendientes o error), para asignarles el valor correspondiente. Se debe dar click en el botón Guardar Valor de la misma fila para registrar el valor Se debe escribir el valor en el formato correcto para registrarlo Si tienes seleccionado el Check, se mostrará un formulario (Gráfico 2) con el historial de la historia de la variable que acaba de ser registrada.		

Prototipo:

Gráfico 1

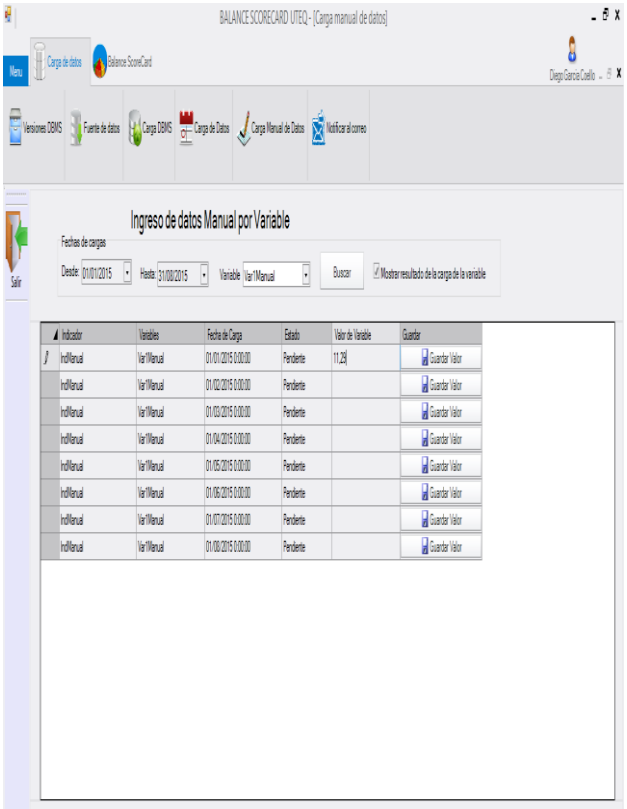
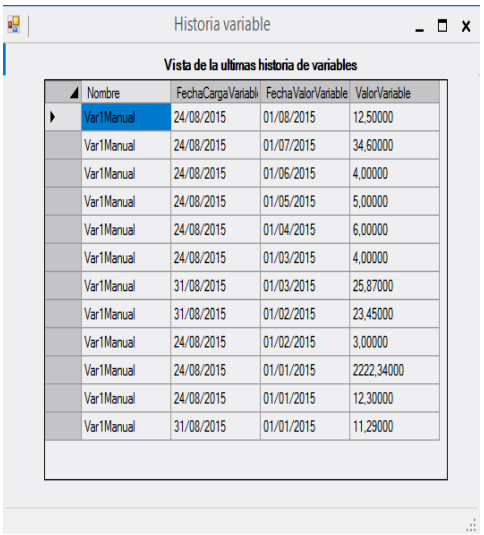


Gráfico 2



Nombre	FechaCargaVariable	FechaValorVariable	ValorVariable
Var1Manual	24/08/2015	01/08/2015	12,50000
Var1Manual	24/08/2015	01/07/2015	34,60000
Var1Manual	24/08/2015	01/06/2015	4,00000
Var1Manual	24/08/2015	01/05/2015	5,00000
Var1Manual	24/08/2015	01/04/2015	6,00000
Var1Manual	24/08/2015	01/03/2015	4,00000
Var1Manual	31/08/2015	01/03/2015	25,87000
Var1Manual	31/08/2015	01/02/2015	23,45000
Var1Manual	24/08/2015	01/02/2015	3,00000
Var1Manual	24/08/2015	01/01/2015	2222,34000
Var1Manual	24/08/2015	01/01/2015	12,30000
Var1Manual	31/08/2015	01/01/2015	11,29000

Tabla 15. Historia de usuario Notificaciones al usuario.

Título:	**Historia de usuario:** Notificaciones al usuario.	Prioridad:	Baja
		Estimación:	10 horas
Descripción:	Como gestor de conexiones quiero gestionar las notificaciones Para informar a los digitadores de información los ingresos de información pendientes		
Criterio de aceptación:	Se debe escoger el rango de fechas en la que se va a buscar el calendario de la variable a notificar. Se debe elegir una variable de la lista. Se debe dar click en botón "Buscar" para presentar en el Grid todas las cargas pendientes y con errores. Se debe dar click en botón "Enviar Notificación", para enviar un correo electrónico (Gráfico 2) al usuario encargado de realizar dicha acción.		

Prototipo:

Gráfico 1

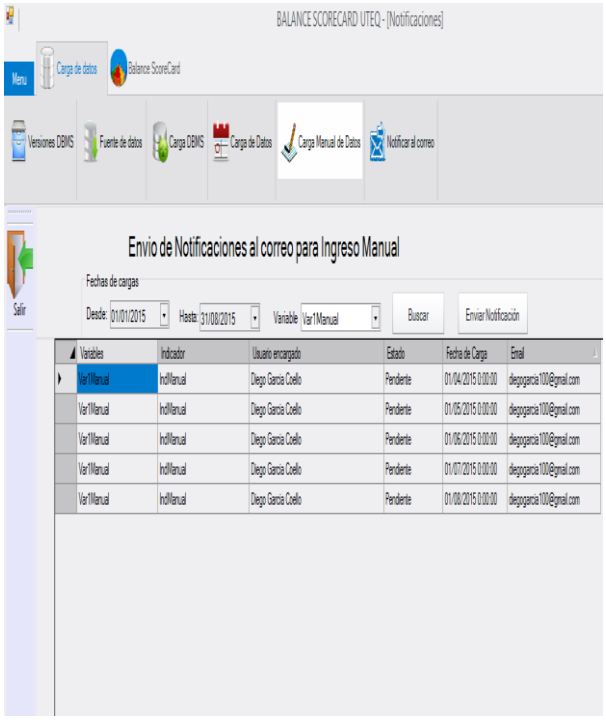
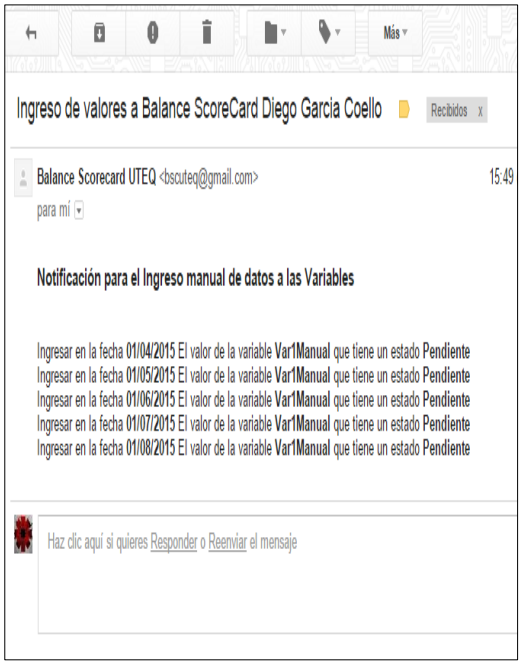


Gráfico 2



3.5.3. Fase de Planificación de la entrega.

En las reuniones realizadas se tomaron en cuenta las prioridades en que se van a desarrollar el **IAN-BSC**, definiendo en siguiente orden:

Primera entrega.- en esta etapa se tienen en cuenta las historias de usuarios que tienen una prioridad media. Entre las cuales tenemos las siguientes:

- Inicio de Sesión.
- Gestionar versiones DBMS
- Administración de versiones DBMS

Segunda entrega. - en esta etapa se tienen en cuenta un grupo de historias de usuarios que tienen una prioridad media. Entre las cuales tenemos las siguientes:

- Gestionar Fuentes de datos
- Administración de Fuente de datos

Tercera Entrega. - en esta etapa se tienen en cuenta las historias de usuarios que tienen una prioridad alta para el desarrollo del proyecto. Entre las cuales tenemos las siguientes:

- Gestionar Cargas DBMS
- Administrar las Cargas DBMS

Cuarta entrega. - en esta última etapa se tienen en cuenta las historias de usuarios que tienen una prioridad alta y prioridad baja para el proyecto. Entre las cuales tenemos las siguientes:

- Gestionar cargas de datos que tiene una prioridad alta.
- Carga manual de datos que tiene una prioridad alta
- Notificaciones al usuario que tienes una prioridad baja.

3.5.4. Fase de Iteraciones.

3.5.4.1. Primera Iteración: Elementos de desarrollo.

En la primera iteración con el usuario interesado en el sistema se determinó la tecnología con la que se va a desarrollar el proyecto, en la cual se definieron los siguientes elementos:

Tabla 16. Elementos de la plataforma de desarrollo

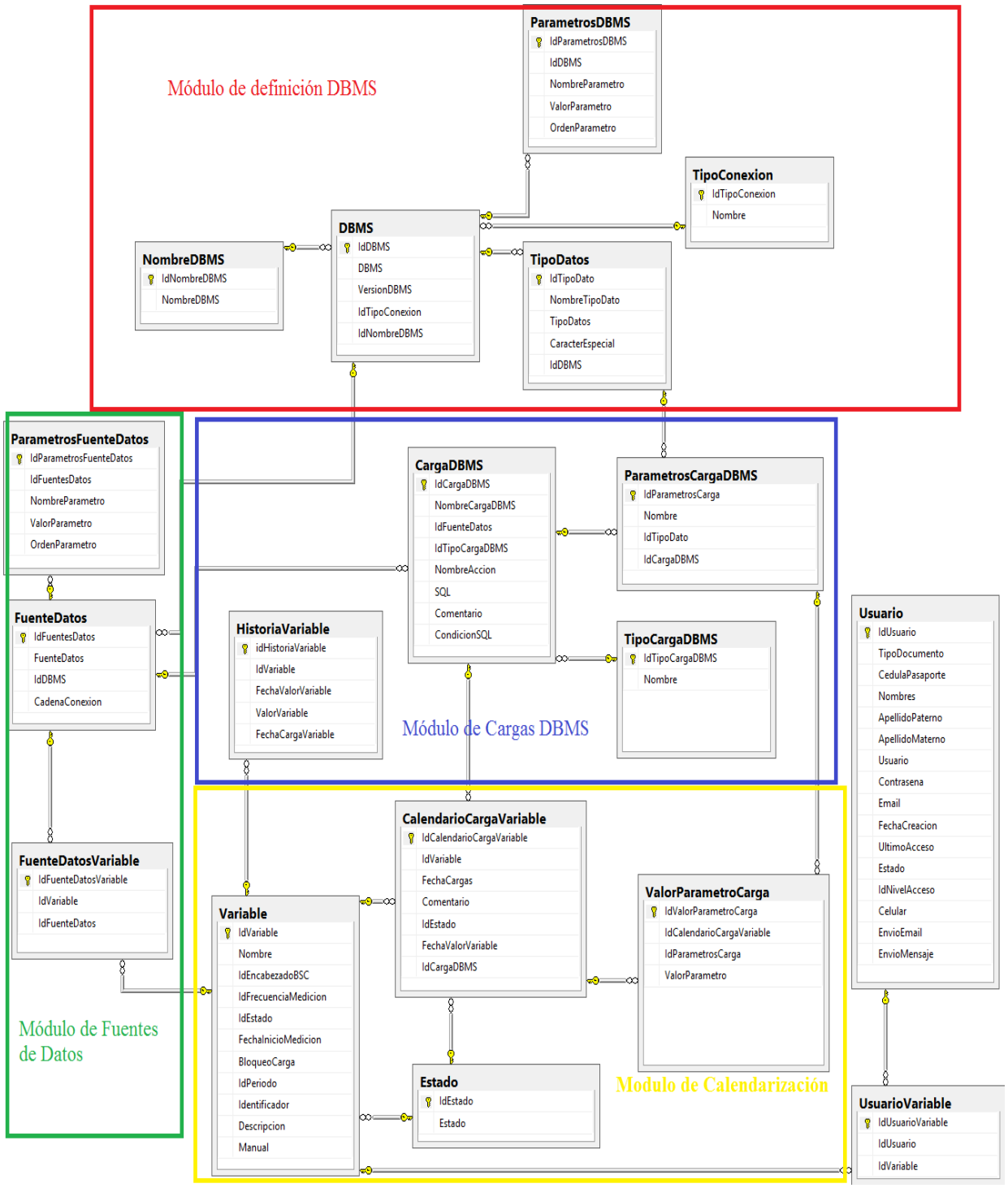
Gestor de Base de Datos:	SQL Server 2012
FrameWork:	Visual Studio 2012
Lenguaje:	C Sharp SQL
Herramientas para creación de Interfaces:	DevComponents.DotNetBar.11.7.0.0_TRIAL
Herramientas para conectividad a las bases de datos	Npgsql.dll Mono.Security.dll MySql.Data.DLL
Herramientas para conectividad ODBC	Driver do Microsoft Excel (*.xls) Microsoft Access Driver (*.mdb,*.accdb) MySQL ODBC 5.2 ANSI Driver PostgreSQL ANSI SQL Server

Para la ejecución del proyecto se necesita un servidor de bases de datos SQLServer 2012 que gestionara toda la información del sistema IAN-BSC.

La investigación permitió guiar el diagrama de bases de datos de sistema con un trabajo de tesis desarrollado anteriormente (Figura 1 Anexos), lo que ayudó a desarrollar el diagrama y ajustarlo a las necesidades de BSC de la UTEQ. En la Figura 1 se observa las tablas del

diagrama de base de datos del sistema, lo cual nos un bosquejo muy claro de los módulos que va tener el IAN-BSC.

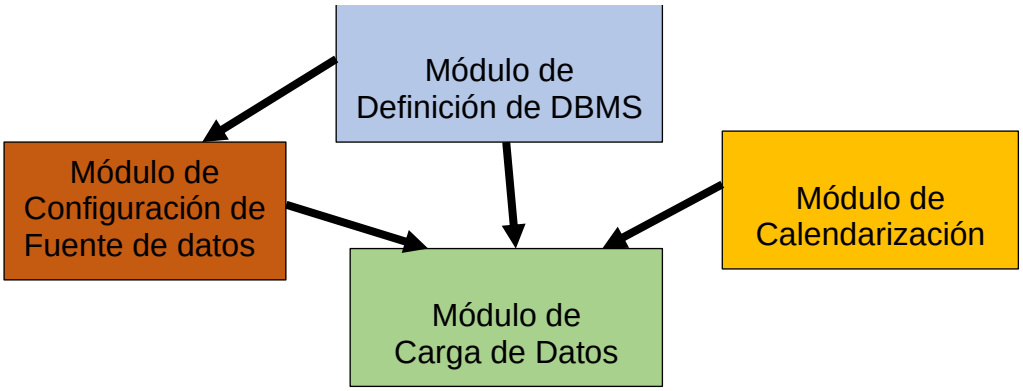
Figura 3. Diagrama de base de datos y su distribución por módulos.



El diagrama del sistema demuestra cómo va a influir el sistema IAD-BSC al Sistema de Balance ScoreCard de la Universidad Técnica Estatal de Quevedo. Las tablas "Variables y HistoriaVariable" son las que se van a utilizar de nexos entre el sistema de IAD-BSC y el Sistemas BSC.

En la figura 4 se explica cómo interactúan los módulos del sistema IAD-BCS, las dependencias que existen al momento de trabajar en cada módulo.

Figura 4. Iteraciones entre módulos del sistema

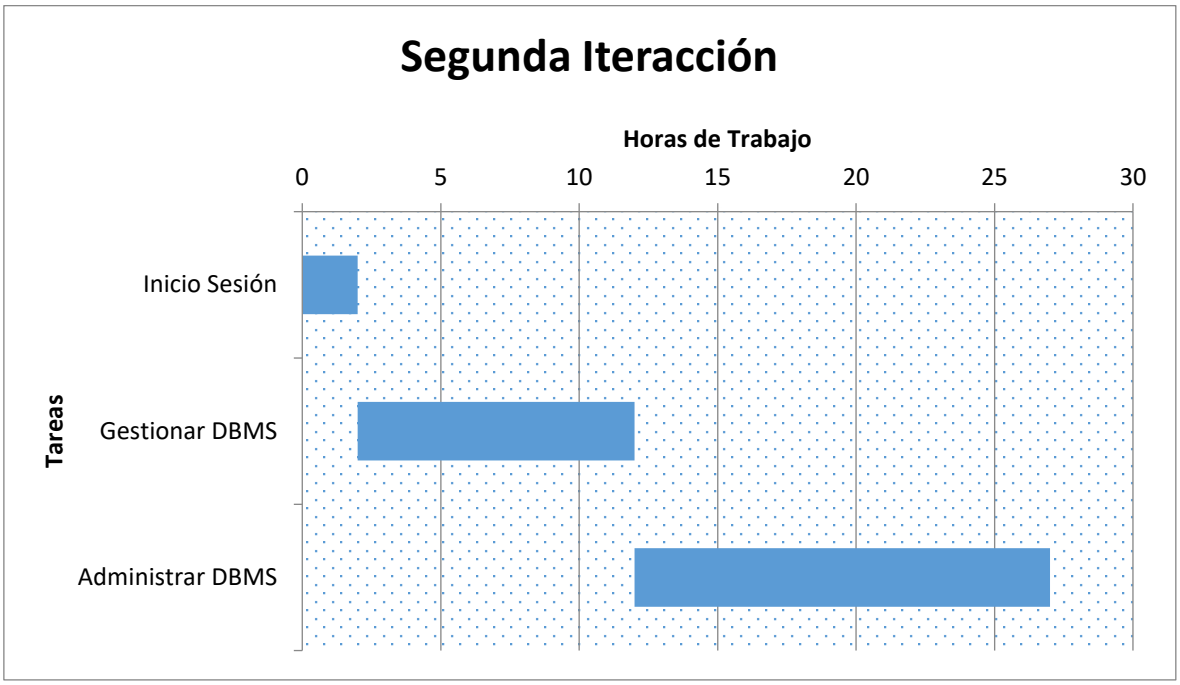


3.5.4.2. Segunda Iteración: Módulo de definición DBMS.

Una vez que se diseñó la base de datos del sistema IAD-BSC, se estableció la arquitectura para los módulos que van a componer el sistema. En la segunda iteración del desarrollo del proyecto se va a trabajar en el módulo de definición de DBMS, que consta de una épica y una historia de usuario que va a estar en la primera entrega.

El proyecto se va desarrollar por una persona esto hace que su creación se vaya dando secuencialmente como se puede observar en el Grafico 1. Las historias de usuarios que van a estar en esta iteración son inicio de sesión, gestionar DBMS y administrar DBMS.

Gráfico 1. Desarrollo de la Segunda Iteración.



El tiempo estimado para el desarrollo de todas las historia de usuarios que componen esta iteración es de 27 horas.

Inicio de Sesión. Para iniciar el desarrollo de esta historia de usuario es necesario definir los tipos de usuarios que van a utilizar el sistema IAD-BSC por lo cual se efectuaron reuniones con las parte interesadas en el proyecto y se definieron los tipos de usuarios que van a acceder al sistemas IAN-BSC diferenciándolos de los usuarios del Sistema de Balance ScoreCard, los cuales mencionaremos continuación:

- **Gestor de Datos.** Es un usuario encargado de administrar todo el sistemas IAD-BSC, debe tener conocimiento sólidos sobre el manejos de conexiones hacia los gestores DBMS, tendrá que crear las cargas de datos para que posteriormente se extraiga información de las bases de datos, vigilar que se cumpla correctamente la ejecución del calendario de carga de datos, enviar las notificaciones correspondiente cuando se produzcan atrasados en la cargas.

- **Digitador de Datos.** Este usuario va a tener dos tareas en el sistema IAD-BSC la primera es cuando se necesite ingresar datos manualmente a la historias de la variable y la segunda es cuando se necesite ingresar datos a los parámetros de la carga de datos.

- **Administrador BSC.** Se encargará de actualizar el calendario de carga de datos de las variables, este usuario no estará vinculado a otro proceso dentro del sistema IAD-BSC. Si las carga no se efectúan por un error se deberá notificar al gestor de datos para que este actué y solucione dicho inconveniente.

Gestionar Versiones DBMS. Este proceso se encargará de facilitar el trabajo de las versiones DBMS registradas en sistema IAD-BSC, dando facilidades a los usuarios al momento de buscar, crear, editar o eliminar una versión DBMS.

Administrar versiones DBMS. Este proceso ayudará a la creación y edición de las versiones DBMS, definirá el tipo de conexión y la versión del DBMS. En el sistema IAD-BSC, se dieron prioridad a los DBMS con los que trabajan los sistemas de la Universidad Técnica Estatal de Quevedo. También se implementó las conexiones a los datos mediante OBDC para acceder a la información que la universidad trabaja en hojas de cálculo (Excel) y no se encuentre en los DBMS de la universidad.

La tabla de la bases de datos del sistema IAD-BSC que van a ser utilizadas por el módulo de definición DBMS se muestran la figura 2.

Figura 5. Tabla del Módulo de definición DBMS.

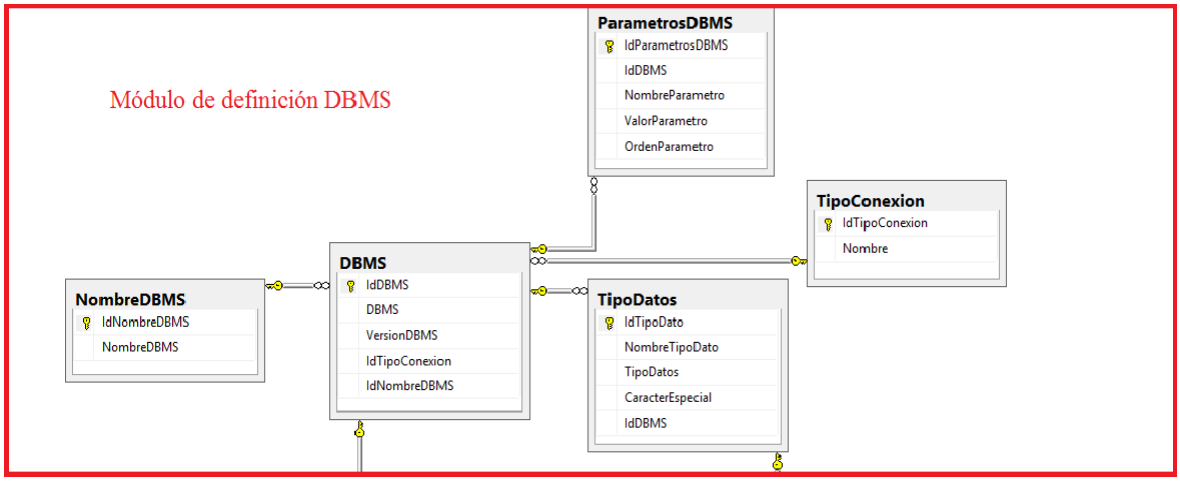
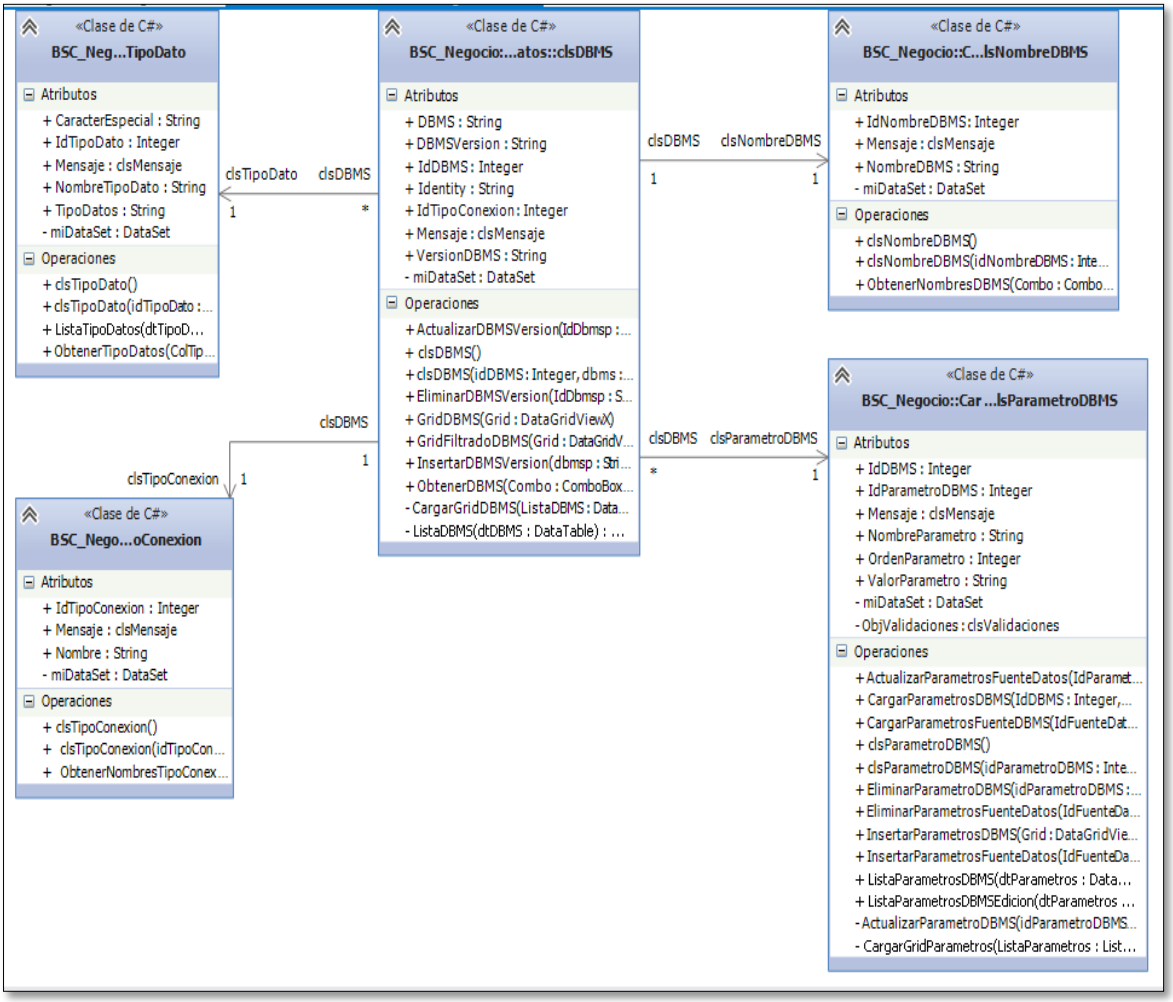


Figura 6. Diagrama de clases del Módulo de definición DBMS

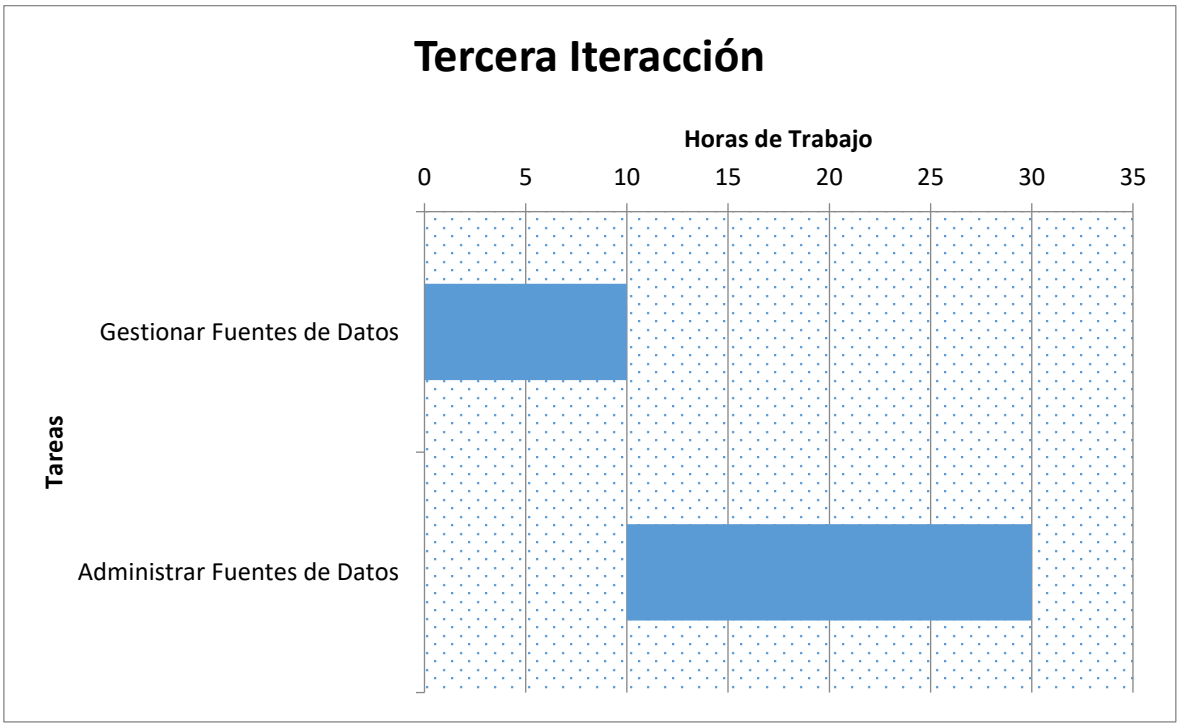


3.5.4.3. Tercera Iteración: Módulo de Fuentes de Datos.

En la tercera iteración se trabaja con las fuentes de datos que van a proveer la información necesaria para el repositorio del Balance ScoreCard, en la reuniones efectuadas con las partes interesadas se establecieron los lineamientos para el desarrollo de las interfaces y las formas que van a funcionar. Las fuentes de datos pueden ser bases de datos a archivos como las hojas de cálculos que se pueden conectar mediante ODBC.

El Gráfico 2. Se pueden observar las historias de usuarios que se establecieron para desarrollar este módulo.

Gráfico 2. Desarrollo de la Tercera Iteración.



El tiempo estimado para el desarrollo de las historia de usuarios gestionar fuentes de datos y administrar fuentes de datos que componen esta iteración es de 30 horas. Es un módulo sencillo pero que va a ayudar a facilitar el trabajo en el módulo de carga DBMS.

Debido a la gran cantidad de información con la que la UTEQ trabaja se hace necesario tener una clasificación que ayude a identificar rápidamente cuando se necesite tener acceso a ella, una de la formas de hacerlo es homologarlas en una lista y que no importe el sistema de almacenamiento que esta utilice. De esa forma es que han planteado las fuentes de datos para extraer la información.

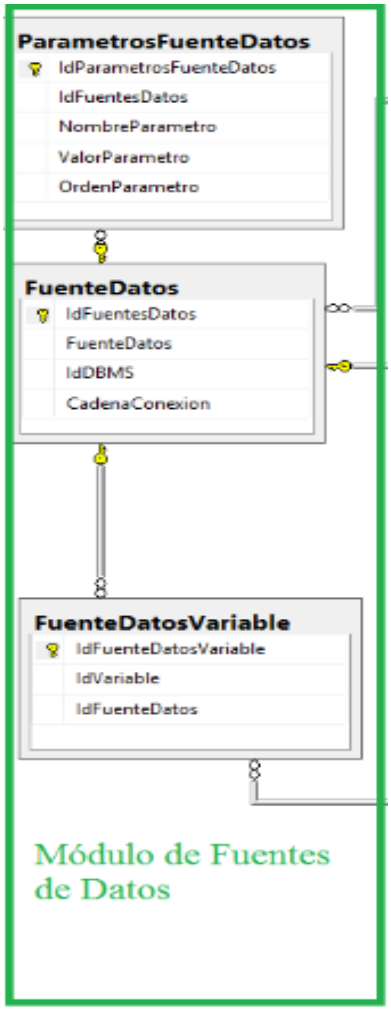
Gestionar Fuentes de Datos. Se presenta una lista que ayudará a la creación, edición y eliminación de las fuentes de datos. Este proceso gestionará todas las fuentes de datos que el sistema IAD-BSC utilizará para alimentar de datos al Balance ScoreCard de la UTEQ.

Administrar Fuentes de Datos. Este proceso ayudará a la creación y edición de las fuentes de datos, definirá los datos para los parámetros que se crearon en la definición DBMS, se tendrá que realizar la conexión correcta y demostrar que existe la fuente de

datos a la que se quiere acceder. En el sistema IAD-BSC tendrá las fuentes de datos con los que trabajan los sistemas de la Universidad Técnica Estatal de Quevedo.

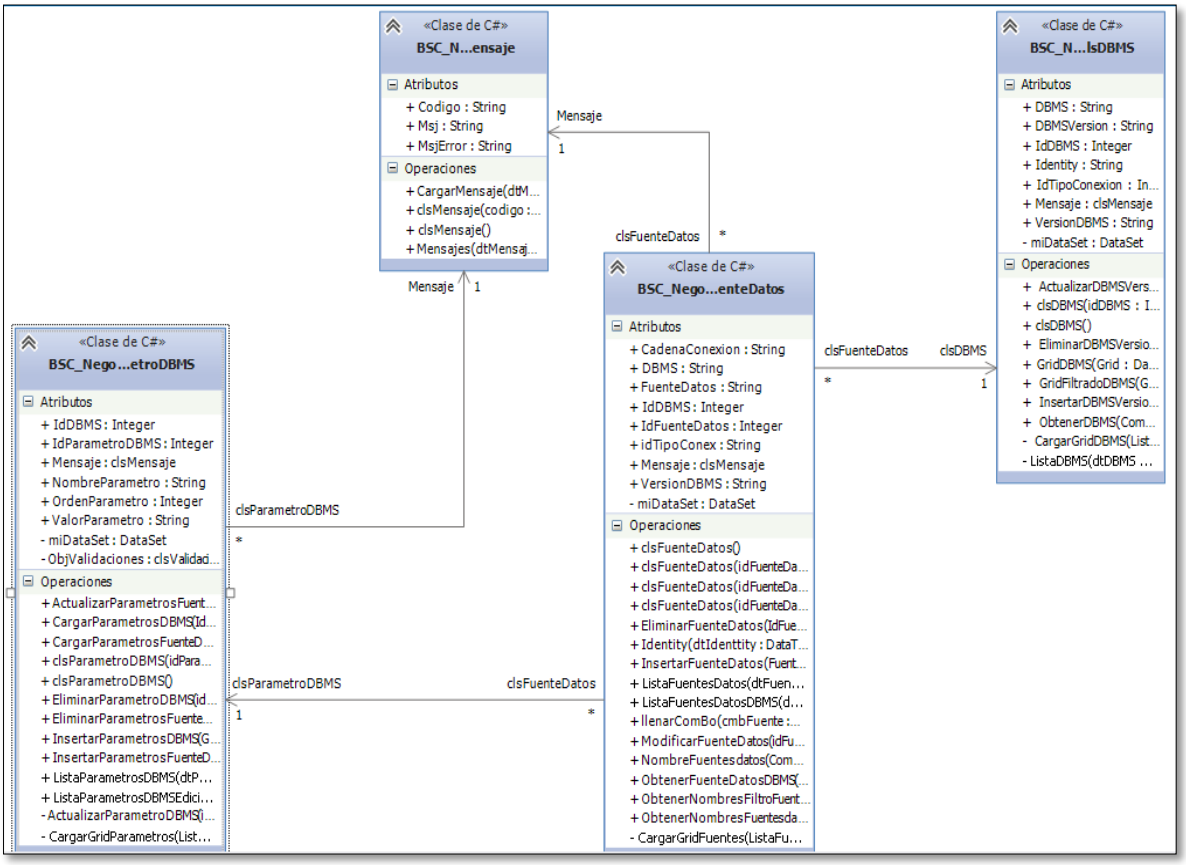
Este módulo dentro del sistema IAD-BSC es el más pequeño como se puede observar en la figura 4 solo se cuenta con tres tablas en la bases de datos para este módulo.

Figura 7. Tabla del Módulo de Fuentes de Datos.



En la figura 5 se muestra el diagrama de clases con el que se constituyó el módulo de fuente de datos.

Figura 8. Diagrama de clases del Módulo de Fuentes de Datos.

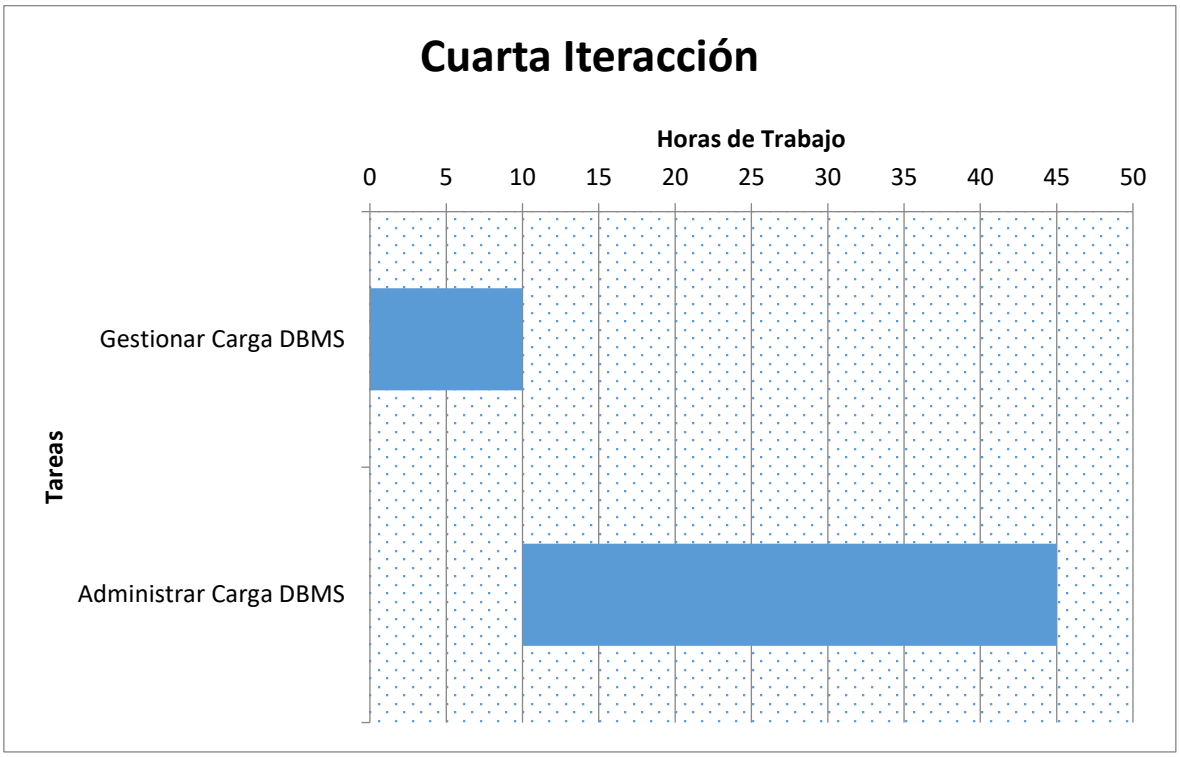


3.5.4.4. Cuarta Iteración: Módulo de Carga DBMS.

En la cuarta iteración se trabaja con las carga DBMS que es el mecanismo que va a permitir extraer información desde las fuentes de datos, este proceso incluye principalmente la parte SQL que va realizar las consultas hacia las fuentes de datos, ya sea estas DBMS u hojas de cálculo que contengan información relevante para el Balance ScoreCard de UTEQ

El Grafico 3 expresa el tiempo que se tomó en desarrollar la cuarta iteración que incluye dos historias de usuario que son: Gestionar cargas DBMS y Administrar cargas DBMS

Gráfico 3. Desarrollo de la Cuarta Iteración.



El tiempo estimado para el desarrollo de todas las historia de usuario gestionar carga DBMS y administrar carga DBMS de 45 horas, es uno de los módulos más importante en el sistema IAD-BSC.

El módulo carga DBMS crea la configuración para acceder y extraer la información que el usuario necesita extraer desde las fuentes de datos, la parte principal es la sentencia SQL para ejecutar la carga y si esta necesita parámetros para funcionar habrá que agregarlos, se definirá el tipo de conexión (Procedimientos almacenados, vista y sentencias SQL) que utilizará la carga.

Gestionar cargas DBMS. Se presenta una lista que ayudará a la creación, edición y eliminación de las cargas DBMS. Este proceso gestionará todas las cargas DBMS que el sistema IAD-BSC para extraer información desde las fuentes de datos al Balance ScoreCard de la UTEQ. Para eliminar una carga esta no deberá tener parámetros.

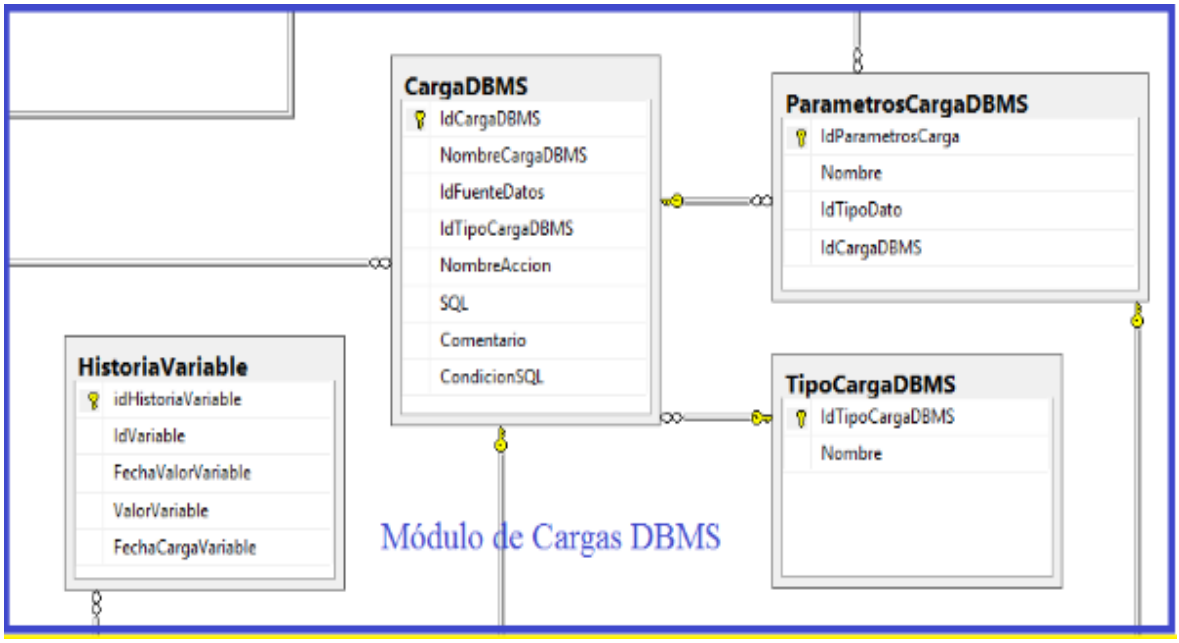
Administrar cargas DBMS. Este proceso ayudará a la creación y edición de las cargas DBMS, dependiendo del tipo de conexión que se va a realizar a la fuente de datos, se fijara la sentencia SQL para carga si esta necesita parámetros se definirán el nombre y el tipo de

datos, el valor de parámetro solo se incluirá para realizar las pruebas sobre el funcionamiento de la carga, este valor no será almacenado, para registrar el cambio en la carga es obligatorio ejecutar la carga de pruebas para validar que esta cumple con los requerimiento. Una fuente de datos puede tener varias cargas DBMS.

Las cargas DBMS son las que sirven de enlace entre el sistema IAD-BSC y el Balance ScoreCard, de donde se debe seleccionar una al momento de crear una variable, se deberá crear una carga DBMS para una variable.

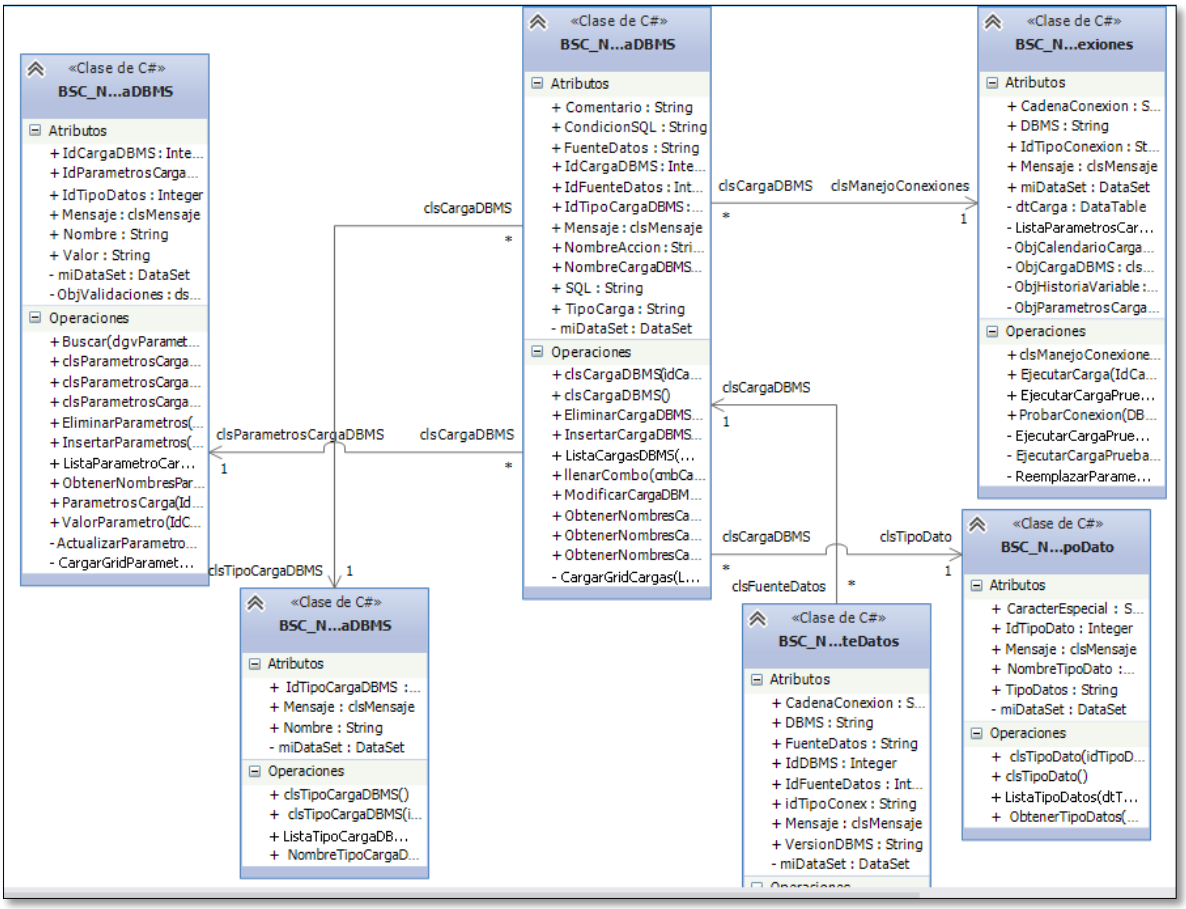
El diagrama de bases de datos con las tablas para el módulo de carga DBMS se muestra en la figura 6, donde se pueden ver las relaciones que existen entre las tablas del módulo.

Figura 9. Tabla del Módulo de Carga DBMS.



El diagrama de clases que indica el funcionamiento del módulo de carga DBMS se muestra en la figura 7

Figura 10. Diagrama de clases del Módulo de Carga DBMS.

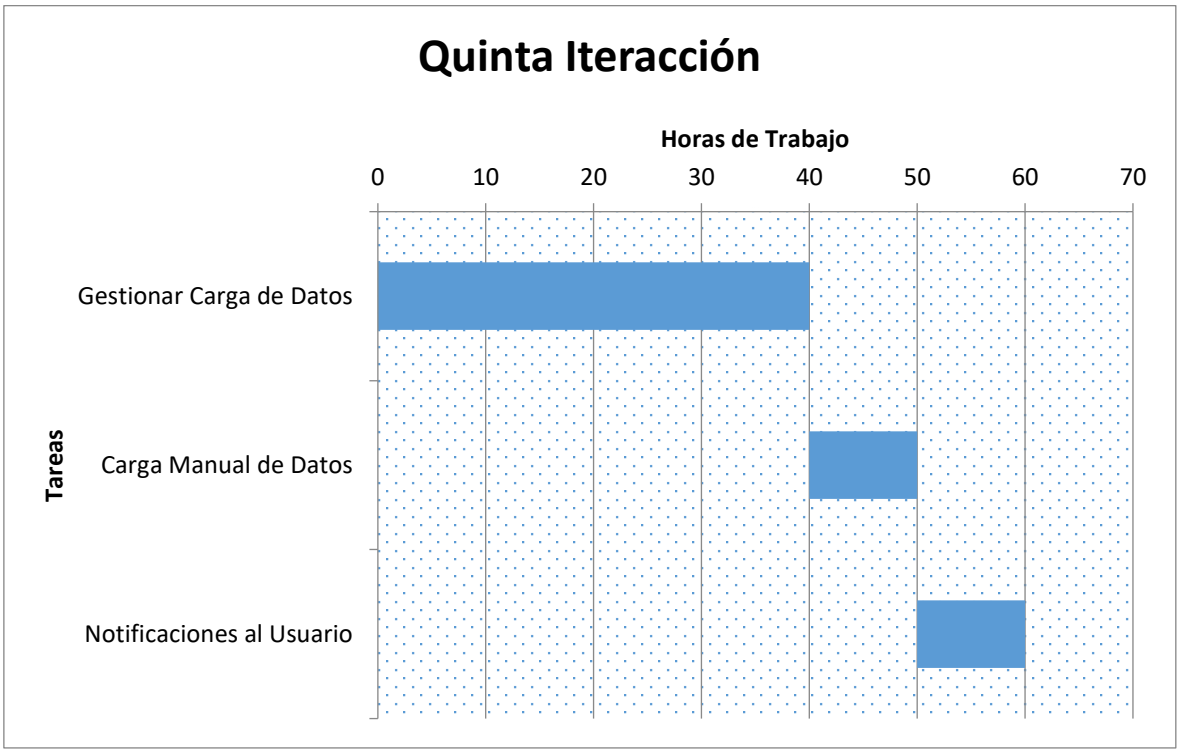


3.5.4.5. Quinta Iteración: Módulo de Calendarización.

En la quinta iteración se trabaja con el calendario de carga de datos automático desde la fuentes de datos, el calendario de carga de datos manual, el envió de notificaciones en caso de errores o atrasos en la carga. Es el mecanismo que va a permitir tener actualizada la información de las historias de la variables, que posteriormente servirán para los indicadores, Este procesos es la parte fundamental para el IAD-BSC porque permite tener actualizada y validada toda las información de las variables ya sea desde la fuentes de datos o la que ha sido cargada manualmente por los usuarios encargados.

El Gráfico 4 expresa el tiempo que se tomó en desarrollar la quinta iteración que incluye tres historias de usuario que son: Gestionar cargas de datos, Carga manual de datos y Notificaciones a los usuarios.

Gráfico 4. Desarrollo de la Quinta Iteración.



El tiempo estimado para el desarrollo de todas las historia de usuarios que componen esta iteración es de 60 horas, siendo uno de los módulos más complejos del sistema IAD-BSC.

El módulo calendarización es el que se encarga de gestionar toda la información que necesitan las variables en su historial, esto lo realiza a través de un calendario que previamente se debe definir cuándo se crea una variable. Se manejarán estados para controlar la carga de información. La información que se extraigan desde las fuentes de datos de la Universidad Técnica Estatal de Quevedo se realizará automáticamente mediantes las cargas DBMS y la información que no se encuentre en las fuentes de datos y que sean necesarios para el Balance ScoreCard se deberá ingresar manualmente.

Gestionar cargas de datos. Se presenta el calendario con las variables pendientes de actualizar su historial de datos, este calendario tiene un filtro de fechas, la fecha actual es la fecha máxima al ser consultada en el filtro y la fecha mínima es la fecha en que inicia el periodo del Balance ScoreCard, este calendario de maneja por estado (Actualizado, Pendiente y Error) que deberá cambiar una vez ejecutada la carga correspondiente.

También consta un TAB que permite consultar el estado de la historia de cada variable según su estado, esto permitirá tener un mejor control en la carga de cada variable.

Carga manual de datos. Este forma de carga de datos se utilizará principalmente para actualizar la información de la historia de las variables que no se encuentran en la fuentes de datos con las que trabaja la Universidad Técnica Estatal de Quevedo pero que es necesaria para el Balance ScoreCard.

Se deberá definir los usuarios se serán los responsables para ingresar la información a las historia de las variables.

Notificaciones al usuario. Se utilizará principalmente para enviar notificaciones a los usuarios a través de correo electrónico cuando existan cargas de datos ya sean automáticas o manuales con errores o pendientes de carga.

El diagrama de bases de datos con las tablas para el módulo de calendarización se muestra en la figura 8, donde se pueden ver las relaciones que existen entre las tablas del módulo.

Figura 11. Tabla del Módulo de Calendarización.

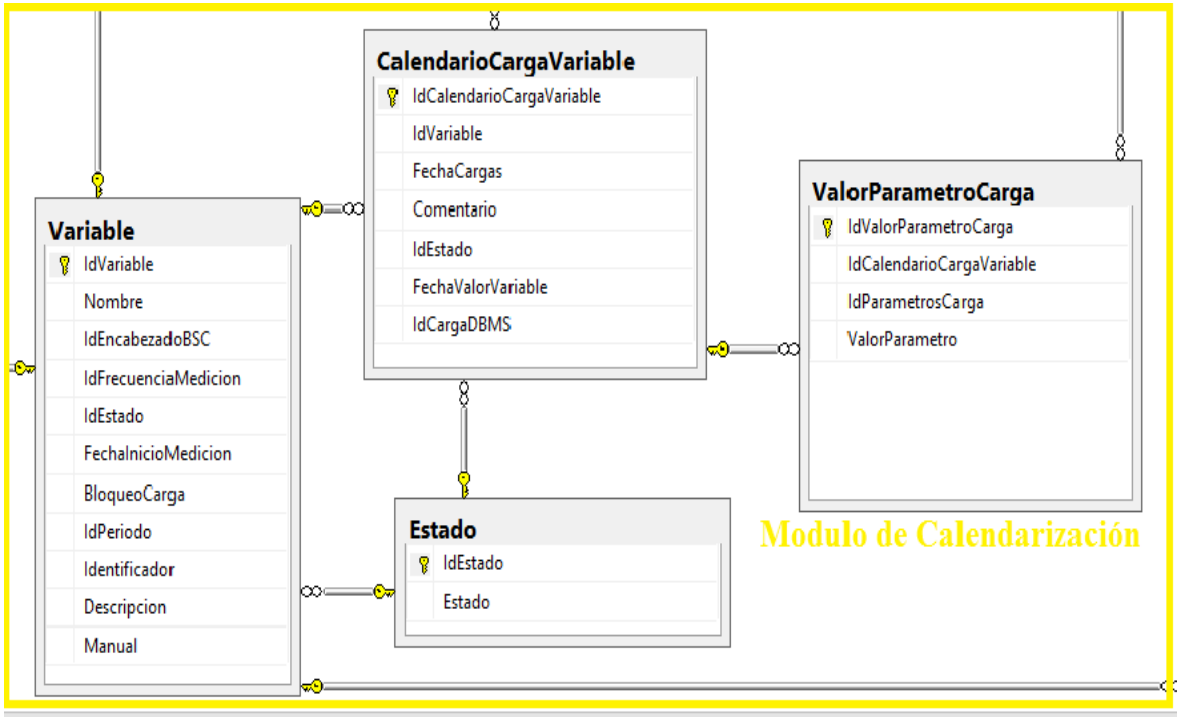
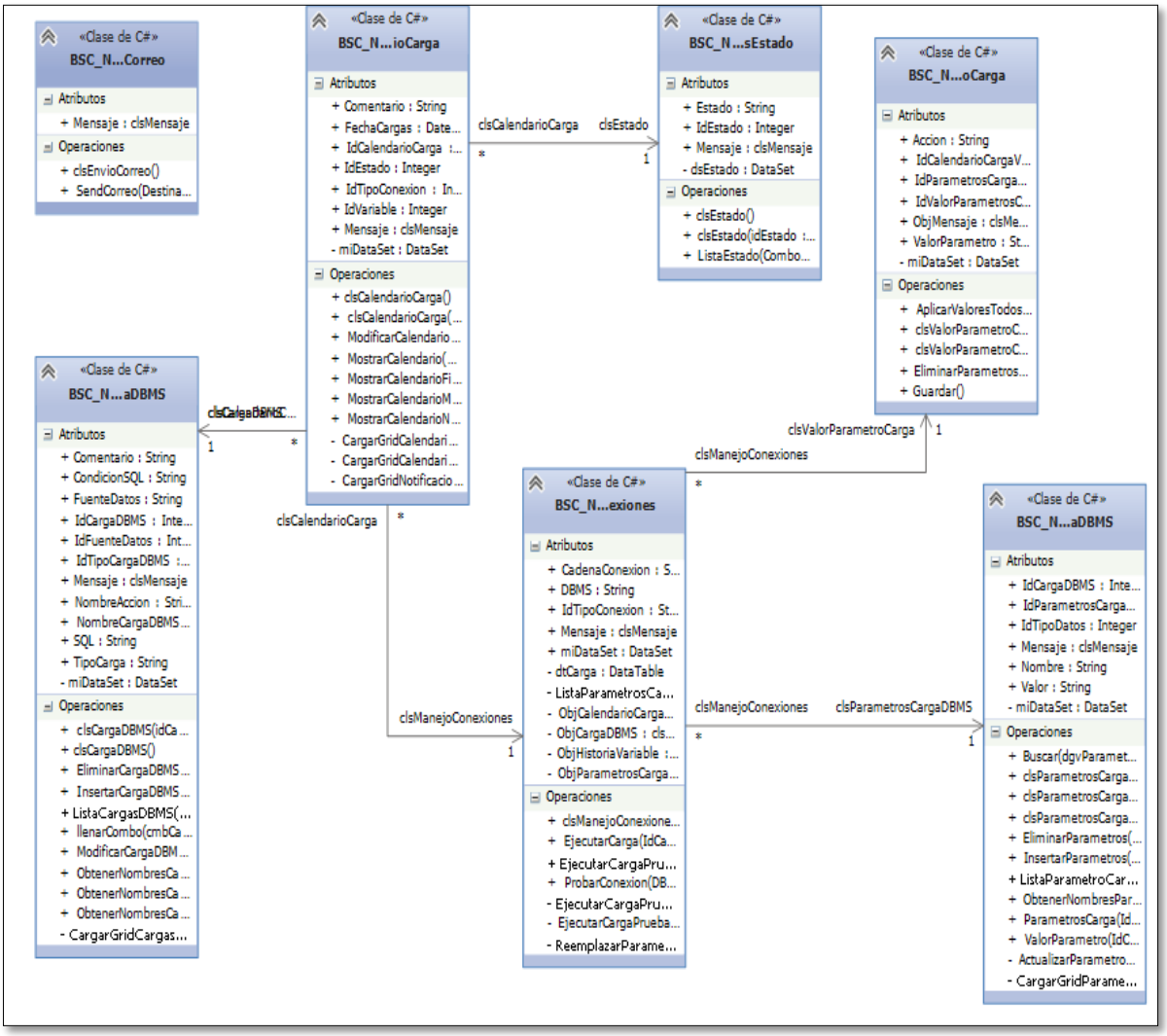


Figura 12. Diagrama de clases del Módulo de Calendarización.



3.5.5. Fase de Producción.

El sistema IAD-BSC se trabajó con el cliente que es la Unidad de Planificación Académica (UPA), donde se establecieron todos los requisitos mediantes las historias de usuarios antes de entrar en el ambiente de producción del sistema, las correcciones que se realizaron se dieron después de hacer las pruebas correspondientes.

Para que el sistema funcione correctamente en producción se debe complementar con el sistema de Balance ScoreCard.

3.5.6. Fase de Mantenimiento.

El sistema IAD-BSC de la Universidad Técnica Estatal de Quevedo fue implementado en la Unidad de Planificación Académica (UPA). El mantenimiento del sistema se debe realizar mediante peticiones a los administradores.

Se deben realizar la peticiones de mantenimiento mediante una solicitud a los administradores (Ver anexo 4), luego serán analizada todas las solicitud que se envíen y se darán prioridades urgentes, normal o critica a las misma.

Según cada requerimiento se deberá analizar si este puede causar variaciones en el diseño o las estructuras del proyecto, o si este afecta al funcionamiento del Balance ScoreCard de la Universidad.

La solicitud de mantenimiento tendrá como plazo un mes para ser atendida y dar las respectivas comunicaciones a los solicitantes. Si cambia el funcionamiento del sistema se deberá informar a todas las partes interesadas, mostrar los prototipos a todos los usuarios.

Cuando entren en funcionamiento las correcciones y observaciones realizadas se deberán realizar las respectivas pruebas.

La primera versión del sistemas es la 1.0.0, cuando se realicen cambios en el funcionamiento se deberá cambiar el código de la versión, por ejemplo la segunda versión sería la 1.0.1, y así posteriormente las siguientes.

3.5.7. Fase de muerte de proyecto.

Se adjunta las respectiva documentación del desarrollo de la metodología XP, se han incluido todas las correcciones pedidas por el cliente y el sistema IAD-BSC se ha integrado completamente con el sistemas Balance ScoreCard de la Universidad Técnica Estatal de Quevedo.

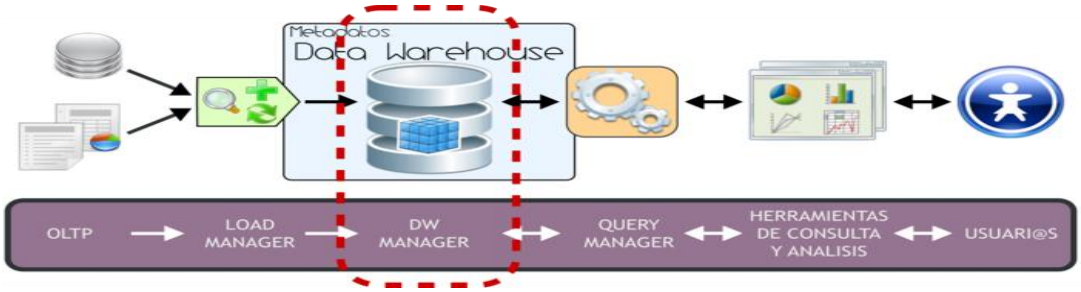
CAPÍTULO IV

RESULTADOS Y DISCUSIÓN

4.1. Resultados y Discusión

La metodología de carga de datos para el Cuadro de Mando Integral de la Universidad Técnica Estatal de Quevedo se fundamenta en un proceso ETL (Extracción, Transformación y Carga) adaptado a la problemática que se plantea en este proyecto. La adaptación ocurre principalmente en el hecho de que las etapas del ETL generalmente constas de un proceso de extracción, transformación y carga de datos. Pero en la adaptación solo se tomó en cuenta las partes que son de utilidad para cargar datos a un BSC. Las tareas propuestas que componen la metodología de carga de datos partió del diseño explicado por Trujillo y Lijan (2003), pero con variantes con respecto a la transformación de las fuentes y dejando en efecto la unión de los atributos de las fuentes de datos con los atributos del destino. En consecuencia la metodología propuesta sería:

Figura 13. Arquitectura Data Warehouse



Fuente: http://www.dataprix.com/datawarehouse-manager

a) Seleccionar los datos para la extracción.

- La selección de datos de datos la debe realizar el Administrador del BSC debido a que debe crear las variables del Cuadro de Mando Integral las cuales van a solicitar la información que se van a extraer.
- El software que administra la integración de datos tiene dos formas de conectarse a las fuentes de datos que son: La OleDB para conexiones a los DBMS y La ODBC que se usa principalmente para conectar con los archivos de Excel y Access.

b) Definir los datos de las fuentes.

- El gestor de datos debe tener hecha toda la configuración para acceder a la información que se requiere.

- Una vez que se crean las variables se debe seleccionar la fuente de datos que contiene la información que se necesita.
- Se deben definir los parámetros de conexión a las fuentes de datos.
- Se define la consulta SQL para extraer la consulta.

c) Definir los datos que no se encuentran en las fuentes.

- El software de Carga de datos ha creado interfaces para se pueda ingresar la información que no se encuentra en ninguna fuente de datos de la UTEQ

d) Unir las fuentes.

- El software administra una biblioteca de Fuentes de Datos de las diferentes Unidades Académicas y Administrativas de la UTEQ.

e) Seleccionar el destino de carga.

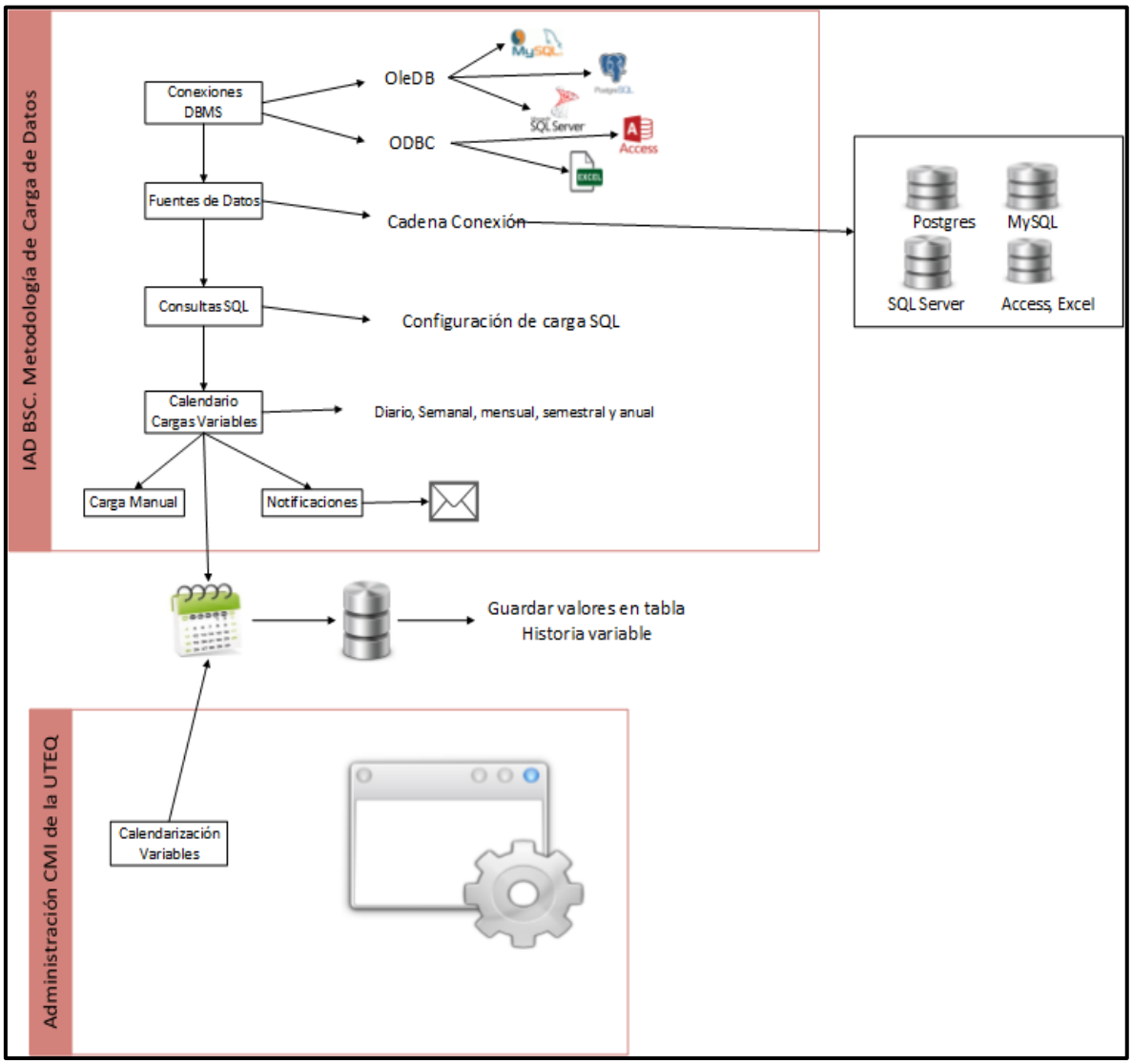
- El destino de carga o almacenamiento de la información extraída siempre va ser la base de datos que utiliza el software que administra el Cuadro de Mando Integral de UTEQ.
- Toda la información va ser almacenada en la Tabla HistoriaVariable.

f) Carga los datos en el repositorio.

- Para carga la información en el repositorio se realiza a través de un calendario personalizado para cada variable del Cuadro de Mando Integral, para las variables que extraen información desde la las fuentes de datos este proceso se realiza de forma automática.
- Las variables que necesitan información que se ingrese de manera manual debe realizarlo por medio de las interfaces desarrolladas para este propósito.

En la figura 12 se explica de forma gráfica la arquitectura del Software de Carga de Datos del Cuadro de Mando Integral, dando un mejor entendimiento de aporte del Software a la Metodología planteada a este proyecto de investigación

Figura 14. Metodología de Carga de Datos aplicada al software.



En el desarrollo del software se basó en el proceso ETL, basado en una arquitectura Data Warehouse, que sirvió como referencia. Sin embargo solo se toma las fases iniciales. Como sugiere Alveiro & Rico [28], esta arquitectura tiene como ventaja almacenar la información extraída en un solo repositorio.

La metodología de carga de datos se integra con la aplicación BSC que realiza el seguimiento del plan estratégico de Universidad Técnica Estatal de Quevedo mediante el calendario personalizado de cada variable del plan estratégico.

Este proyecto de investigación se planteó ante la necesidad de tener una herramienta, que proporcionara información actualizada, validada y oportuna, para el Cuadro de Mando

Integral de la Universidad. La alimentación de datos va a estar menos expuesta a errores que se pueden cometer cuando se realiza la carga de información de forma manual, se extraen datos desde las diferentes bases de datos para simplificar a lo máximo el trabajo de los usuarios. Por ese motivo se automatizan los procesos en la carga de datos, que es una de las ventajas que se presenta al momento de utilizar esta herramienta porque se puede reducir costos y ayudan a tener un mejor manejo en los errores que se pueden producir en este proceso de carga como lo mencionan Carhuatocoto, Noriega &Hildebrando [29] y Gonzáles [30] en sus trabajo.

Se comprobó que la carga automatizada facilita la productividad del software desarrollado en este proyecto como lo señala Betancur & Moreno [24] que realizaron una aplicación que recopila la información de cuatro sistemas, parecido al software de este proyecto que agrupa la información de varios sistemas de la universidad.

Se lleva un cronograma detallado por variable para cargar la información necesaria para el cuadro de mando integral debido a que se utiliza la metodología de carga de datos implementada en un software específico para realizar esta tarea [31], para evitar que registren atrasos se cuenta con la opción de notificación periódicas por correo electrónico a los encargados de registrar información a la variables, donde se busca que los usuarios realicen este proceso de carga en los tiempos programados. La plataforma de correo de electrónica para enviar las notificaciones es GMAIL que es la misma plataforma con la que trabaja la UTEQ.

Se libera en gran parte a los departamentos de la Universidad de la carga de trabajo de enviar información para el Balance Scorecard como se lo realizaba anteriormente al desarrollo de la metodología, el software que administra esta metodología se encarga de realizar esta tarea bajo la supervisión de los encargados del mismo.

Analizando la metodología de carga de datos se puede tener una visión general del planteamiento del sistema gestor de conexiones para el Cuadro de Mando Integral de la Universidad Técnica Estatal de Quevedo, de donde se pueden sacar varias conclusiones funcionalidades del software aplicado para este proyecto, entre las cuales tenemos:

- El sistema Balance Scorecard requiere que el acceso a datos sea lo más simple y optimo posible para las personas que están encargadas de realizar el seguimiento del plan estratégico.
- Se deberán identificar todas las fuentes de datos de la Universidad para definir y extraer la información que el Cuadro de Mando Integral necesita para sus indicadores.
- Las configuraciones de acceso a los datos las realiza el Gestor de Datos, por lo cual se hace un proceso transparente para los demás usuarios del Sistema BSC.
- Las cargas de datos se pueden realizar de varias fuentes de datos (DBMS, Archivos como Excel), dependiendo del tipo de conexión al que se adapte por ejemplo OleDB u ODBC.
- Las cargas de datos pueden llevar parámetros, que se utilizaran para realizar las consultas a las fuentes de datos, el gestor de datos deberá realizar las pruebas necesarias para que la carga de datos no sufra inconvenientes a momento de que los digitadores de datos realicen las carga automáticas de información.
- Se planificará un calendario personalizado para cada variable del Cuadro de Mando Integral lo que permitirá llevar un control en las carga de datos.
- La medición de los indicadores dependerá de las cargas de datos periódicas que se le realicen a las variables por medio del calendario que se planifico al momento de crear cada variable.
- Cuando las variables necesiten datos que no se encuentre en una fuente de datos de la Universidad, el digitador de datos registrará manualmente la información mediante las interfaces creadas para este proceso según el calendario planificado.
- Cuando la carga de datos no se realice de acuerdo a la planificación ya sea por falta de datos en los parámetros, errores o atraso en las cargas, errores en las conexiones se le notificara al usuario encargado de digitar los datos de la variable. Para realizar las notificación se utilizó el medio del correo electrónico esto obliga que al momento de registrar un usuario es indispensable que se registre el correo electrónico.
- Se utilizó la plataforma de correo de Google que es la misma plataforma de correo que utiliza la Universidad.

CAPÍTULO V

CONCLUSIONES Y RECOMENDACIONES.

5.1. Conclusiones.

- La definición de la metodología propuesta para la alimentación de datos desarrollada, fundamentada en un proceso ETL adaptada a las necesidades del Cuadro de Mando Integral de Universidad Técnica Estatal de Quevedo, resultó exitosa; logrando un mejor manejo en el control de la información que se extrae desde las diversas fuentes de datos hacia las variables que tiene el plan estratégico institucional.

- La utilización de librerías, la parametrización de datos en la cadenas de conexión, la clasificación de la fuentes de datos y la definición de metadatos en las consultas SQL aportaron a consolidar un mecanismo sólido para extraer la información desde la diferentes fuentes de datos existente y futuras en la Universidad Técnica Estatal de Quevedo.

- La implementación de cronogramas personalizados para cargar la información a cada variable mejora el seguimiento y control por parte del personal de la Unidad de Planeamiento Estratégico (UPE). Esto permitió implementar notificaciones por medio de correo electrónico a las personas responsable de realizar esta tarea. Se definió que esta funcionalidad puede ser utilizada para realizar acciones de modo preventivo (cuando se envían advertencias para realizar oportunamente la carga) y modo correctivo (cuando existen atrasos o errores en la carga).

- La implementación de las interfaces para realizar las cargas manuales de datos ayudan a optimizar el seguimiento del Plan Estratégico Institucional debido a que se lleva un mejor control de errores en el registro de datos en el historial de la variable y también se responsabiliza a los usuarios por cada variable.

5.2. Recomendaciones.

- Para administrar el software del proyecto se requiere que los usuarios estén capacitados en administración de Bases de datos para llevar el manejo en las cadenas de conexiones, la creación en los enlaces con los ODBC y en la creación de las sentencias SQL, vista y procedimientos almacenados.

- Se aconseja que se establezcan políticas de seguridad en las bases de datos y servidores que van a ser fuente de información, de preferencia, que se creen usuarios a los que solo se les permita ejecutar consultas select para que no afecte a la integridad de la información almacenadas.

- El sistema fue configurado para trabajar específicamente con DBMS como MySQL, SQL Server y Postgres en la función para ejecutar procedimiento almacenados por lo que se aconseja agregar en un futuro nuevas conexiones a DBMS. Como por ejemplo la de Oracle.

CAPÍTULO VI

BIBLIOGRAFÍA.

[1] Universidad de Talca, «Sistemas de Información: Estrategia para la Gestión Empresarial,» 2013.

[2] F. Jaramillo, «LA COMPETENCIA EN EL MANEJO DE LA INFORMACIÓN: UN IMPERATIVO EN LA SOCIEDAD Y UNIVERSIDAD MODERNA,» 2013.

[3] R. Kaplan y D. Norton, «Balanced Scorecard,» Global Network Content Services LLC, DBA Noticias Financieras LLC, Miami, 2014.

[4] CEAACES, «MODELO PARA LA EVALUACIÓN DE,» 2010.

[5] UPE, Interviewee, *Balance Scorecard UTEQ.* [Entrevista]. Mayo 2015.

[6] Instituto Potosino, «Utilización del Balanced Scorecard en el Instituto Potosino de Investigación Científica y Tecnológica, A.C. (Presentation),» de *IIE Annual Conference. Proceedings*, Norcross, 2013.

[7] A. Fernandez, «El balanced Scoredcard ayudando a implantar la estrategia,» *Revista de antiguos alumnos,* Marzo 2001.

[8] MECI, *Evaluación al Plan de Desarrollo Institucional Tablero de Comando: Balanced Score Card,* Tolima: Universidad del Tolima, 2010.

[9] C. Graver, *Diseño de un modelo de gestión estratégica, Balanced Scorecard, en minera el tesoro.,* Santiago: Universidad de Chile, 2007.

[10] C. Azabache, Sistema de Información Gerencial para la Gestión Estratégicaen ventas por medio de un cuadro de mando integral, Quevedo, 2014.

[11] D. Howe, «Free on-line dictionary of computing.».

[12] J. Senso y A. De la Rosa, «El concepto de metadato. Algo más que descripción de recursos electrónicos,» Universidad Granada, Brasilia, 2003.

[13] Z. Ercegovac, «Introduction. Journal of the American Society for Information Science».

[14] M. Albert y M. Hernadez, «El Cuadro de Mando Integral para la implementación y el control estratégico. Experiencias.,» ISPJAE, La Habana, 2004.

[15] R. Villarroel, Y. Gomez y C. Krause, «Incorporación de Seguridad en el Modelado Conceptual de Procesos Extracción-Transformación-Carga,» Pontificia Universidad Católica de Valparaíso, Valparaiso, 2013.

[16] R. Villarroel, Y. Gómez y C. Krause, «Incorporación de Seguridad en el Modelado Conceptual de Procesos Extracción-Transformación-Carga.,» 2013.

[17] L. Muñoz, J. Mazón y J. Trujillo, *Revisión sistemática y comparación de modelado*

de, 2010.

[18] J. Trujillo y J. Luján, *A UML Based Approach for Modeling ETL*, Chicago, 2003.

[19] L. Etcheverry, P. Gatto y S. Tercia, *Análisis del proceso de carga del Sistema de Data Warehousing de Enseñanza de la Facultad de Ingeniería*, 2005.

[20] A. Lavado y H. Noriega, *Automatización del proceso de carga de datos académicos para la Escuela de Ingeniería de Sistemas y Computación.*, Universidad Peruana de Ciencias Aplicadas - UPC, 2014.

[21] D. Sánchez, A. Pan y A. Viña, *EII, un nuevo paradigma para la integración de información dispersa y heterogénea en la administración. Experiencia: vixía, un servicio de vigilancia tecnológica*, Murcia, 2004.

[22] K. Laudon y J. Laudon, Sistemas de Información Gerencial, 12 ed ed., Pearson Education, 2012.

[23] M. Rodríguez, C. Macías y D. Casero, «El Cuadro de Mando Integral: La respuesta a la implantación exitosa de la estrategia en las universidades públicas españolas,» *LATIN AMERICAN AND CARIBBEAN JOURNAL OF ENGINEERING EDUCATION*, vol. 4, 2010.

[24] D. Betancur y J. Moreno, *Una Aproximación Multi-Agente para el Soporte al Proceso de Extracción-Transformación-Carga en Bodegas de Datos*, 2012.

[25] J. Cardona y A. Vallejo, *Desarrollo de un sistema de información computarizado basado en datos maestros, utilizando una metodología de indicadores balanceados de desempeño (balanced, scorecard), para apoyar la estrategia corporativa de la Universidad de Manizales*, Manizales, 2015.

[26] P. Letelier, «Métodologías ágiles para el desarrollo de software: eXtreme Programming (XP),» 2006. [En línea]. Disponible: http://www.cyta.com.ar/ta0502/v5n2a1.htm#.

[27] J. Cortizo, D. Expósito y M. Ruiz, «eXtreme Programming,» 2010, p. 17.

[28] R. Alveiro y D. Rico, *INTELIGENCIA DE NEGOCIOS: ESTADO DEL ARTE*, 2010.

[29] L. Carhuatocto, A. Noriega y H. Hildebrando, *Automatización del proceso de carga de datos académicos para la Escuela de Ingeniería de Sistemas y Computación*, 2014.

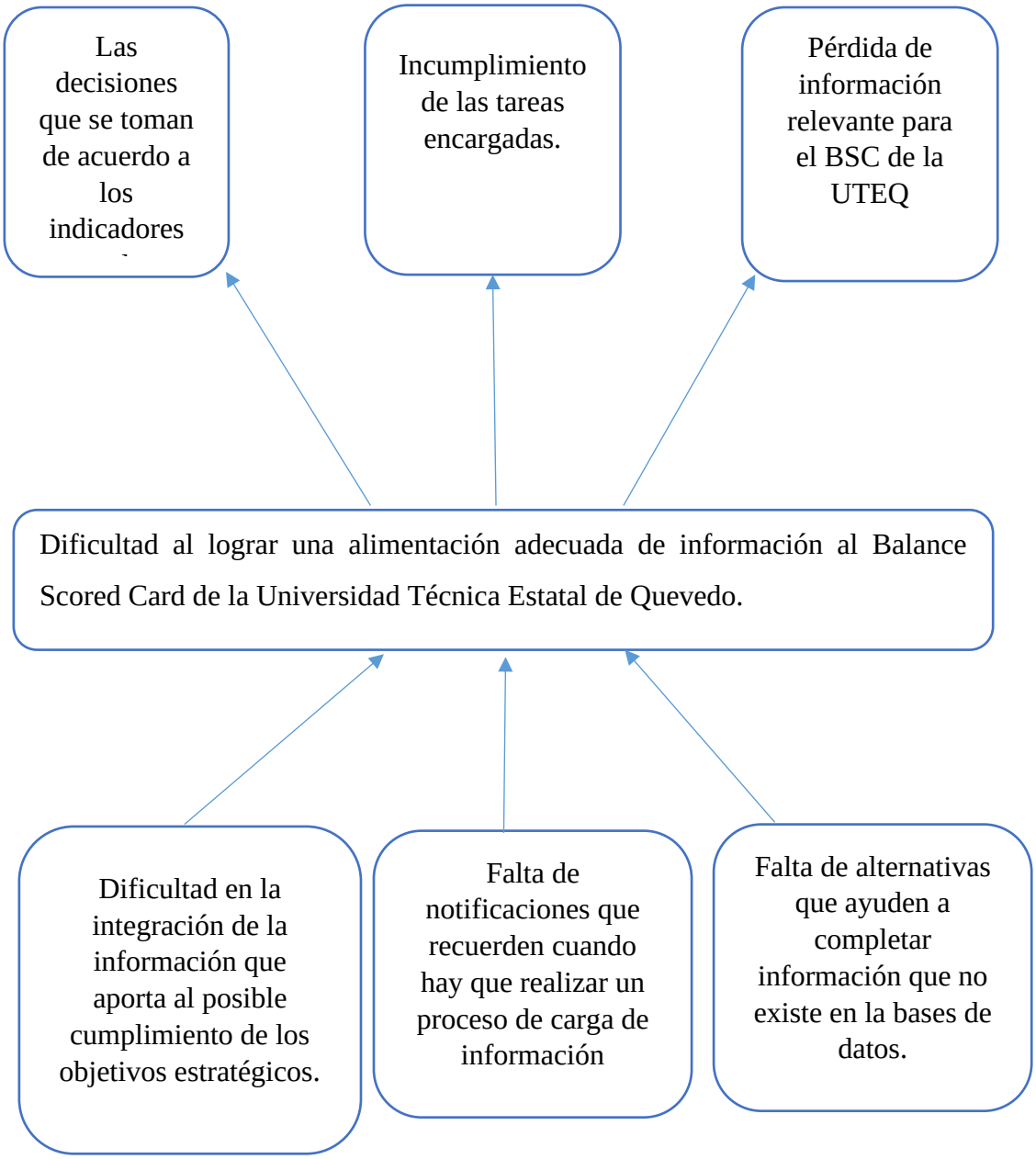
[30] J. Gonzáles, *Cómo automatizar Procesos con Excel hacia una Base de Datos*, 2010.

[31] IBM, «Comprensión de la planificación de la carga de trabajo dinámica,» 2014. [En línea]. Disponible: http://www-01.ibm.com/support/knowledgecenter/SSGSPN_8.5.1/com.ibm.tivoli.itws.doc_8.5.1/dqx1welcome.htm%23dqx1welcome?lang=es.

CAPÍTULO VII

ANEXOS.

Anexo 1. Árbol del problema



Anexo 2. Matriz de relación Problema, Objetivos

Problema General	Objetivos General
¿Qué metodología es la adecuada para la carga de datos en el Cuadro de Mando Integral de la Universidad Técnica Estatal de Quevedo?	Determinar la metodología apropiada para la carga de datos al Cuadro de Mando Integral de la Universidad Técnica Estatal de Quevedo.
Problemas Derivados	**Objetivos Específicos**
¿Cómo integrar la información almacenada en las diferentes fuentes de datos de la institución?	Definir los mecanismos para extraer la información de las diferentes fuentes de datos de la institución.
¿Cómo mejorar el cumplimiento de la carga oportuna de datos para el Cuadro de Mando Integral?	Desarrollar los métodos necesarios para calendarizar y notificar la cargar adecuada de los datos.
¿De qué forma se puede implementar la carga de información que no se encuentra en las fuentes de datos de la UTEQ?	Implementar las interfaces correspondiente para la carga datos que no se encuentran en las fuentes de datos de la UTEQ.

Anexo 3. Estructuras de las Historias de usuarios

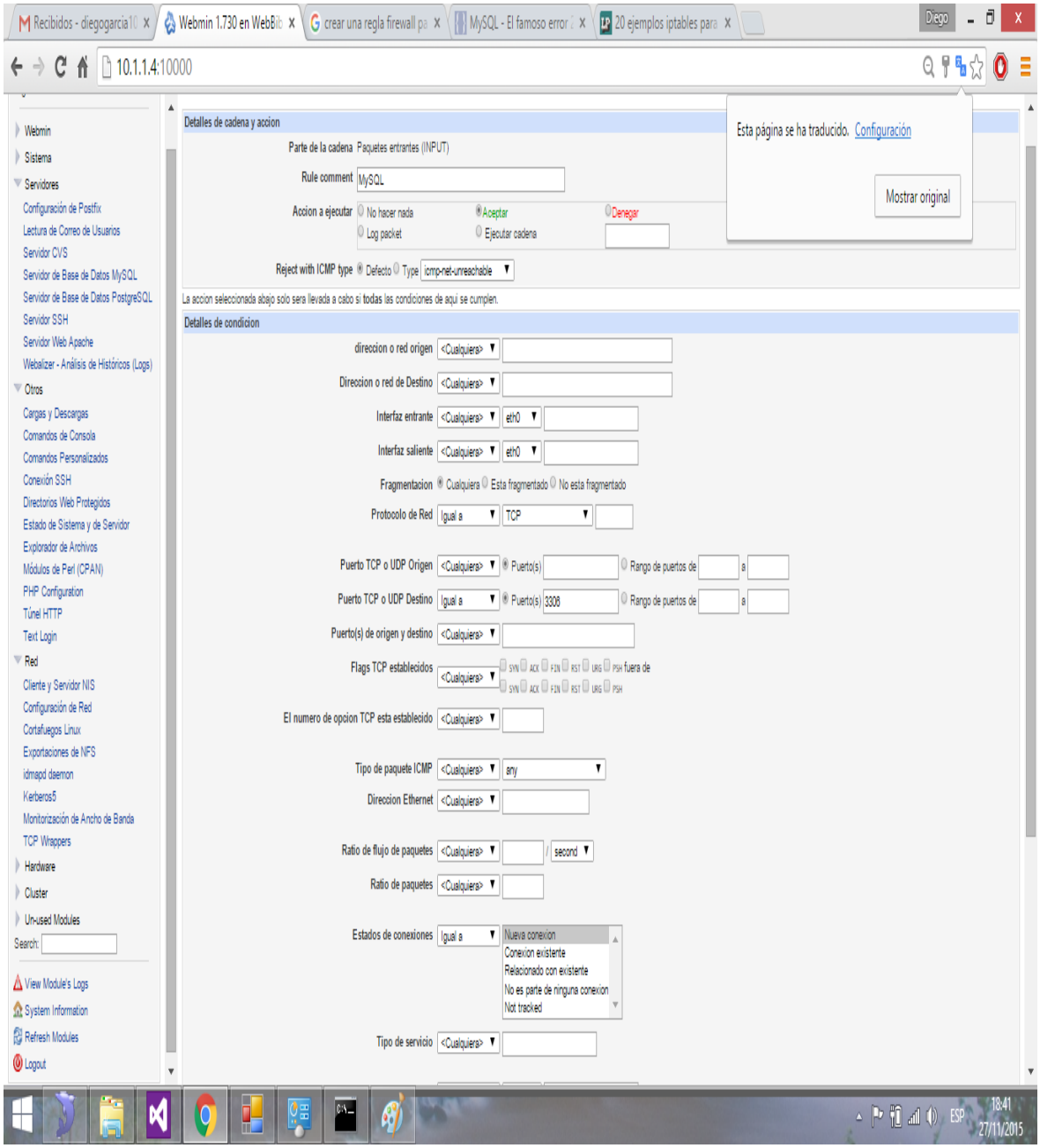
<table>
<tr><td rowspan="2">Título:</td><td rowspan="2">< Breve enunciado sobre lo que trata la historia de usuario ></td><td>Prioridad:</td><td><Valor numérico que se estima después de analizar todas las historias de usuarios></td></tr>
<tr><td>Estimación:</td><td><esfuerzo (horas / programador) calculado></td></tr>
<tr><td>Descripción:</td><td colspan="3">Como < rol de usuario>
quiero …<acción a ejecutar>
para …<finalidad></td></tr>
<tr><td>Criterio de aceptación:</td><td colspan="3"><cuáles requisitos deben cumplirse y a la vez las pruebas que se deben realizar></td></tr>
<tr><td>Prototipo:</td><td colspan="3"><Bosquejo (usando GUIs) de la interface></td></tr>
</table>

Anexo 4. Formato de solicitud de Mantenimiento del Sistema

Solicitud para mantenimiento del sistemas de Carga de Datos del BSC			
Fecha		Teléfono	
Solicitante			
Email		Cargo	
Prioridad	Urgente ☐ Normal ☐ Critica ☐		
Detalle de trabajo a realizar			
Fallo a revisar			

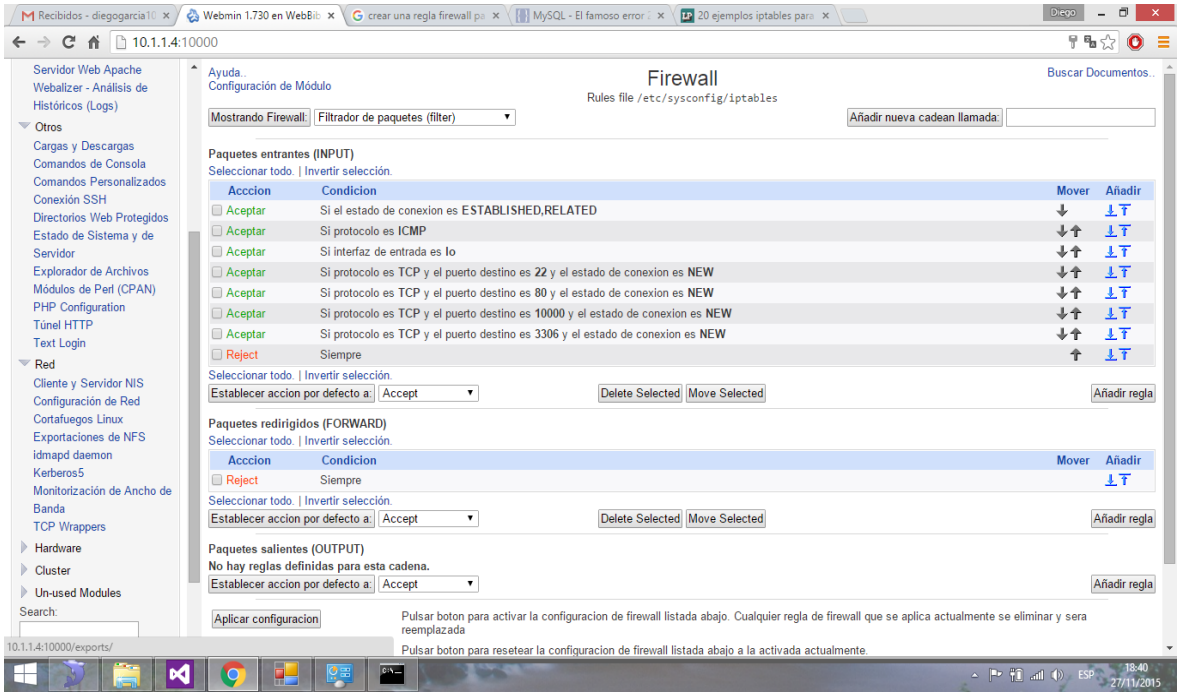
Anexo 5. Prueba realizada de conexión, extracción y carga de datos de la fuente de datos Biblioteca

Paso 1. Se configuro el firewall añadiéndole una regla con el puerto 3306 de MySQL



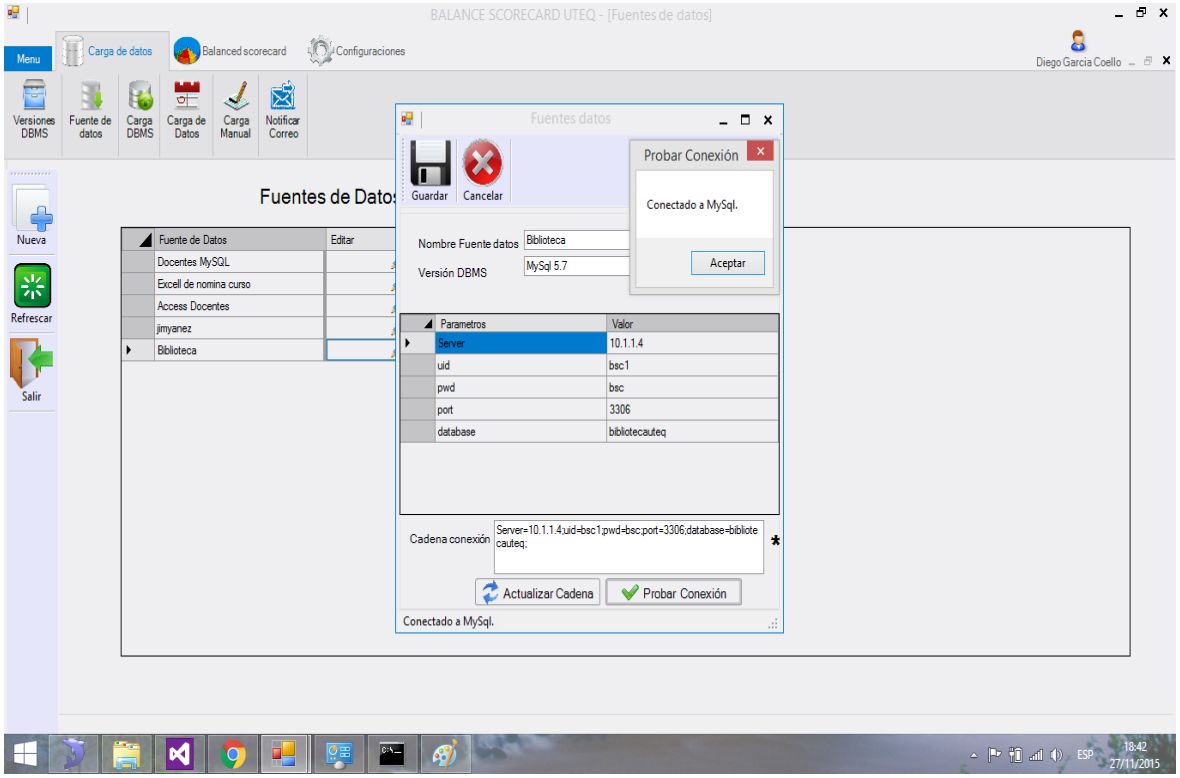
Este proceso se realizó debido a que el servidor donde se encuentra el BDMS MySQL de Biblioteca no tenía una regla que incluya el puerto 3306, lo que bloqueaba las conexiones remotas a la base de datos.

Paso 2 Lista del firewall con la regla incluida



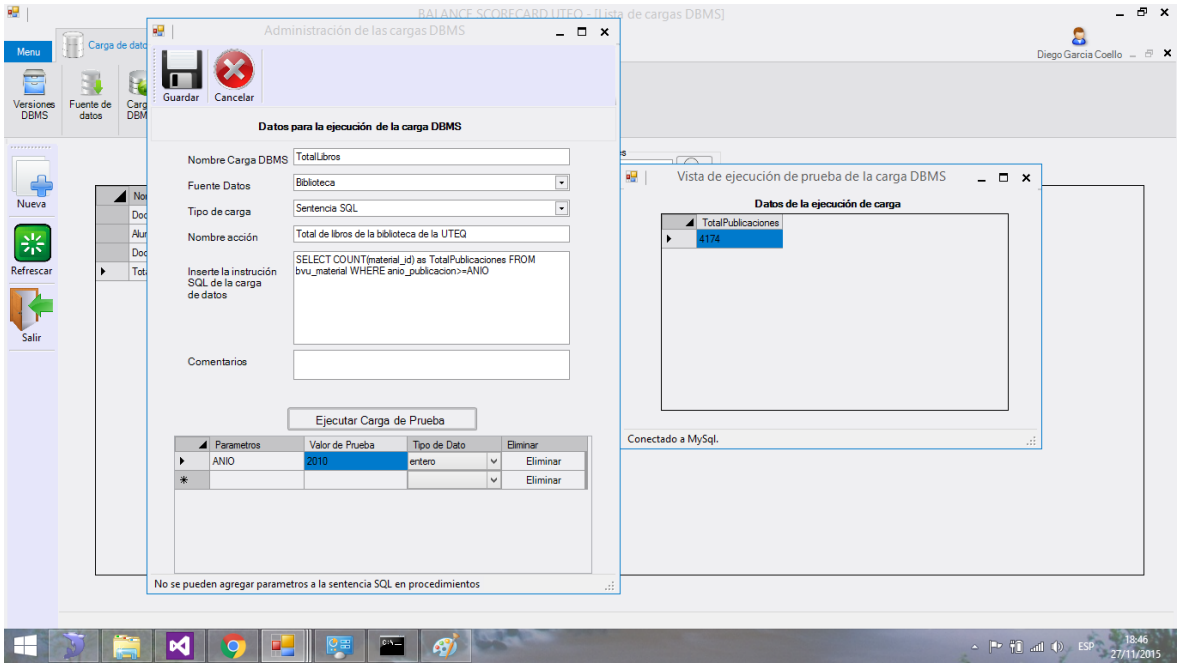
Vista que muestra la regla incluida en las lista de firewall del servidor.

Paso 3. Prueba con los parámetros de conexión a la fuente de datos



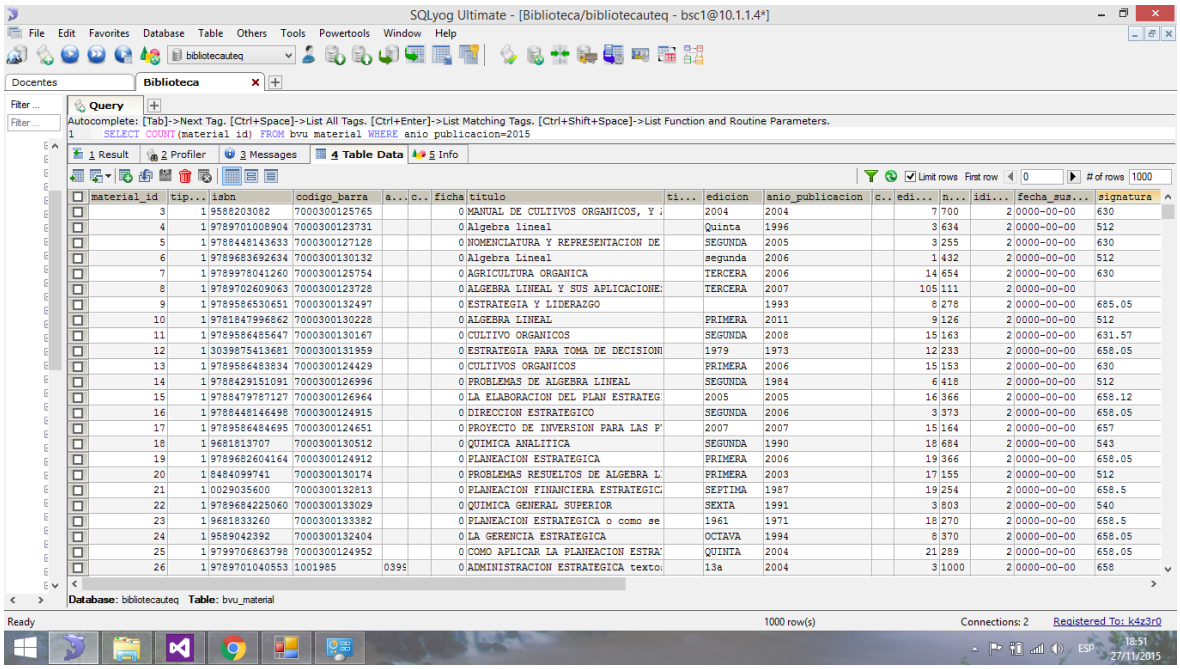
Prueba realizada desde el software que administra la metodología de alimentación de datos donde se añadieron todos los parámetros necesarios para realizar la conexión.

Paso 4. Extracción de datos



Vista donde se muestra una prueba de extracción de datos desde la fuente biblioteca, se debe tomar en cuenta que los datos que se consultan solo deben contener un registro (Una fila y una columna) para guardar la carga.

Paso 5. Tabla de fuente de datos.



Vista que muestra la tabla bvu_material de la fuente de datos biblioteca que contiene los datos que fueron utilizados para realizar las consulta de extracción

Anexo 6. CERTIFICACIÓN DE REDACCIÓN.

Yo, **LCDA. MARJORI TORRES BOLAÑOS** con **C.I. 070182756-0,** Docente de la Facultad de Ciencias de la Ingeniería de la Universidad Técnica Estatal de Quevedo, certifico que he revisado el proyecto de investigación del Egresado **MARCOS DIEGO GARCIA COELLO** con **C.I. 120573633-1** previo a la obtención del título de Ingeniero en Sistemas, titulado **"DESARROLLO DE UNA METODOLOGÍA DE ALIMENTACION DE DATOS PARA EL CUADRO DE MANDO INTEGRAL DE LA UTEQ"**, habiendo cumplido con la redacción y corrección ortográfica que se ha indicado.

Lcda. Marjori Torres Bolaños

C.I. 070182756-0

Anexo 7. CERTIFICADO - ABSTRACT

Yo, **LCDO. MIGUEL MONTALVO ROBALINO** con **C.I. 020049726-1,** Docente de la Facultad de Ciencias de la Ingeniería de la Universidad Técnica Estatal de Quevedo, certifico que he revisado la traducción del idioma inglés del resumen del proyecto de investigación del Egresado **MARCOS DIEGO GARCIA COELLO** con **C.I. 120573633-1** previo a la obtención del título de Ingeniero en Sistemas, titulado **"DESARROLLO DE UNA METODOLOGÍA DE ALIMENTACIÓN DE DATOS PARA EL CUADRO DE MANDO INTEGRAL DE LA UTEQ"**, lo realizo bajo mi supervisión y se encuentra correctamente traducido bajo la estructura del lenguaje inglés.

Lcdo. Miguel Montalvo Robalino

C.I. 020049726-1