

Universidad Internacional de La Rioja (UNIR)

**Escuela Superior de Ingeniería y
Tecnología**

**Máster Universitario en Diseño y Gestión de
Proyectos Tecnológicos**

Diseño de un DataMart
(BI) orientado al área
Administrativa para la
toma de decisiones del
CEC Universidad Técnica
de Machala

Trabajo Fin de Máster

Presentado por: Arreaga Santacruz, Sally Astrid

Director/a: Dr. Infante Moro, Alfonso

Ciudad: Machala, Ecuador

Fecha: 26/02/2020

Resumen

El Business Intelligence (BI) desempeña un rol indispensable para la toma de decisiones con directriz en la calidad, y a su vez resolver los requerimientos de inteligencia corporativa en las organizaciones. Los instrumentos de BI como DataMart son importantes porque gestionan la información que se genera en una entidad, y más aún en el área Administrativa del Centro de Educación Continua (CEC) en la Universidad Técnica de Machala; porque permite a la alta dirección consultar, interpretar y validar los datos procesados dentro de la institución y así obtener conocimiento funcional que permita tomar decisiones sustentadas y de prospectiva consiente. Otra herramienta favorable que deriva del BI es el Data Warehouse (DW) porque integra varias fuentes de datos, principalmente bases de datos transaccionales estructuradas, formatos estructurados, o fuentes no estructuradas para hacer más eficiente el análisis de la información. En congruencia a esta premisa, el TFM tiene como propósito diseñar un sistema DW con DataMart y BI para ofrecer soporte a la alta dirección en la toma de decisiones utilizando datos históricos estructurados o no estructurados. La metodología de inteligencia de negocios utilizada fue la metodología Kimball junto con la metodología para la gestión del proyecto que fue la metodología PMBOK la misma que detalla los lineamientos pertinentes para la administración y desarrollo de un proyecto tecnológico; entre las principales ideas de la metodología destacan la gestión del alcance, tiempo, costos, calidad y riesgos muy necesarios para cumplir satisfactoriamente; en ambas metodologías, la consecución del DataMart y los requerimientos del usuario. Al término del TFM la conclusión más importante obtenida es que un DataMart como estrategias de inteligencia de negocios, pueden mejorar la toma de decisiones porque junto con la base de datos multidimensionales junto con su conexión estrella forman un cubo dimensional de información que muestra consultas directas de datos que hacen que las partes interesadas puedan observar los mejores caminos para llegar a diversas soluciones de un problema, en el caso de CEC, se mitigarán las dificultades que tenían los administradores al momento del registro y admisión de estudiantes, los resultados financieros o el aprovechamiento académico; marcando un impacto positivo en los estudiantes que quieren una educación continua sin tapujos en la administración de la información que procesa la Universidad Técnica de Machala.

Palabras Clave: Alta dirección, Business Intelligence, Data Warehouse, DataMart, toma de decisiones, OLTP, OLAP.

Abstract

Business Intelligence (BI) plays an indispensable role for decision making with quality guidelines, and in turn, meet the requirements of corporate intelligence in organizations. BI instruments such as DataMart are important because they manage the information generated in an entity, and even more so in the Administrative area of the Center for Continuing Education (CEC) at the Technical University of Machala; because it allows senior management to consult, interpret and validate the data processed within the institution and thus obtain functional knowledge that allows for sustained decisions and prospective consent. Another favorable tool that derives from BI is the Data Warehouse (DW) because it integrates several data sources, mainly structured transactional databases, structured formats, or unstructured sources to make information analysis more efficient. Consistent with this premise, the TFM aims to design a DW system with DataMart and BI to offer support to senior management in decision making using structured or unstructured historical data. The business intelligence methodology used was the Kimball methodology along with the project management methodology that was the PMBOK methodology that details the relevant guidelines for the administration and development of a technological project; among the main ideas of the methodology are the management of the scope, time, costs, quality and risks that are very necessary to comply satisfactorily; in both methodologies, the achievement of the DataMart and the user's requirements. At the end of the TFM, the most important conclusion obtained is that a DataMart, such as business intelligence strategies, can improve decision making because together with the multidimensional database together with its star connection they form a dimensional cube of information that shows direct queries of data that make it possible for interested parties to observe the best ways to reach various solutions to a problem, in the case of CEC, the difficulties faced by administrators at the time of student registration and admission, financial results or achievement will be mitigated academic; marking a positive impact on the students who want a continuous education without blunt in the administration of the information processed by the Technical University of Machala.

Keywords: Senior management, Business Intelligence, Data Warehouse, DataMart, decision making, OLTP, OLAP.

INDICE DE CONTENIDO

CONTENIDOS	PÁGINAS
Resumen	2
Abstract	3
CAPÍTULO 1. INTRODUCCIÓN	14
1.1.- Motivación	14
1.2.- Planteamiento del trabajo	15
1.3.- Estructura del trabajo	18
CAPÍTULO 2. CONTEXTO Y ESTADO DEL ARTE	19
2.1.- Análisis del contexto	19
2.2.- Estado del arte	19
2.2.1.- Business Intelligence	19
2.2.1.1.- Importancia del BI	20
2.2.1.2.- Componentes del BI	21
2.2.2.- Data Warehouse	22
2.2.2.1.- El modelo de datos del Data Warehouse	24
2.2.2.2.- Técnicas de modelado Data Warehouse	25
2.2.2.3.- Modelado de diseño de bases de datos Data Warehouse	25
2.2.2.4.- Desarrollar un Data Warehouse	26
2.2.3.- Data Mart	27
2.2.3.1.- Clasificación de los Data Mart	28
2.2.3.2.- Carga de datos en Data Mart	30
2.2.4.- Consulta y herramientas de reporte	30
2.2.5.- Sistemas tradicionales de inteligencia empresarial	30
2.2.6.- OLAP versus OLTP	32

2.2.7.- Enfoques de diseño del Data Warehouse	33
2.2.7.1.- Arquitectura Bill Inmon	33
2.2.7.2.- Arquitectura Ralph Kimball	35
2.2.7.3.- Metodología de Kimball	37
2.2.7.4.- Modelo multidimensional	40
2.2.7.5.- Diferencias entre los enfoques de Inmon y Kimball	43
2.2.7.6.- Factores decisivos	43
2.2.8.- Hipótesis	44
CAPÍTULO 3. OBJETIVOS Y METODOLOGÍA DEL TRABAJO	45
3.1.- Objetivos	45
3.1.1.- Objetivo general	45
3.1.2.- Objetivos específicos	45
3.2.- Descripción del proyecto	45
3.3.- Metodología de trabajo	47
3.3.1.- Metodología de gestión del proyecto	47
3.3.2.- Metodología de la solución de BI	51
CAPÍTULO 4. DESARROLLO ESPECÍFICO DE LA CONTRIBUCIÓN	53
4.1.- Identificación de interesados	53
4.1.1.- Interesados externos	53
4.1.2.- Interesados internos	54
4.1.3.- Interesados de desarrollo	54
4.2.- Definición de requerimientos	58
4.3.- Definición del alcance del proyecto	61
4.3.1.- Alcance de la solución	61
4.3.2.- Exclusiones	63

4.3.3.- Supuestos	64
4.3.4.- Restricciones	65
4.4.- Plan de trabajo	65
4.4.1.- Paquete de trabajo 1: Gestión del proyecto	65
4.4.2.- Paquete de trabajo 2: Análisis y especificación de requisitos	82
4.4.3.- Paquete de trabajo 3: Diseño y desarrollo de la arquitectura	93
4.4.4.- Paquete de trabajo 4: Desarrollo de un prototipo demostrador	94
4.4.4.1.- Datos del CEC	94
4.4.4.2.- Requisitos del sistema	94
4.4.4.3.- Diseño del sistema	95
4.4.5.- Paquete de trabajo 5: Divulgación y explotación	116
4.4.6.- Cronograma	118
4.5.- Estructura de gestión	120
4.6.- Plan de calidad	121
4.6.1.- Comunicación y colaboración	124
4.6.2.- Aseguramiento de la calidad	125
4.7.- Plan de gestión del riesgo	129
4.8.- Estrategia de divulgación	138
4.9.- Estrategia de explotación de resultados	139
4.9.1.- Plan de internacionalización de los resultados	140
4.9.2.- Plan de gestión de la propiedad intelectual	141
4.10.- Presupuesto	142
4.11.- Impacto socioeconómico y medio ambiental	148
4.11.1.- Creación de empleo	151
4.11.2.- Medidas orientadas a la igualdad de género	152

4.11.3.- Medidas orientadas a la inclusión social	153
4.11.4.- Medidas orientadas a la mejora de la sostenibilidad	154
CAPÍTULO 5. CONCLUSIONES Y TRABAJO FUTURO	157
5.1.- Principales conclusiones	157
5.2.- Líneas de trabajo futuro	159
REFERENCIAS BIBLIOGRÁFICAS	161
BIBLIOGRAFÍA	163
ANEXOS	165

INDICE DE TABLAS

TABLA	PÁGINAS
Tabla 1: Diferencias entre la base de datos OLTP y la base de datos OLAP	33
Tabla 2: Cuadro comparativo de las metodologías de BI	51
Tabla 3: Personas interesadas del diseño del DataMart (BI)	55
Tabla 4: Funciones system owner	56
Tabla 5: Funciones Administrator DataMart	56
Tabla 6: Funciones Developer	57
Tabla 7: Tipos y roles de Usuario	58
Tabla 8: Dimensión de años académicos	98
Tabla 9: Dimensión de los semestres académicos	99
Tabla 10: Dimensión de la fecha	99
Tabla 11: Dimensión de los programas de estudio	101
Tabla 12: Dimensión colegiatura	102
Tabla 13: Dimensión de departamentos	102
Tabla 14: Dimensión de cursos	102
Tabla 15: Dimensión de los años de colegiatura	103
Tabla 16: Dimensión de los grados de colegiatura	103
Tabla 17: Dimensión de las secciones de colegiatura	104
Tabla 18: Dimensión de estados estudiantiles	104
Tabla 19: Dimensión de niveles de los estudiantes	105
Tabla 20: Dimensión de ubicaciones geográficas	105
Tabla 21: Dimensión de los estudiantes	106
Tabla 22: Dimensión de las becas	107

Tabla 23: Dimensión de horas de registro de estudiantes	107
Tabla 24: Divulgación de la solución	116
Tabla 25: Explotación de la solución	117
Tabla 26: Cronograma de actividades	118
Tabla 27: Estructura de gestión del proyecto	120
Tabla 28: Comunicación y colaboración en el proyecto	124
Tabla 29: Planificación del aseguramiento de la calidad	126
Tabla 30: Estrategia de divulgación	138
Tabla 31: Estrategia de explotación de resultados	139
Tabla 32: Internacionalización de los resultados	140
Tabla 33: Objetivos estratégicos de gestión de la propiedad intelectual	141
Tabla 34: Recurso Humano	143
Tabla 35: Caracterización del Hardware para el desarrollo	143
Tabla 36: Hardware adicional para el desarrollo	144
Tabla 37: Costo total del hardware para el desarrollo	144
Tabla 38: Caracterización del Hardware para la implantación	144
Tabla 39: Hardware adicional para la implantación	145
Tabla 40: Costo total del hardware para la implantación	145
Tabla 41: Caracterización del Software para el desarrollo	145
Tabla 42: Caracterización del Software para la implantación	146
Tabla 43: Capacitación del Recurso Humano	146
Tabla 44: Costo del diseño del DataMart	147
Tabla 45: Costo total del diseño del DataMart	147
Tabla 46: Rango de niveles de impacto positivo y negativo	148

Tabla 47: Métricas del impacto socioeconómico	148
Tabla 48: Métricas del impacto medioambiental	149
Tabla 49: Métricas del impacto en forma general	150

INDICE DE FIGURAS

FIGURA	PÁGINAS
Figura 1: Componentes principales de la BI	22
Figura 2: Modelo de ciclo de vida de desarrollo de Data Warehouse (DWLC)	26
Figura 3: DataMart Independiente	29
Figura 4: DataMart Dependiente	29
Figura 5: Arquitectura tradicional de un sistema BI	31
Figura 6: Enfoque de Bill Inmon para el Data Warehousing	34
Figura 7: Enfoque de Ralph Kimball para el Data Warehousing	35
Figura 8: Ciclo de vida de la metodología Ralph Kimball	38
Figura 9: Modelo estrella	41
Figura 10: Modelo copo de nieve	42
Figura 11: Modelo constelación	42
Figura 12: Arquitectura de la metodología PMBOK	48
Figura 13: Arquitectura de la metodología SCRUM	49
Figura 14: Personas interesadas del DataMart (BI)	55
Figura 15: Requerimientos para el diseño de un DataMart	59
Figura 16: Pirámide de datos que detallan la granularidad	61
Figura 17: Ejemplo de modelo dimensional	62
Figura 18: Triple restricción	66
Figura 19: Arquitectura de la solución	93
Figura 20: Modelo en espiral del ciclo de vida del almacén de datos	96
Figura 21: DataMart resultados de colegiatura	108
Figura 22: DataMart admisión de estudiantes	109

Figura 23: Vistas de admisión de estudiantes	110
Figura 24: DataMart de registro de estudiantes	111
Figura 25: Vistas de registro de estudiantes	112
Figura 26: DataMart ubicaciones de los estudiantes	112
Figura 27: Vistas de ubicaciones de estudiantes	113
Figura 28: DataMart promedio del estudiante	114
Figura 29: Vistas de promedios de colegiatura estudiantes	115

INDICE DE ANEXOS

ANEXO	PÁGINAS
ANEXO I: PRUEBAS DE ACEPTACIÓN	165
ANEXO II: CIERRE DEL PROYECTO Y PLANTILLAS	167

CAPÍTULO 1. INTRODUCCIÓN

1.1.- Motivación

La motivación del diseño de un DataMart para mejorar la toma de decisiones en el área Administrativa del Centro de Educación Continua (CEC) de la Universidad Técnica de Machala concuerda con los lineamientos de la plataforma de Business Intelligence (BI) porque ayudará a disponer de información coherente en reportes e indicadores del rendimiento de la entidad con respecto a métricas de control interno como la educación de calidad, la investigación, el clima organizacional, la inteligencia corporativa y las funciones que el CEC cumple con respecto a la profesionalización y formación superior de la sociedad. Con la herramienta DataMart la generación de reportes será automática, eliminando el error humano en la compilación, procesamiento y ejecución de la información.

El desarrollo de un Data Warehouse que integra el DataMart permitirá obtener una base de conocimientos que apoyarán a la toma de decisiones de la Alta Dirección del CEC para que puedan determinar de forma fundamentada, si se está cumpliendo con los objetivos estratégicos, misión y visión del centro y en directriz de una educación de calidad. En este sentido, el TFM trata de inducir a la Universidad en el paradigma de la automatización de los procesos de consultas e indicadores y la estructuración lógica del procesamiento de la información para lograr indicios de la inteligencia de negocios (BI) y tener un aprovechamiento eficaz de la tecnología de la información y las estrategias tecnológicas de datos, las cuales proporcionan mayor rapidez en el flujo de reportes, la accesibilidad y el soporte en la toma de decisiones.

El diseño del DataMart facilita al personal administrativo del CEC de la Universidad de Machala disponer de un conocimiento óptimo de identificación de la situación actual del CEC y la prospectiva del mismo, porque la información se manipulará de forma precisa y entendible de tal forma que se puedan tomar decisiones oportunas, con conciencia de mejora continua. Un pilar fundamental del diseño de un DataMart es la seguridad que se procesa en la información, un punto muy importante para entidades que manipulan información tan delicada como son las universidades, gestoras del desarrollo profesional de las sociedades.

Las herramientas de BI toman gran importancia porque pueden convertir los datos contenidos en las bases de datos transaccionales del sistema informático del CEC, en información transformada en estructuras gráficas y organizativas con conocimiento útil en el proceso de

toma de decisiones de la alta dirección del CEC y lógicamente de la Universidad Técnica de Machala. El diseño del Data Warehouse es importante para el desarrollo de la entidad de educación superior porque ayuda a la toma de decisiones con conciencia prospectiva por medio de la formulación de preguntas, consultas, análisis e interpretación de datos instantáneos o en tiempo real, sin necesidad de tener arduos conocimientos en programación o informática avanzada.

El diseño del DataMart como estrategia de BI se logra dos beneficios reales como es la integración y acceso de la información por medio de datos, depurando gran cantidad de información poco relevante para la alta dirección; y la generación de un compilador eficiente de la información útil para la toma de decisiones. En este sentido, económicamente se eliminarán los costos de la elaboración de reportes manuales; tecnológicamente, el DataMart permitirá procesar y presentar toda la información necesaria para atender los requerimientos de gestión del conocimiento del CEC.

En síntesis, se puede deducir que el diseño del DataMart proporcionará a la alta dirección del CEC un instrumento tecnológico de carga y descarga de datos de forma intuitiva, para luego reportar los datos históricos en el menor tiempo posible, en un formato entendible por el usuario y cumpliendo con las métricas de BI, usabilidad e interoperabilidad en distintos dispositivos y plataformas. Esta premisa, satisface las necesidades de la sociedad con respecto a la calidad en la educación porque mantiene la vanguardia tecnológica que disponen las instituciones de educación superior posicionadas en niveles prioritarios del ranking mundial. Desde otra perspectiva, el diseño del DataMart también cumple con el medio ambiente y su cuidado, porque contribuye en la disminución del uso de papel impreso debido a que la información se procesa y gestiona de forma digital.

1.2.- Planteamiento del trabajo

Las universidades son organizaciones con una importante responsabilidad social, en ellas se genera y transmite gran parte del conocimiento que apoya el desarrollo económico de las sociedades. La trascendencia del encargo social de las universidades y el alto costo de la enseñanza, sobre todo en el campo tecnológico, demandan eficiencia, eficacia y calidad en los procesos que en éstas se desarrollan. Para conseguir este propósito, la gestión de los procesos debe ser efectiva, basada ampliamente en el uso de las tecnologías y con métodos sujetos a constante perfeccionamiento. Con el objetivo de mejorar la gestión y el procesamiento del conocimiento que se genera y acumula sobre sus actividades y

desempeño. Es con este fin que la gestión de la información y el conocimiento devienen de herramientas importantes como los DataMart y los Data Warehouse para dar soporte a la toma de decisiones en las universidades.

El proceso de formación y profesionalización de la sociedad, principal proceso en las universidades, es un proceso complejo, debido a la gran variedad en las características y necesidades de educación de los estudiantes, que son el objeto a transformar, y a la variedad de condiciones que confluyen en el proceso para consumir esta transformación; los gestores de este proceso, denominados como administradores del conocimiento, entendidos a su vez, como aquellos que usan la información como una principal directriz de transformación a través de su conocimiento para alinear la toma de decisiones y planificar las estrategias para gestar un pertinente y consolidado procesamiento de este conocimiento que debe transfigurar la vida de los estudiantes.

Para que educadores y directivos sean más productivos, deben un amplio conocimiento, no sólo de las asignaturas que enseñan, sino también de las condiciones bajo las cuales se desarrolla el proceso de educación, para poder elaborar decisiones acertadas acerca de la metodología y pedagogía a utilizar para alcanzar los diferentes resultados en honor a la calidad educativa. El dominio cognitivo que se desea de los directivos y educadores de una institución de educación superior, requiere una correcta gestión de la información y el conocimiento, que facilite la toma de decisiones así como la concepción de estrategias eficaces. En este sentido, la inteligencia de negocios (BI) es una invención que permite diseñar, desarrollar e implantar una herramienta tecnológica que ayuda al análisis de información y da soporte en la toma de decisiones de los administradores y educadores de estas entidades de la educación superior.

En respuesta a la exigencia de BI y los cambios continuos de los ambientes universitarios, que han enfrentado el reto de tomar decisiones basándose en información que generalmente se encuentra dispersa y en diferentes formatos, e integrarla; representa mucho esfuerzo y tiempo, así como la gestión de recursos. La BI permite integrar la información de todas las dependencias del área administrativa, para que desde una interfaz sencilla, se pueda entender cómo está funcionando para detectar oportunidades de mejora continua. Analizando indicadores clave y configurando herramientas para profundizar en el suficiente detalle de la perspectiva que necesita la educación superior con la toma de decisiones oportunas y acertadas. Estas herramientas tecnológicas y de configuración que regula la BI son las DataMart en alineación con las Data Warehouse, cuyo campo es relativamente nuevo en el

ámbito de la información, la tecnología; y como resultado, no hay una gran cantidad de documentación analítica de los procesos de Data Warehouse.

Los avances en el dominio de la tecnología de la información han guiado la evolución de los sistemas de procesamiento de datos desde las primeras etapas con un soporte único y aplicaciones independientes hasta los sistemas analíticos y de BI integrales para el entorno informativo actual. Dentro de este amplio contexto, el Data Warehouse define una amplia combinación de tecnologías surgidas a principios de la década de 1990 como resultado de los avances en el área de las capacidades de procesamiento de datos logradas en los sistemas de información basados en computadora.

El Data Warehouse representa un componente del marco general de BI, que cubre una amplia gama de aplicaciones y herramientas que ayudan a analizar grandes cantidades de datos y transformarlos en información comprensible y conocimiento organizacional. Está diseñado para manejar un entorno informativo en el que una serie de componentes permiten la recopilación e integración de datos de toda la organización, para el caso el área Administrativa del Centro de Educación Continua (CEC), proporcionando a estos usuarios datos estructurados y consolidados, con los que puedan mejorar la toma de decisiones.

Un DataMart configurado en un Data Warehouse representa un modelo de datos universitarios, especialmente estructurado para facilitar los procesos de consulta y análisis de datos integrados y consolidados. Es un componente esencial y dominante de los sistemas de soporte de decisiones basados en datos y su objetivo principal es permitir a los usuarios de las instituciones de educación superior tomar decisiones tácticas y estratégicas efectivas basadas en datos fehacientes. Para lograr su objetivo, el Data Warehouse se define mediante modelos de datos particulares que especifican la estructura de datos en el DataMart. Estos modelos de datos, optimizados para consultas y análisis, se crean de manera estable, consistente y predecible a través de diferentes técnicas de modelado de datos como son las bases de datos.

Teniendo en cuenta su nivel de complejidad, el desarrollo de soluciones de Data Warehouse exige un enfoque estructurado y planificado, definido en forma de metodología, así como un marco arquitectónico apropiado. Las metodologías están diseñadas para lograr resultados predecibles de acuerdo con requisitos bien definidos y para proporcionar procesos de desarrollo repetibles, heurísticos y consistentes. Las arquitecturas representan las estructuras que unen todos los componentes del Data Warehouse y proporcionan una base sólida para la integración de datos. La selección de una metodología y arquitectura adecuada determinan

el éxito general de la implementación de la solución de Data Warehouse y DataMart. Otro aspecto esencial en el desarrollo del Data Warehouse se refiere al empleo de un marco capaz de proporcionar un conjunto de pautas para describir sus componentes y su interoperabilidad y para respaldar la existencia de un entorno de procesos reutilizables, integración, consistencia y flexibilidad en la entrega de información.

1.3.- Estructura del trabajo

Los temas que se presentan a partir del siguiente capítulo establecen:

- **Capítulo 2:** Se realiza el estudio del contexto y el estado del arte; es decir, se fundamenta científicamente el diseño de un DataMart, sus componentes, metodología, características y especificaciones de desarrollo. Además, se analiza y compara la diferente información que implica la inteligencia de negocios para la toma de decisiones.
- **Capítulo 3:** En este capítulo se realiza un análisis de los objetivos y la metodología del trabajo. En primera instancia se postula el objetivo general que guiará el diseño del DataMart para ser complementado con la proposición de los objetivos específicos. Además, se estipula la descripción del proyecto y la metodología de trabajo, que para el caso será la metodología de gestión de proyectos PMBOK (Project Management Body of Knowledge) que define desde el ciclo de vida del proyecto a las estrategias y conceptos de gestión que se pueden aplicar al proyecto.
- **Capítulo 4:** A partir del contexto y el estado del arte del diseño del DataMart del capítulo primero entrelazado a los objetivos y la metodología de trabajo del tercer capítulo, se describe el desarrollo específico de la contribución, considerando la identificación de interesados, la definición de requisitos, el alcance del proyecto, el plan de trabajo, la estructura de gestión, el plan de calidad, el plan de gestión de riesgo, la estrategia de divulgación, la estrategia de explotación de resultados, el presupuesto y el impacto socioeconómico y medio ambiental.

Finalmente, se sugiere el diseño del DataMart como modelo conceptual de sistema de BI, para que a partir de este, se adapte a las necesidades particulares del CEC de la Universidad Técnica de Machala.

- **Capítulo 5:** Se finaliza el TFM con las principales conclusiones y las respectivas líneas de trabajo futuro.

CAPÍTULO 2. CONTEXTO Y ESTADO DEL ARTE

2.1.- Análisis del contexto

Antes de profundizar en la metodología aplicada en el TFM, se detalla el estado del arte, en relación a distintos conceptos que se tratan en el documento. En este capítulo una revisión bibliográfica sobre la inteligencia de negocios, el contexto del DataMart y Data Warehouse, así como los indicadores de la gestión en general y en la gestión de proyectos en particular.

En primer lugar se señala los fundamentos de la Inteligencia de Negocios o Business Intelligence (BI) por sus siglas en inglés, como disciplina de la gestión de proyectos tecnológicos y de la Gestión del Conocimiento organizacional como son los DataMart y el Data Warehouse, estableciendo paradigmas concretos producidos para el diseño del DataMart.

Una vez introducida la disciplina del DataMart en el que se centra el TFM, se procede a revisar las especificaciones y planificación, así como los indicadores de la gestión de proyectos tecnológicos para el diseño del DataMart en el CEC de la Universidad Técnica de Machala.

2.2.- Estado del arte

Para poder plantear conceptualmente un sistema de BI que fortalezca el control de gestión y la toma de decisiones dentro del CEC de la Universidad Técnica de Machala, es necesario establecer estructuras y modelos de gestión para el diseño del DataMart que mejore las condiciones administrativas del Centro de Educación Continua. En los siguientes epígrafes se realiza una descripción de la inteligencia de negocios y la información pertinente al diseño del DataMart, sin el afán de llegar a particularidades técnicas, se presentan conceptos para entender las características de este tipo de sistemas.

2.2.1.- Business Intelligence

En los últimos años, los sistemas de soporte de decisiones llamados inteligencia empresarial (BI) se ha convertido en una parte integral de la estrategia de toma de decisiones de las organizaciones. El BI es el proceso de convertir los datos en información y luego en conocimiento (Golfarelliet, 2004). Actualmente, las organizaciones compiten en el mercado global. Para que una empresa obtenga una ventaja competitiva sobre las demás y también para ayudar a tomar mejores decisiones, BI ahora está desempeñando un papel importante

en la toma de decisiones estratégicas. Ayuda a las empresas a tomar mejores decisiones, racionalizar los flujos de trabajo y proporcionar mejores servicios al cliente.

La BI es una colección de tecnologías de soporte de decisiones y conocimiento para la empresa, cuyo objetivo es permitir que los trabajadores del conocimiento, como ejecutivos, gerentes y analistas, tomen decisiones mejores y más rápidas. Las últimas dos décadas han visto un crecimiento explosivo, tanto en la cantidad de productos y servicios ofrecidos como en la adopción de estas tecnologías por la industria. Este crecimiento se ha visto impulsado por el costo decreciente de adquirir y almacenar grandes cantidades de datos provenientes de fuentes tales como transacciones de clientes en la banca, el comercio minorista y también en los negocios electrónicos, etiquetas RFID para el seguimiento de inventario, correo electrónico y registros de consultas para la Web, sitios electrónicos, blogs y catálogos digitales de productos. Las organizaciones de hoy recogen datos en una granularidad más fina, que por lo tanto es de un volumen mucho mayor.

Frente a lo argumentado sobre la BI, Adelman (2002) describe la BI como un término que abarca una amplia gama de software analítico y soluciones para recopilar, consolidar, analizar y proporcionar acceso a la información de una manera que se supone que permite a los usuarios de una organización tomar mejores decisiones comerciales. Por su parte, Malhotra (2000) señala a los beneficios de la BI como facilitadores de conexiones en la organización, proporcionando información en tiempo real a repositorios centralizados; además, facilitando su análisis de soporte que pueden explotarse en todos los niveles horizontales y verticales dentro y fuera de la empresa. En congruencia, Golfarelli (2004) resume sobre la BI que incluye un Data Warehouse efectivo y también un componente reactivo capaz de monitorear los procesos operativos críticos de tiempo para permitir a los tomadores de decisiones tácticos y operativos ajustar sus acciones de acuerdo con la estrategia de la organización. En síntesis a lo estipulado por los autores, Gangadharan & Swamy (2004) definen la BI como el resultado del análisis en profundidad de datos comerciales detallados, incluidas las tecnologías de bases de datos y aplicaciones, así como las prácticas de análisis; también, incluyen potencialmente la gestión del conocimiento, ERP, sistemas de soporte de decisiones y minería de datos.

2.2.1.1.- Importancia del BI

Las organizaciones están aprovechando su activo de datos de manera agresiva mediante la implementación y la experimentación con técnicas de análisis de datos más sofisticadas para impulsar las decisiones comerciales y ofrecer nuevas funcionalidades, como ofertas y

servicios personalizados a los clientes. Hoy en día, es difícil encontrar una empresa exitosa que no haya aprovechado la tecnología de BI para su negocio. Por ejemplo, la tecnología BI se utiliza en la fabricación para el envío de pedidos y la atención al cliente, en el comercio minorista para el perfil de los usuarios, servicios para análisis de reclamos y detección de fraude, en transporte para la gestión de flotas, en telecomunicaciones para identificar razones de pérdida de clientes, en servicios públicos para análisis de uso de energía, en la educación para mejorar el rendimiento académico y estructurar las estrategias pedagógicas en base a análisis neuroeducativos, o en la atención médica para análisis de resultados.

Los estudios de la industria han puesto de relieve este importante desarrollo. Por ejemplo, basado en una encuesta realizada a más de 4.000 profesionales de tecnología de la información (TI) de 93 países y 25 industrias, el Informe de tendencias tecnológicas de IBM (2018) identificó el análisis empresarial como una de las cuatro principales tendencias tecnológicas en la década de 2017. En una encuesta sobre el estado de la analítica empresarial realizada por Bloomberg Businessweek (2018), se descubrió que el 97 por ciento de las empresas con ingresos superiores a \$ 100 millones usaban alguna forma de analítica empresarial. Un informe del Instituto McKinsey Global predijo que para 2018, solo Estados Unidos enfrentará una escasez de 140.000 a 190.000 personas con profundas habilidades analíticas, así como una escasez de 1.5 millones de gerentes conocedores de datos con los conocimientos para analizar BIG DATA para tomar decisiones efectivas (Manyika, 2011).

2.2.1.2.- Componentes del BI

La BI incluye varios software para extracción, transformación y carga (ETL por sus siglas en inglés), data warehousing, consulta de bases de datos e informes, análisis de datos de procesamiento analítico multidimensional en línea (OLAP), minería de datos y visualización de información (Berson, 2002).

Los expertos observan la BI de diferentes maneras. Los expertos en data warehousing describen la BI como una plataforma tecnológica para la aplicación de soporte de decisiones. Para los expertos en minería de datos, la BI es un conjunto de sistemas avanzados de soporte de decisiones con técnicas de minería de datos y aplicaciones de algoritmos. Para los estadísticos, la BI se considera una herramienta de análisis multidimensional y de pronóstico que incursiona en el paradigma prospectivo de las organizaciones.

Las herramientas de BI son ampliamente aceptadas como un nuevo middleware entre las aplicaciones transaccionales y las aplicaciones de soporte de decisiones, desacoplando así

los sistemas diseñados para un manejo eficiente de las transacciones comerciales de los sistemas diseñados para un soporte eficiente de las decisiones comerciales. Las capacidades de BI incluyen soporte de decisiones, OLAP, análisis estadístico, pronóstico y minería de datos.

La figura 1 muestra los componentes principales que constituyen la BI (Medina, 2018, p.18), el diseño está caracterizado por el uso de fuentes de información, ETL, Data Warehouse, Servidor OLAP y herramientas Front-end para mostrar resultados a los usuarios y personas interesadas.

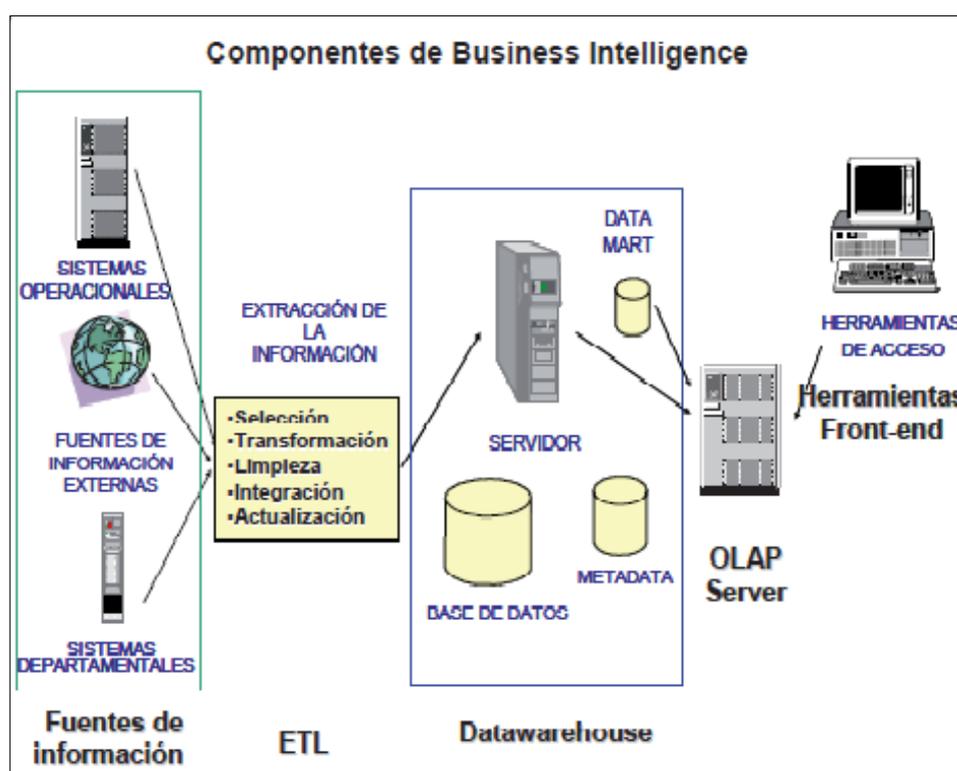


Figura 1. Componentes principales de la BI. Fuente: (Medina, 2018, p.18).

2.2.2.- Data Warehouse

El Data Warehouse es el componente importante de BI. El Data Warehouse admite la propagación física de datos al manejar los numerosos registros empresariales para tareas de integración, depuración, agregación y consulta. También puede contener los datos operativos que se pueden definir como un conjunto actualizable de datos integrados utilizados para la toma de decisiones tácticas en toda la organización o de un área temática particular. Contiene datos en vivo, no instantáneas, y conserva un historial mínimo.

El almacenamiento de datos o Data Warehouse es el proceso de recopilación de datos para ser almacenados en una base de datos administrada en la que los datos están orientados al tema e integrados, varían en el tiempo y no son volátiles para el soporte de la toma de decisiones (Inmon, 1993). Los datos de las diferentes operaciones de una corporación se concilian y almacenan en un depósito central (un Data Warehouse) de donde los analistas extraen información que permite una mejor toma de decisiones (Cho y Ngai, 2003).

Los datos pueden luego agregarse o analizarse, y desplegarse en segmentos según sea necesario para proporcionar información (Fox, 2004). Hay dos autores principales que son conocidos en el mundo del diseño de Data Warehouse, sus enfoques son diferentes; William Inmon y Ralph Kimball. El enfoque de Inmon es el diseño de arriba hacia abajo, mientras que el de Kimball es el diseño de abajo hacia arriba. La mayoría de los profesionales del Data Warehouse se suscriben a cualquiera de los dos enfoques.

Según Inmon (1993), un Data Warehouse es una recopilación de datos no volátil, orientada por temas, integrada, variante en el tiempo y utilizada para apoyar los procesos de toma de decisiones. “Orientada al sujeto” significa que un Data Warehouse se enfoca en las entidades de alto nivel de la empresa y los datos se organizan de acuerdo al sujeto (Zeng, 2003). “Integrada” significa que los datos se almacenan en formatos consistentes, convenciones de nomenclatura, en la medición de variables, estructuras de codificación, atributos físicos de datos o restricciones de dominio. Por ejemplo, mientras que una organización puede tener cuatro o cinco esquemas de codificación únicos para el origen étnico, en un Data Warehouse solo hay un esquema de codificación (Chan, 1999). “Variante de tiempo” significa que el almacenamiento proporciona acceso a un mayor volumen de información más detallada durante un período más largo (Zeng, 2003) y que los datos están asociados con un punto en el tiempo (Chan, 1999) como mes, trimestre o año. Los datos del Data Warehouse no son volátiles, ya que los datos que ingresan a la base de datos rara vez se modifican una vez que se ingresan al almacenamiento. Los datos en el almacenamiento son de solo lectura; las actualizaciones de los datos ocurren de manera periódica, incremental o completa. Finalmente, “no volátil” significa que los datos no cambian (Chan, 1999).

Por su parte, Kimball (2002), asegura que el Data Warehouse es el conglomerado de todos los mercados de datos dentro de la organización. La información siempre se almacena en el modelo dimensional. Según el autor, el Data Warehousing es un componente de los DataMarts. Los DataMarts se centran en entregar objetivos comerciales para los departamentos de la organización. Y el Data Warehouse es una dimensión conformada de los DataMarts. Para Kimball (2002) un DataMart es un subconjunto del Data Warehouse. El Data

Warehouse es la suma de todos los DataMarts, cada uno de los cuales representa un proceso de negocio en la organización por medio de un esquema en estrella, o una familia de esquemas en estrella de diferente granularidad.

La principal diferencia entre el enfoque de Kimball (2002) y la de Inmon (1993) es que las dimensiones conformadas de Kimball están des-normalizadas, mientras que Inmon utiliza un modelo de base de datos central altamente normalizado. Los DataMarts de Inmon almacenan una segunda copia de los datos de las tablas del Data Warehouse centralizado, mientras que las dimensiones de Kimball utilizadas en los DataMarts no son copias de las dimensiones conformadas, sino la tabla de dimensiones en sí. Kimball (1996) se refieren al conjunto de dimensiones conformadas como el bus del Data Warehouse.

Existe relatividad entre estas dos ideas, ya que representan diferentes filosofías de Data Warehousing. En realidad, el Data Warehouse en la mayoría de las empresas está más cerca de la idea de Ralph Kimball. Esto se debe a que la mayoría de los Data Warehouses comenzaron como un esfuerzo departamental y, por lo tanto, se originaron como un DataMart. Solo cuando se construyen más DataMart en lo posterior, se convierten en un Data Warehouse.

Según Basaran (2005) detalla algunas de las características del Data Warehouse:

- Está orientado a temas.
- No es volátil.
- Permite la integración de varios sistemas de aplicación.
- Es compatible con el procesamiento de información mediante la consolidación de datos históricos.
- Los datos se almacenan en un formato estructurado para consultas y análisis.
- Los datos se resumen. Los Data Warehouse generalmente no conservan tantos detalles como los sistemas orientados a transacciones.

2.2.2.1.- El modelo de datos del Data Warehouse

Según Inmon (1993) argumenta que hay tres niveles en el proceso de modelado de datos: modelado de alto nivel (llamado ERD, nivel de relación de entidad) que presenta entidades, atributos y relaciones; modelado de nivel medio (llamado conjunto de elementos de datos) que son datos establecido por departamento; y el modelado de bajo nivel (denominado modelo físico) que permite optimizar el rendimiento.

El modelo de datos físicos se crea a partir del modelo de datos de nivel medio simplemente extendiendo el modelo de datos de nivel medio para incluir claves y características físicas del modelo. En este punto, el modelo de datos físicos se parece a una serie de tablas, a veces llamadas tablas relacionales (Basaran, 2005).

En congruencia, Mullins (2007) explica lo que se puede utilizar para diferenciar el Data Warehouse de una base de datos de archivo ordinaria que puede convertirse fácilmente en un vertedero de información. Los datos se conforman (los elementos de datos se conforman para que las definiciones de “cliente” o “ingresos” signifiquen lo mismo sin importar dónde se originó), los datos son históricos (vista de la empresa en un momento determinado), los datos se comparten (se puede consultar o acceder de otra manera tiene poco valor), los datos son completos (se pueden capturar y consolidar desde múltiples sistemas).

2.2.2.2.- Técnicas de modelado Data Warehouse

El autor Ballard (1998) exploró la evolución del concepto de Data Warehousing, ya que se relaciona con el modelado de datos para el Data Warehouse, definiendo que el modelado del Data Warehouse es el proceso de construir un modelo para los datos con el fin de almacenarlos en el Data Warehouse . Existen dos técnicas de modelado de datos que son relevantes en un entorno de Data Warehousing: el modelado de relación de entidad (ER) y el modelado dimensional.

El modelado ER produce un modelo de datos del área de interés específico, utilizando dos conceptos básicos: entidades y las relaciones entre esas entidades. Los modelos ER detallados también contienen atributos, que pueden ser propiedades de las entidades o las relaciones. Por su parte, el modelado dimensional utiliza tres conceptos básicos: medidas, hechos y dimensiones. El modelado dimensional es potente para representar los requisitos del usuario comercial en el contexto de las tablas de la base de datos. Las medidas son valores numéricos que se pueden agregar y calcular.

2.2.2.3.- Modelado de diseño de bases de datos Data Warehouse

Hay tres niveles de modelado de datos; estos son conceptuales, lógicos y físicos. Para el propósito del TFM, se discutirá únicamente los dos primeros. El diseño conceptual maneja conceptos que están cerca de la forma en que los usuarios perciben los datos; el diseño lógico trata con conceptos relacionados con cierto tipo de DBMS; El diseño físico depende del DBMS específico y describe cómo se almacenan realmente los datos.

El objetivo principal del modelado de diseño conceptual es desarrollar un diseño formal, completo y abstracto basado en los requisitos del usuario.

El diseño lógico de Data Warehouse implica la definición de estructuras que permiten un acceso eficiente a la información. El diseñador construye estructuras multidimensionales teniendo en cuenta el esquema conceptual que representa los requisitos de información, las bases de datos de origen y los requisitos no funcionales (principalmente de rendimiento). Esta fase también incluye especificaciones para herramientas de extracción de datos, procesos de carga de datos y métodos de acceso al Data Warehouse. Al final de la fase de diseño lógico, se debe crear un prototipo funcional para el usuario final (Basaran, 2005, p.78).

2.2.2.4.- Desarrollar un Data Warehouse

Según Demarest (2008) asegura que la planificación del desarrollo y la implementación de un Data Warehouse estándar debe tomarse como un proyecto de Tecnología de la Información (TI), por lo tanto, lo que hace que falle un proyecto de TI también se aplica al desarrollar el Data Warehouse, de ahí la necesidad de planificar el proyecto y seguir el ciclo de vida de desarrollo de un sistema TI (p.41). Existe la necesidad de una cuidadosa planificación, especificación de requisitos, diseño, creación de prototipos e implementación. El modelo cíclico implica cinco etapas que se muestran en la figura 2.

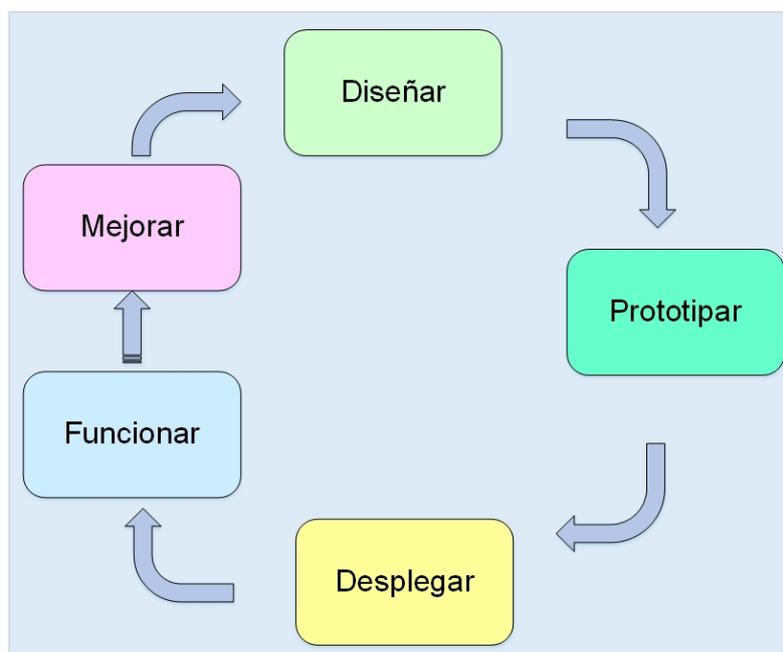


Figura 2. Modelo de ciclo de vida de desarrollo de Data Warehouse (DWLC). Fuente: (Demarest, 2008, p.41)

Donde la etapa de diseño toma información de los inventarios de datos disponibles y los requisitos de los analistas y las necesidades analíticas, de modelos de datos robustos y los convierte en marts de datos e información inteligente. La etapa de implementación del prototipo, donde un grupo de formadores de opinión y el soporte de usuarios finales, se ponen en contacto con un modelo de trabajo del Data Warehouse o diseño de Data Mart adecuado para su uso real. El propósito de la creación de prototipos cambia, ya que el equipo de diseño se mueve de un lado a otro entre el diseño y el prototipo. La etapa de implementación es la etapa de formalización del prototipo aprobado por el usuario y la recopilación de sus necesidades. La fase de despliegue es el mantenimiento diario del Data Warehouse, los servicios de entrega de datos y las herramientas del cliente que brindan a los analistas su acceso al almacenamiento y la gestión de los procesos continuos de extracción, transformación y carga que mantienen el Data Warehouse actualizado con respecto a los sistemas fuente transaccionales autorizados. La etapa de mejora es donde las condiciones comerciales externas cambian de manera discontinua, o las propias organizaciones experimentan cambios discontinuos, la mejora se traslada sin problemas al diseño fundamental, si el diseño inicial y el prototipo no cumplen con los requisitos del mercado y las necesidades del usuario.

2.2.3.- Data Mart

Un mercado de datos o DataMart como lo describe Inmon (1999) es una colección de áreas temáticas organizadas para el soporte de decisiones basadas en las necesidades de un determinado departamento de una organización. Las finanzas tienen su Data Mart, el marketing tiene el suyo y el comercio tiene el suyo, etc. Según Inmon (1999), lo más importante para los DataMarts es que los departamentos individuales poseen el hardware, software, datos y programas que constituyen su funcionalidad. Cada departamento tiene su propia interpretación de cómo debería ser un DataMart y el DataMart de cada departamento es peculiar y específico para sus propias necesidades.

Los DataMarts contienen datos operativos que ayudan a la alta dirección de las organizaciones a elaborar estrategias basadas en análisis de tendencias y experiencias pasadas. La diferencia clave es que la creación de una DataMart se basa en una necesidad específica y predefinida de una determinada agrupación y configuración de datos seleccionados. Puede haber múltiples mercados de datos dentro de una organización. Un DataMart puede admitir una función comercial, un proceso comercial o una unidad comercial en particular.

De forma explícita, para entender el significado de DataMart se postulan las siguientes definiciones:

- Los DataMart son Data Warehouse departamentales edificados rápidamente para proporcionar soluciones a las directrices de un departamento dentro de una organización, y a su vez, de la propia entidad.
- En este mismo sentido los DataMart se pueden considerar como un subconjunto de Data Warehouse diseñados para complacer las necesidades específicas de un área de las organizaciones, industrias, universidades o instituciones.
- En unísono al procesamiento de la información y los datos, se puede considerar al DataMart como la implementación de un Data Warehouse con funciones restringidas, que se utiliza en un área específica de la organización, para el caso del TFM en el área Administrativa de del Centro de Educación Continua (CEC) de la Universidad Técnica de Machala.
- También se puede decir, que un DataMart por su aplicación, es un almacén de datos que brinda soporte a un área específica dentro de una organización; este DataMart puede ser independiente, parte de una red distribuida de DataMart o dependiente de los datos de un Data Warehouse.
- En el sentido comercial, existe la analogía de que el DataMart es una tienda local que proporciona un servicio a un vecindario, en comparación con un supermercado que atiende a una ciudad o una gran población.

En semejanza a los principios y fundamentación de los Data Warehouse, los DataMart tienen las mismas características de integración, orientación temática y no propiciar un estamento volátil. En este sentido, se puede asegurar que los DataMart presentan dos tipologías; los DataMart dependientes y los DataMart independientes.

2.2.3.1.- Clasificación de los Data Mart

Como se dedujo en el anterior numeral existen dos tipologías de DataMart como son independientes y los dependientes.

DataMart Independientes: Son los DataMart que no dependen de un Data Warehouse central, porque pueden recibir los datos directamente del ambiente operacional, ya sea mediante procesos internos de las fuentes de datos o de almacenes de datos operacionales (ODS).

Es decir, los DataMart independientes procesan la información en primera línea, gestionan la información que se genera en la organización desde sus bases y la depura para poder establecer un almacén de datos útil en la toma de decisiones, la información que procesa los DataMart independientes constituyen todo el funcionamiento del área o departamento de una organización, sin la dependencia de otros DataMart o de los Data Warehouse. En la figura 3 se muestra un diagrama de los componentes de un DataMart independiente.

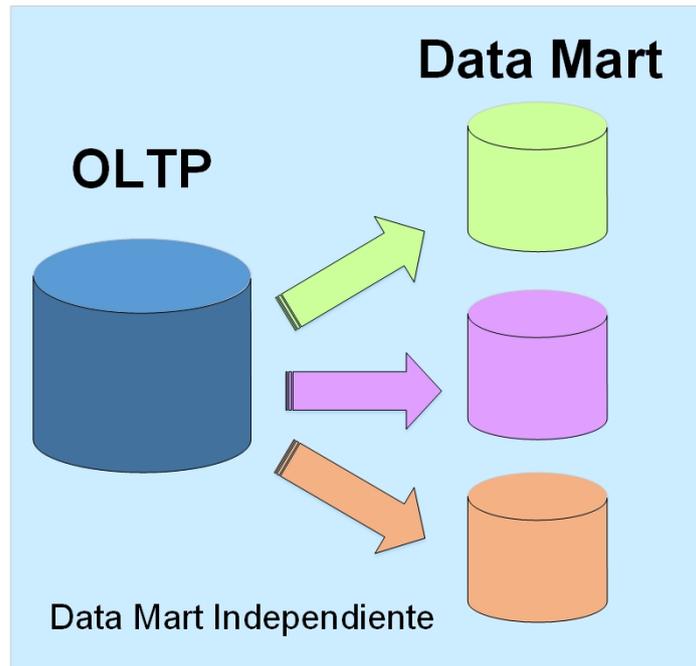


Figura 3. DataMart Independiente. Fuente: (Demarest, 2008, p.41)

DataMart Dependientes: Son los que se construyen a partir de una Data Warehouse central; es decir, que reciben sus datos de un repositorio organizacional principal.

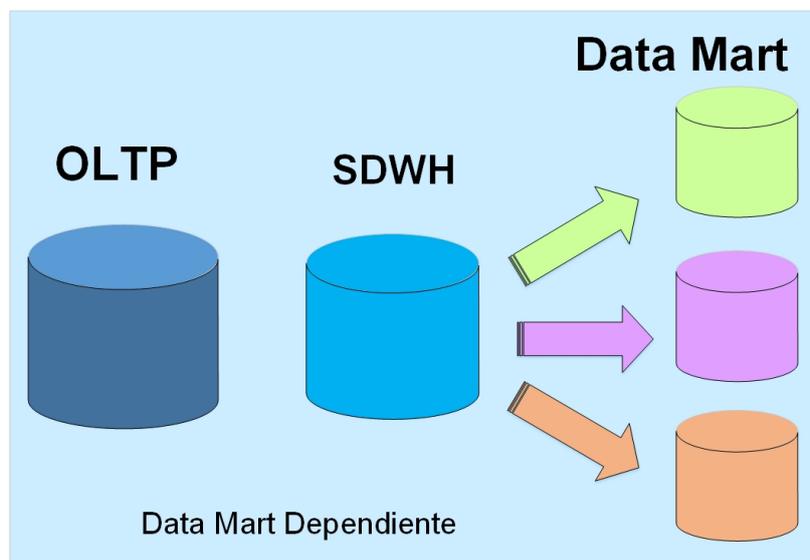


Figura 4. DataMart Dependiente. Fuente: (Demarest, 2008, p.41)

2.2.3.2.- Carga de datos en Data Mart

Para cargar los datos hacia el DataMart se pueden utilizar técnicas de carga para las herramientas OLAP, pero se debe tener en cuenta la capacidad de soportar extracción de grandes volúmenes de datos desde las fuentes, para sí no exigir en extremo la recopilación y almacenamiento de datos. Además, el tiempo de carga y la calidad de los datos a ser transportados hacia el DataMart.

La información que manipula los DataMart deben tener un procesamiento de depuración, para esto se puede usar software que acceden y administran los datos de orígenes diversos, hacia transferencia de múltiples plataformas. Una de las herramientas para la carga de datos y gestión de los mismos, por ejemplo en Microsoft SQL Server son los DTS, con los cuales se puede realizar una carga por medio de la fuente y el destino de la información; para este propósito es indispensable tener cuidado que los datos sean coherentes; es decir, sean los mismos que se necesitan, para esto los procesos de depuración establecen que es necesario eliminar inconsistencias y desarrollar una estrategia o metodología para un mantenimiento efectivo de la limpieza de datos que gestiona el DataMart.

2.2.4.- Consulta y herramientas de reporte

OLAP proporciona vistas multidimensionales y resumidas de datos comerciales y se utiliza para informes, análisis, modelado y planificación para optimizar la organización. Las técnicas y herramientas OLAP se pueden usar para trabajar con Data Warehouse o DataMart diseñados para sofisticados sistemas de inteligencia empresarial. Estos sistemas procesan las consultas necesarias para descubrir tendencias y analizar factores críticos. El software de informes genera vistas agregadas de datos para mantener a la alta dirección informada sobre el estado de los procesos operativos de la organización.

Se utilizan otras herramientas de BI para almacenar y analizar datos, como minería de datos y Data Warehouse; sistemas de soporte de decisiones y pronósticos; almacenes de documentos y gestión de documentos; conocimiento administrativo; mapeo, visualización de información y tablero de instrumentos; gestión de sistemas de información, sistemas de información geográfica; análisis de tendencia; software de servicio, entre otros.

2.2.5.- Sistemas tradicionales de inteligencia empresarial

La clave principal para un sistema BI exitoso es consolidar los datos de los diferentes sistemas operativos empresariales en un Data Warehouse organizacional, como se ha venido

mencionando en los diferentes epígrafes. Muy pocas organizaciones tienen un Data Warehouse organizacional completo. Esto se debe al amplio alcance del esfuerzo para consolidar toda la información empresarial. Las organizaciones se distinguirán por la capacidad de aprovechar la información sobre su mercado, clientes y operaciones para aprovechar las oportunidades comerciales. Moss & Atre (2003) describen el BI como una integración perfecta de aplicaciones operativas de front-office con aplicaciones operativas de back-office. Las organizaciones pueden tomar mejores decisiones, decisiones correctas en particular sobre sus clientes, proveedores, empleados, logística, infraestructura y recopilar, almacenar, acceder y analizar grandes cantidades de registros solo con BI. La figura 5 muestra la arquitectura tradicional de un sistema BI.



Figura 5. *Arquitectura tradicional de un sistema BI.* Fuente: (Demarest, 2008, p.41)

Los enfoques actuales de almacenamiento de datos y BI son ampliamente aceptados como una capa de middleware para sistemas de soporte de decisiones de última generación (Seufert & Schiefer, 2005). Sin embargo, no brindan suficiente apoyo para enfrentar los desafíos futuros, como la toma de decisiones en tiempo real y en circuito cerrado. Sin una BI efectiva para apuntar a organizaciones orientadas a procesos para su soporte, no es posible. Las empresas tienen que rediseñar sus procesos comerciales para administrar y controlar la funcionalidad de la organización de manera efectiva. Para permitir un desempeño eficiente e identificar oportunidades para mejorar continuamente, es esencial recopilar y conciliar todos los datos operativos relacionados con los procesos operativos, para la toma de decisiones analíticas. Esta toma de decisiones analíticas es esencial y necesario la integración de

componentes de los DataMart en cada área de la organización. Los componentes para una arquitectura de BI efectiva requieren un Data Warehouse bien desarrollado, un DataMart efectivo y una gestión de metadatos, herramientas analíticas como minería de datos y OLAP y otras herramientas de informes de consultas y también una base sólida de hardware que puede soportar la escalabilidad tanto de los datos, así como las perspectivas del usuario.

Pero para lograr una ventaja competitiva, las organizaciones también se esfuerzan por reducir el tiempo necesario para reaccionar a las operaciones comerciales relevantes. Al organizar y desplegar BI según las características propias de la organización, se puede liberar el valor completo de los datos almacenados en toda la empresa.

2.2.6.- OLAP versus OLTP

El Data Warehouse también se puede denominar sistema de Procesamiento Analítico en Línea (OLAP) porque ayuda a los tomadores de decisiones en el ámbito del análisis de datos. Por otra parte, los sistemas operativos se pueden denominar sistemas de Procesamiento de Transacciones en Línea (OLTP) porque se encargan de las transacciones y procesos diarios de la organización. La diferencia se resume en que el objetivo del sistema OLTP es obtener datos en bases de datos, mientras que el Data Warehouse está construido para obtener información de la base de datos.

Según Han, Kamber & Pei (2012) establecen que el sistema OLTP es orientado al cliente, en contraste, un Data Warehouse está orientado al mercado. No es una buena práctica combinar las tareas OLAP y OLTP en un sistema, porque la base de datos OLAP está optimizada para consultas y la base de datos OLTP está optimizada en tiempo real para reportes. El Data Warehouse se basa en un modelo de diseño dimensional que es mucho más eficaz para la recuperación de datos, mientras que los sistemas OLTP se basan en un diseño de modelo relacional. Además, los Data Warehouse pueden contener datos de diversas fuentes de datos con formatos de datos heterogéneos que se manipulan antes de cargarse en el Data Warehouse.

Las organizaciones, generalmente, eligen separar los sistemas OLTP y OLAP. Según Poe, Klauer & Brobst (1996) porque los datos en los sistemas OLAP se utiliza para extraer conocimiento e información a través de comparaciones o mediante la detección de patrones y tendencias. Dicha información no se puede descubrir fácilmente del sistema OLTP debido a la complejidad. Además, sostienen que la idea del Data Warehouse debe basarse en requisitos comerciales. Para Kimball (1996) este concepto es idóneo, por su parte Inmon

(2002) tiene un enfoque diferente para el desarrollo del Data Warehouse; él autor sostiene que a pesar de que los sistemas OLTP se desarrollan en función de los requisitos comerciales, los datos se desarrolla del sistema de almacenamiento, por su parte los DataMart están diseñados para atender a las necesidades de los individuos y los procesos de negocio.

Desde otra perspectiva, Santos & Ramos (2009) definen el Data Warehouse como una base de datos que se administra independientemente de una base de datos operativa (OLTP) (p.33). La tabla 1 muestra que ambos tipos de bases de datos tienen diferentes propósitos. La base de datos operacional apunta a manejar transacciones casi en tiempo real (es decir, transacciones de tarjetas de crédito) mientras los datos del Data Warehouse está más dedicado a analizar un volumen de datos de bugger. Todas estas diferencias deben tenerse en cuenta en el proceso de selección de una nueva base de datos.

Tabla 1. Diferencias entre la base de datos OLTP y la base de datos OLAP

Base de Datos OLTP	Warehouse OLAP
Propósitos operacionales	Historial de registros
Acceso de lectura / escritura	Acceso de solo lectura
Acceso a transacciones predefinidas	Informes periódicos y acceso ad hoc.
Acceso a una pequeña cantidad de registros	Acceso a una gran cantidad de registros
Actualización de datos casi en tiempo real	Cargas de datos programadas
Estructura optimizada para actualizaciones	Estructura optimizada para procesar problema

Fuente: Elaboración propia

2.2.7.- Enfoques de diseño del Data Warehouse

Bill Inmon y Ralph Kimball son dos pioneros que construyeron dos diferentes enfoques para el almacenamiento de datos y que toman su nombre para denominar la arquitectura.

2.2.7.1.- Arquitectura Bill Inmon

Bill Inmon propuso un enfoque de arriba hacia abajo para el almacenamiento de datos que se muestra en la figura 6, conocida como Corporate Information Factory (CIF) o fábrica de información corporativa (Inmon, 2002). Los componentes de un CIF incluyen un Data Warehouse que se construye en una tercera forma normalizada e individual des-normalizada de DataMart que se fundamentan desde el Data Warehouse. Estos DataMart satisfacen las

necesidades individuales de los procesos organizacionales. Los cubos de informes se crean según sea necesario en la parte superior de los DataMart.

En la figura 6, la alimentación del Data Warehouse comienza por el proceso de Extracción, Carga y Transformación (ETL) de diferentes fuentes. Primero, los datos se extraen y transforman en el área de ensayo. A continuación, una vez que los datos estuvieron listos, se cargan en el Data Warehouse empresariales centralizado en la tercera forma normal (3NF). Finalmente, una vez que se cargan los datos, los Data Mart se diseñan de acuerdo con los requisitos comerciales.

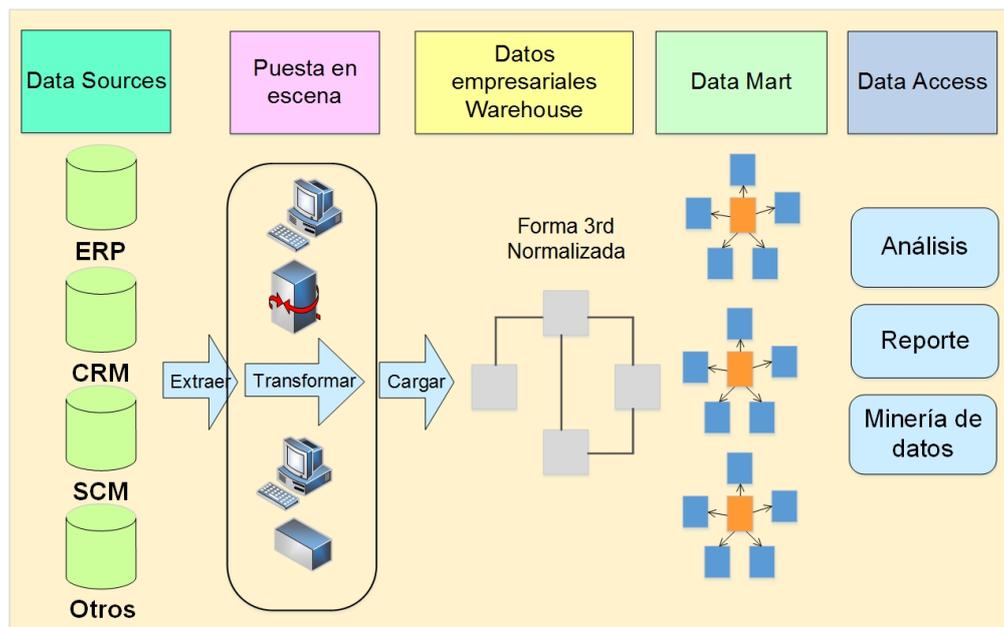


Figura 6. Enfoque de Bill Inmon para el Data Warehousing. Fuente: Rangarajan (2016)

El Data Warehouse, tal como lo define Bill Inmon, contiene datos empresariales sin redundancia al más bajo nivel de detalle; es decir, datos transaccionales en 3NF. Según Rangarajan (2016) resume las ventajas y desventajas del enfoque de Inmon de la siguiente manera:

Las ventajas clave del enfoque de Inmon son:

- El Data Warehouse realmente sirve como la única fuente verdadera para la organización, ya que es la única fuente para los DataMart y todos los datos en el Data Warehouse están integrados.
- Se evitan las anomalías en la actualización de datos debido a la baja redundancia. Esto hace que el proceso ETL sea más fácil y menos propenso a errores.
- Los procesos comerciales se pueden entender fácilmente, ya que el modelo lógico representa las entidades comerciales detalladas.

- Muy flexible, a medida que cambian los requisitos comerciales o cambian los datos de origen, es fácil actualizar el Data Warehouse ya que la información está en un solo lugar.
- Puede gestionar diversas necesidades de informes en toda la organización.

A continuación se presentan algunas de las desventajas del método Inmon:

- El modelo y la implementación pueden volverse complejos con el tiempo, porque implica más tablas y combinaciones.
- La necesidad de recursos expertos en modelado de datos y en los procesos de la organización. Este tipo de recursos puede ser difícil de encontrar y a menudo son caros.
- La configuración inicial y la entrega tomarán más tiempo, y la gerencia debe ser consciente de esto.
- Se necesita más trabajo ETL porque los DataMart se crean desde el Data Warehouse.
- Se necesita un equipo bastante grande de especialistas para gestionar con éxito el medio ambiente (Breslin, 2004).

2.2.7.2.- Arquitectura Ralph Kimball

El segundo enfoque presentado por Ralph Kimball, conocido como arquitectura de depósito de datos dimensionales, consideró un diseño ascendente como se muestra en la figura 7 (Kimball & Ross, 2013, p.63).

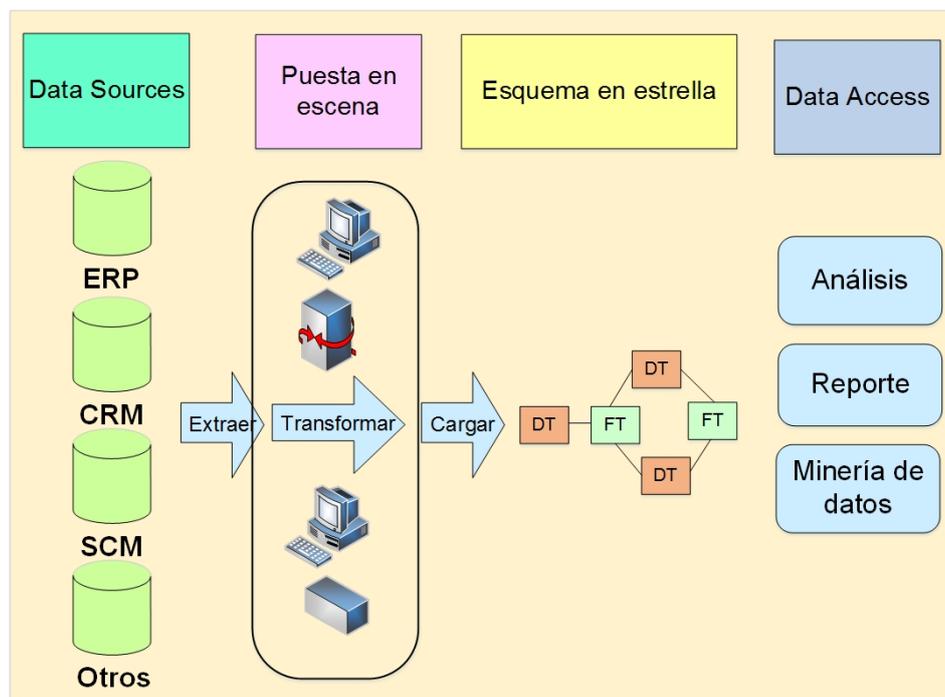


Figura 7. Enfoque de Ralph Kimball para el Data Warehousing. Fuente: Rangarajan (2016)

En este enfoque, los usuarios empresariales tienen una estructura dimensional simple al principio, y cuando se combinan crearán un amplio Data Warehouse. El factor clave del modelado dimensional es la simplicidad. Kimball & Ross en (2013) argumentan que el uso del modelado dimensional en el Data Warehouse hace que los datos sean más fáciles de entender incluso para usuarios comerciales, no solo para expertos en Data Warehousing.

Dado que el enfoque de Kimball es un enfoque de abajo hacia arriba, el Data Warehousing en la figura 7 comienza con el proceso ETL para uno o más procedimientos de negocios. Los datos se extraen primero de diferentes fuentes de sistemas de organización, luego se transforman en el área de preparación para cada DataMart por separado. A continuación se carga en las tablas de hechos/dimensiones correspondientes en el DataMart. Finalmente, los informes están diseñados para obtener datos directamente de los DataMart.

La idea principal del enfoque de Kimball es construir el Data Warehouse de forma incremental a lo largo del tiempo combinando los Data Mart.

Según Rangarajan (2016) resume las ventajas y desventajas del enfoque de Inmon de la siguiente manera:

Estas son algunas de las ventajas del método Kimball:

- Rápido de configurar y construir, y la primera fase del proyecto de Data Warehousing se entregará rápidamente.
- El esquema de estrella puede ser fácilmente entendido por los usuarios comerciales y es fácil de usar para generar informes. La mayoría de las herramientas de BI funcionan bien con el esquema en estrella.
- La huella del entorno de Data Warehousing es pequeña; ocupa menos espacio en la base de datos y facilita la gestión del sistema.
- El rendimiento del modelo de esquema en estrella es muy bueno. El motor de la base de datos realizará una “unión en estrella” en la que se creará un producto cartesiano utilizando todos los valores de dimensión y la tabla de hechos se consultará finalmente para las filas selectivas. Esto se sabe que es una operación de base de datos muy efectiva.
- Un pequeño equipo de desarrolladores y arquitectos es suficiente para mantener el Data Warehousing funcionando de manera efectiva.
- Funciona realmente bien para las métricas en función del departamento y el seguimiento de KPI, ya que los mercados de datos están orientados a la presentación de informes por departamento o por proceso empresarial.

- Profundiza, donde una herramienta de BI atraviesa múltiples esquemas en estrella para generar un informe se puede lograr con éxito utilizando dimensiones confirmadas.

Estas son algunas de las desventajas del método Kimball:

- La esencia de la “única fuente de verdad” se pierde, ya que los datos no se integran completamente antes de satisfacer las necesidades de informes.
- Los datos redundantes pueden causar anomalías en la actualización de datos.
- Agregar columnas a la tabla de registros puede causar problemas de rendimiento. Esto se debe a que las tablas de registros están diseñadas para ser muy profundas. Si se van a agregar nuevas columnas, el tamaño de la tabla de registros se vuelve mucho más grande y no funcionará bien.
- No se pueden manejar todas las necesidades de informes de la organización porque el modelo está orientado hacia los procesos comerciales en lugar de la organización en su conjunto.
- La integración de datos heredados en el Data Warehouse puede ser un proceso complejo.

En base a la arquitectura de Ralph Kimball, se puede establecer una metodología, denominada metodología de Kimball.

2.2.7.3.- Metodología de Kimball

La metodología se basa en lo que Kimball denomina Ciclo de Vida Dimensional de Negocio (Business Dimensional Lifecycle), este ciclo de vida de la solución de Data Warehouse y DataMart, está basado en cuatro principios básicos:

Centrarse en el negocio: Hay que concentrarse en la identificación de los requerimientos del negocio y su valor asociado, y usar estos esfuerzos para desarrollar relaciones sólidas con el negocio, agudizando el análisis del mismo y la competencia consultiva de las partes interesadas.

Construir una infraestructura de información adecuada: Diseñar una base de información única, integrada, fácil de usar, de alto rendimiento donde se reflejará la amplia gama de requerimientos de negocio identificados.

Realizar entregas en incrementos significativos: Crear el Data Warehouse en incrementos entregables en plazos de 6 a 12 meses. Hay que usar el valor de negocio de cada elemento identificado para determinar el orden de aplicación de los incrementos. En esto la metodología se parece a las metodologías ágiles de construcción de software.

Ofrecer la solución completa: Proporcionar todos los elementos necesarios para entregar valor a los usuarios de negocios. Para comenzar, esto significa tener un Data Warehouse sólido, bien diseñado, con calidad probada, y accesible. También se deberá entregar herramientas de consulta ad hoc, aplicaciones para informes y análisis avanzado, capacitación, soporte, sitio web y documentación.

La razón de ser de los proyectos de Business Intelligence y de muchos otros, es el negocio, por lo tanto, uno de los puntos importantes es tener claro que las necesidades del negocio son las que guiarán a lo largo de toda la solución. En general se contempla que el ciclo de vida dimensional del negocio se puede expresar en términos de lo que se representa en la figura 8.

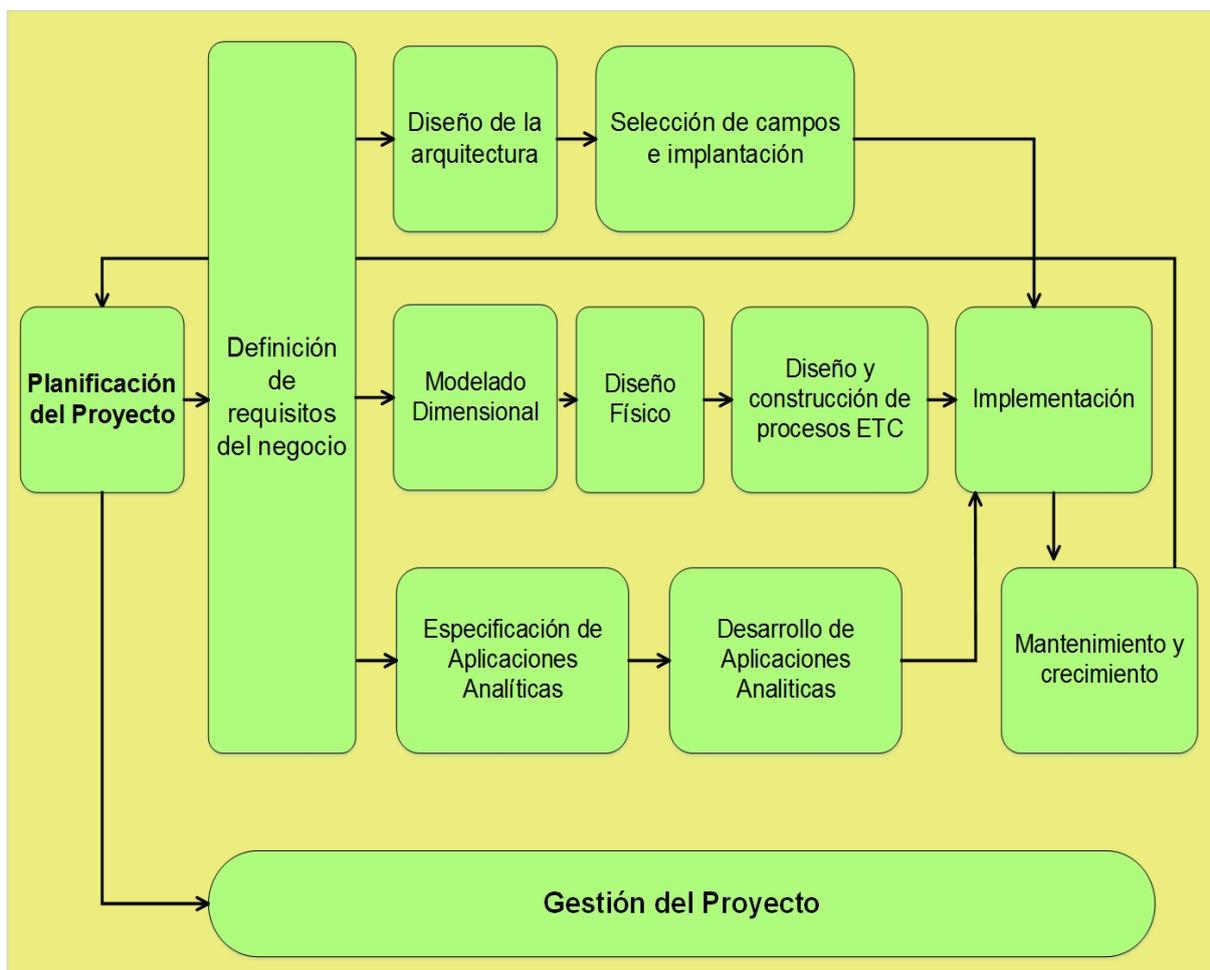


Figura 8. Ciclo de vida de la metodología Ralph Kimball. **Fuente:** Elaboración propia basado en Gutiérrez (2014)

Esta metodología es la más utilizada para el diseño e implementación de DataMart y se compone de 4 pasos esenciales.

1.- Seleccionar el proceso de negocios que se desea modelar.

2.- Declarar la granularidad del proceso de negocios.

3.- Escoger las dimensiones que describen cada fila de la tabla de hechos.

4.- Identificar los hechos numéricos que poblaran cada fila de la tabla de hechos.

1.- Seleccionar el proceso de negocio:

Seleccionar el proceso de negocio permite delimitar la información que estará contenida en el DataMart. Todos los datos que se consideren importantes para el proceso de negocio deben ser incorporados en el DataMart.

El enfocarse en procesos de negocio, en lugar de departamentos de negocio, permite crear un modelo dimensional transversal a la organización, permitiendo la integración de todos los departamentos y áreas que participan en el proceso que se desea modelar, asegurando así, la consistencia en la información. Múltiples flujos de información, a través de múltiples modelos dimensionales, conducen indudablemente a la duplicidad de la información, y por ende, a su inconsistencia. La mejor manera de asegurar la consistencia es publicar los datos una sola vez.

2.- Declarar la granularidad:

Cuando se habla de granularidad, se refiere al nivel máximo de detalle de los datos relacionados con el proceso de negocios. La importancia de la granularidad radica en que define el nivel máximo de detalle que estará disponible para realizar el análisis, lo que permite dimensionar además la cantidad de datos que contendrá el DataMart. En este sentido, es importante determinar que existe un trade-off entre nivel de granularidad y cantidad de datos a guardar.

La granularidad es un paso crítico, y que a la vez facilita el desarrollo de los pasos 3 y 4. Sin embargo, si estando en cualquiera de los pasos siguientes, se observa que la declaración de granularidad es errónea, siempre se puede volver al paso 2 y re definir la granularidad del modelo, para continuar con los siguientes pasos.

3.- Escoger las dimensiones:

Cada proceso de negocio está determinado por datos numéricos que permiten hacer el análisis del proceso y datos sobre estos otros datos numéricos que permiten describir el contexto del proceso o también denominado metadata del proceso. El tercer paso se refiere a seleccionar las dimensiones que se van a aplicar a cada fila en la tabla de hechos. Se

persigue acompañar la tabla de hechos con un conjunto robusto de dimensiones que puedan representar todas las posibles descripciones que toman los valores simples, en el contexto de cada métrica. Si se tiene clara la granularidad del modelo, entonces las dimensiones pueden ser identificadas muy fácilmente. Al escoger cada dimensión, se listarán todos los atributos discretos y de tipo texto que permitirán rellenar la dimensión con la información descriptiva que las métricas requieren.

4.- Escoger los hechos:

El cuarto paso se refiere a seleccionar los valores numéricos sobre los que se realizará el análisis del proceso de negocio. Todos los candidatos a métrica en el diseño del modelo, deben estar acorde a la granularidad definida en el paso 2. Por lo tanto, los hechos que claramente se encuentren en un nivel granular distinto, deben ser incluidos en una nueva tabla de hechos.

2.2.7.4.- Modelo multidimensional

Luego de cumplir estos pasos el modelo resultante se implementa como un modelo multidimensional, para el cual existen dos tipos: modelo estrella y modelo copo de nieve. En un modelo dimensional, los hechos numéricos del proceso de negocio se incluyen en las tablas de hechos, las cuales se relacionan con un conjunto de tablas que corresponden al contexto del proceso mismo.

Modelo estrella:

El modelo estrella consiste en ubicar una tabla de hechos en el centro del modelo, relacionándose con el resto de las tablas dimensionales, de tal modo que todas ellas se relacionen sólo a través de esta tabla de hechos.

El tema principal en un modelo estrella es el diseño de las dimensiones, puesto que para que éstas se relacionen sólo con la tabla de hechos y así evitar la relación con otras dimensiones, es estrictamente necesario ubicar en una sola tabla, todo aquello que se pueda deducir del elemento más granular de la tabla considerado por la dimensión, así como también las jerarquías que podrían surgir en torno a este elemento.

Dentro de las ventajas de un esquema estrella, se encuentra su simplicidad y velocidad para ser utilizados en análisis multidimensionales, permitiendo acceder a datos agregados como también a datos de detalle. La simplicidad de este esquema es una ventaja muy apreciada por los usuarios finales. Las consultas no son complicadas, puesto que sólo involucran a la

tabla de hechos y las dimensiones de manera directa, evitando encadenar uniones y condiciones a dos o más niveles. Por otra parte, este esquema optimiza la velocidad al permitir indexar las dimensiones de forma individualizada, sin que esto impacte en el rendimiento de la base de datos. La figura 9 muestra la estructura del modelo estrella.

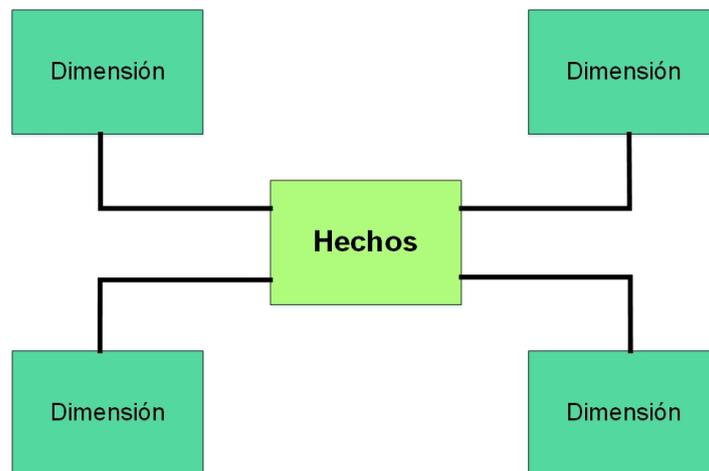


Figura 9. Modelo estrella. Fuente: Elaboración propia

Modelo copo de nieve:

El modelo copo de nieve está basado en un modelo estrella, donde las dimensiones se normalizan en múltiples tablas. Esta normalización se basa en las jerarquías que presenta cada dimensión, dejando cada nivel jerárquico explicitado en una nueva tabla dimensional. Con esto, la tabla de hechos deja de ser la única vía para relacionar tablas, lo cual dificulta la comprensión del modelo para los usuarios finales, además de disminuir el rendimiento, debido a los Joins que se deben emplear para realizar las consultas.

A favor de este esquema, como todo modelo normalizado, disminuye la redundancia de datos, ahorrando espacio y contribuyendo a su mantención. El ahorro de espacio es un tema que se debe evaluar con el usuario final, puesto que con los avances en hardware, esta preocupación ha disminuido. Este esquema se recomienda cuando se tienen grandes dimensiones, y donde sus registros presentan distintos comportamientos.

Un ejemplo donde conviene normalizar, es para una tabla calendario de una organización retail con presencia en distintos países. En estos casos, cada país posee distintos calendarios, basados en su forma de ver la semana, de acuerdo a los días festivos, etc. por este motivo, se recomienda dejar en la dimensión base todos los campos en común, dejando los atributos particulares de cada calendario en nuevas dimensiones para cada calendario. La figura 10 muestra la estructura del modelo copo de nieve, su característica principal es que se estructura por varias tablas dimensionales relacionadas por una tabla de hechos.

2.2.7.5.- Diferencias entre los enfoques de Inmon y Kimball

Ahora, en el caso del enfoque de Inmon del Data Warehousing, la arquitectura sugiere que se construya un Data Warehouse 3NF como primer paso, que contendría todos los datos de la organización, y luego construir una capa de DataMart para admitir la capa de informes basada en la arquitectura del Data Warehouse.

Sin embargo, en el caso del enfoque de Kimball, la idea es construir la capa de DataMart justo después de la capa de preparación. Estos DataMart se adaptan a los procesos comerciales identificados de manera individual. Todos los DataMart juntos forman el Data Warehouse como lo define Kimball. Sin embargo, las dimensiones comunes entre los procesos operativos de la organización se comparten entre ellos, sin crear una versión separada del mismo, para mantener una única versión de la verdad y facilitar la actualización. Estas se llaman dimensiones conformadas. Con este enfoque, no se necesita una segunda capa de preparación, y dado que los DataMart son específicos de un proceso comercial, se pueden generar informes a partir de él, sin esperar a que se diseñen e implementen el resto de DataMart.

2.2.7.6.- Factores decisivos

En base al argumento de los pros y los contras de los enfoques de Kimball e Inmon, surge una pregunta. ¿Qué y cuándo debe usarse un enfoque? Los arquitectos del Data Warehouse enfrentan esta pregunta cada vez que comienzan a construir un Data Warehouse. Estos son los factores decisivos que pueden ayudar a un arquitecto a elegir entre los dos métodos:

Requisitos de informes: Si los requisitos de informes son estratégicos y se necesitan informes integrados para toda la organización, entonces Inmon funciona mejor. Si los requisitos de presentación de informes son tácticos y orientados al proceso de negocio/equipo, entonces el método de mejor función es el de Kimball.

Proyecto de urgencia: Si la organización tiene tiempo suficiente para esperar la primera entrega del Data Warehouse (por ejemplo, de 4 a 9 meses), se puede seguir el enfoque de Inmon. Si el tiempo es limitado para que el Data Warehouse esté en funcionamiento (por ejemplo, de 2 a 3 meses), entonces el enfoque de Kimball es el mejor (Breslin, 2004).

Plan de dotación de personal futuro: Si la organización puede permitirse tener un equipo de especialistas de gran tamaño para mantener el Data Warehouse, entonces se puede seguir

el método Inmon. Si el plan futuro para el equipo es pequeño, entonces el método de Kimball es el más adecuado.

Frecuencia de cambios: Si se espera que los requisitos de informes cambien más rápidamente y se sepa que los sistemas de origen son volátiles, entonces el enfoque de Inmon funciona mejor, ya que es más flexible. Si los requisitos y los sistemas fuente son relativamente estables, se puede utilizar el método Kimball.

Cultura de la organización: Si los patrocinadores del Data Warehouse y la alta dirección de la organización entienden la propuesta de valor del Data Warehouse y están dispuestos a aceptar un valor permanente de la inversión del Data Warehouse, el enfoque de Inmon es mejor. Si los patrocinadores no se preocupan por los conceptos pero quieren una solución para mejorar en la presentación de informes, entonces el enfoque de Kimball es suficiente.

2.2.8.- Hipótesis

- El diseño de un DataMart (BI) optimizará la gestión de la información y el conocimiento institucional del área Administrativa del Centro de Educación Continua (CEC) de la Universidad Técnica de Machala maximizando la seguridad, racionalidad y estrategia prospectiva de los estudiantes en garantía a la toma de decisiones para consolidar la calidad educativa

CAPÍTULO 3. OBJETIVOS Y METODOLOGÍA DEL TRABAJO

3.1.- Objetivos

3.1.1.- Objetivo general

- Diseñar un DataMart BI (Business Intelligence) para la toma de decisiones factible al desarrollo del área Administrativa del Centro de Educación Continua (CEC) de la Universidad Técnica de Machala.

3.1.2.- Objetivos específicos

- Identificar y diagnosticar los procesos de toma de decisiones para el análisis de requerimientos de información acorde a la perspectiva, procesos y control de la gestión del conocimiento con la tecnológica en el área Administrativa del CEC.
- Definir el alcance, objetivos, requisitos y partes interesadas claves para el desarrollo de la base de datos transaccional junto con las fuentes de datos que incita el Data Warehouse.
- Fundamentar la utilización del DataMart para la inteligencia de negocios inherente a mejorar la calidad educativa y el procesamiento de la información en el CEC de la Universidad Técnica de Machala.
- Diseñar un DataMart en base al método de Ralph Kimball que cumpla con los requisitos necesarios para obtener indicadores deseados inherentes a la toma de decisiones en el CEC de la Universidad Técnica de Machala.
- Proponer un prototipo de interfaz de usuario para la evaluación del DataMart y desplegar el cubo OLAP eficiente a la toma de decisiones del CEC de la Universidad Técnica de Machala.

3.2.- Descripción del proyecto

La solución a desarrollar en el presente TFM consiste en el diseño de un DataMart para el Área Administrativa del Centro de Educación Continua de la Universidad Técnica de Machala, institución enmarcada en la educación superior de calidad y la formación permanente de la sociedad; a fin de que se convierta en una herramienta viable que permita a los usuarios de esta área en la toma de decisiones. Sin embargo, la solución no sólo consiste en el modelo

multidimensional del DataMart; esta solución abarca varios puntos que operando de manera integrada forman la solución integral de la inteligencia de negocios:

El DataMart debe permitir cumplir con las necesidades de información requerida por el CEC; es decir, debe contener la funcionalidad correcta; sin ella, el DataMart no cumple con su objetivo principal y el proyecto no logra proponer una solución de negocios. Por ello, el DataMart debe estar orientado a satisfacer las necesidades del Área Administrativa y la formación permanente de los estudiantes con la educación de calidad.

El DataMart necesita del modelo multidimensional porque es la base fundamental de la solución de inteligencia de negocio. El modelo multidimensional a desarrollar será del tipo estrella y tendrá las dimensiones y hechos necesarios para abarcar los posibles escenarios y supuestos del CEC de la Universidad Técnica de Machala. Para consumir esta actividad se levantará información y definido los requerimientos de los usuarios del Área Administrativa. Únicamente así se puede proceder con el diseño del DataMart. En el diseño del DataMart, primero se elaborará el respectivo modelo lógico para ver las relaciones e interacciones entre tablas de datos. Después, se realizará el modelo físico en el cual se definirá los tipos de datos y variables de dimensión o tamaño más efectivo para los campos de las tablas.

En este mismo sentido descriptivo, un DataMart es una base de datos de un departamento o área específica, y como tal requiere información, por ello es importarle la carga de datos fehacientes y válidos. Para este propósito, la solución incluye un proceso ETL el cual extraerá la información de una o más fuentes de datos, adaptará la estructura de datos a fin de que coincida con la estructura del modelo del DataMart y realizará la carga de datos al mismo. La fuente de datos puede ser una base de datos, archivos o documentación digital. El proceso ETL extraerá la información de estas tablas y se realizará las transformaciones necesarias para cargar los datos al gestor de base de datos seleccionado.

Por otra parte, si el modelo multidimensional constituye en gran parte la solución, es la herramienta de explotación finalmente con la que el usuario puede interactuar y visualizar el DataMart. La herramienta de explotación es un software que recibe una base de datos de modelo dimensional y permite visualizarla de una manera fácil e intuitiva. Se debe tener en cuenta que la solución no incluye la implementación de una herramienta de explotación. Existen diversas herramientas de este tipo en el mercado, para ello se seleccionará el software de mejores facilidades y prestaciones, la misma que se configurará para que acepte al modelo multidimensional diseñado. Finalmente con la herramienta de explotación configurada y con

la información cargada el usuario ya podría empezar a cruzar las diversas dimensiones y crear sus propios reportes que se consideren útiles para la toma de decisiones.

3.3.- Metodología de trabajo

En este epígrafe se hará un análisis de la metodología utilizada para la planificación y gestión del proyecto, así como la metodología empleada para el diseño del DataMart, o en su propósito conocida como metodología de inteligencia de negocios.

3.3.1.- Metodología de gestión del proyecto

Un aspecto fundamental en el diseño de proyectos ligados a la inteligencia de negocios y aspectos tecnológicos es la metodología de trabajo que se utilice. Existe una variedad de metodologías de gestión de proyectos, las metodologías más reconocidas como PMBOK, PRINCE2, APMOK, entre otras; y también metodologías ágiles como SCRUM o XP, entre otras. Las metodologías ágiles pretenden brindar el máximo beneficio posible al proyecto, mientras que las metodologías tradicionales tienen como propósito elaborar un producto en un tiempo previsto en el costo estimado, para conseguir esto realiza tareas de vigilancia y seguimiento para mitigar los retrasos. En base a este argumento, se puede decir que las dos metodologías pueden colaborar porque no son contrarias e inducir un mayor beneficio en el desarrollo de un proyecto tecnológico.

En base a lo argumentado, se hace un breve análisis de las metodologías de trabajo planteadas. La metodología APMOK es una metodología orientada a gestionar proyectos en el sector de la industria, PRINCE2 es un método de gestión de proyectos que se basa en la gestión de procesos y se enfoca en productos, CCPM es una metodología que se basa en la metodología y algoritmos derivados de la teoría de restricciones. XP es una metodología de desarrollo ágil que potencia las relaciones interpersonales como base para el éxito del desarrollo de software, promover el trabajo en equipo y el clima laboral.

Si bien son cierto todas las metodologías planteadas apoyan a la gestión y desarrollo de un proyecto de software, sin embargo, se ha considerado PMBOK y SCRUM para el presente trabajo debido a que poseen una generación de documentación completa en el caso de PMBOK y fácil de implementar en el caso de SCRUM.

Metodología PMBOK: Esta metodología es patrocinada por PMI (Project Management Institute) para gestión de proyectos; PMBOK es una norma reconocida en la administración

de proyectos, puesto que postula métodos, procesos y prácticas establecidas, estandariza el vocabulario a nivel de proyectos. Esta metodología sirve para administrar varios proyectos y organizarlos adecuadamente en portafolios (Pérez, 2012). La figura 12 muestra la arquitectura de la metodología PMBOK.

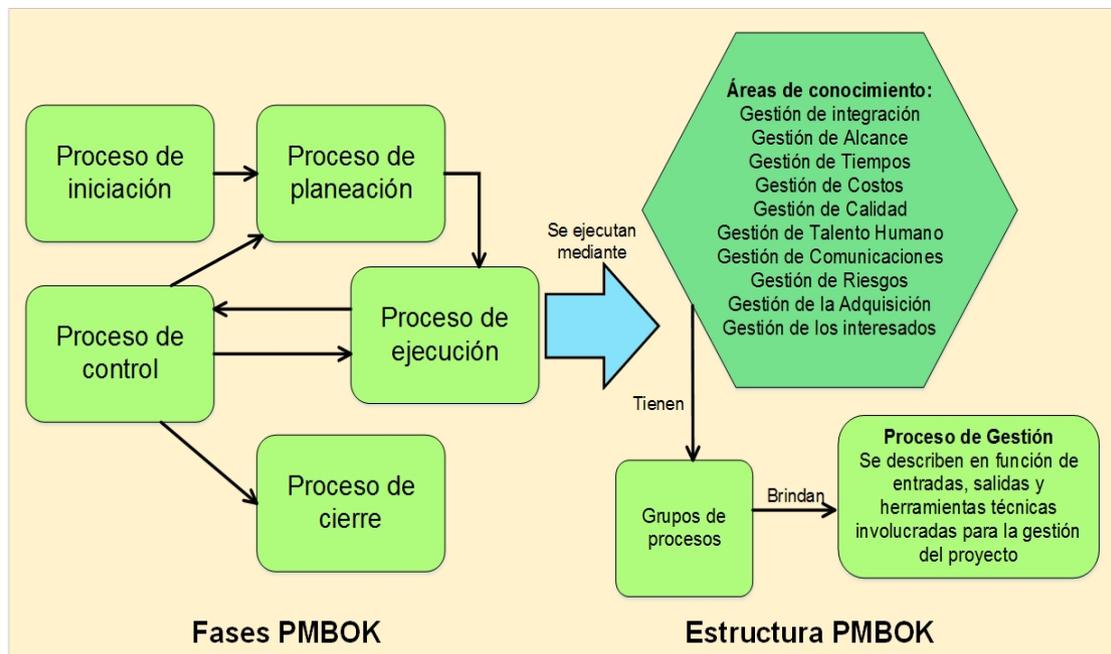


Figura 12. Arquitectura de la metodología PMBOK. Fuente: Elaboración propia adaptado de Pérez (2012).

PMBOK define grupos de procesos para la gestión de proyectos desde la visión administrativa, entre estos grupos están: procesos de iniciación, de planeación, de control, de ejecución y de cierre. Estos grupos de procesos agrupan procesos de gestión que se deben considerar en las áreas de conocimiento que requiere un proyecto; es decir, contemplan todas las tareas a realizar en cada proceso administrativo de una forma integral.

PMBOK posee diez áreas de conocimiento, las cuales se deben considerar en el proyecto para que tenga éxito, estas áreas son: la gestión de integración, del alcance, del tiempo, del costo, de la calidad, del Talento Humano, de la comunicación, de los riesgos, de las adquisiciones y la gestión de los interesados del proyecto. Para cada área se requieren realizar actividades que implican contar con entradas y salidas que son documentos que se adjuntan al proyecto para su realización (PMI, 2013).

Metodología SCRUM: El término SCRUM (Hirotaka & Nonaka, 1986) tiene el significado de “melé”, que es un tipo de jugada del Rugby. Se acoge este nombre para describir a los marcos de desarrollo ágiles que: adoptan una estrategia de desarrollo incremental, basan la calidad

del resultado más en el conocimiento tácito de las personas, en equipos auto-organizados, y utilizan solapamiento para las fases de desarrollo en lugar de un ciclo de cascada o ciclo secuencial (Palacio, 2011).

En la figura 13 se visualiza que SCRUM está formado por roles, artefactos y eventos. Los roles se dan a las personas interesadas o actores que van a operar la solución, estos roles son: propietario del producto (Product Owner), facilitador (Scrum Master) y equipo de desarrollo (Team).

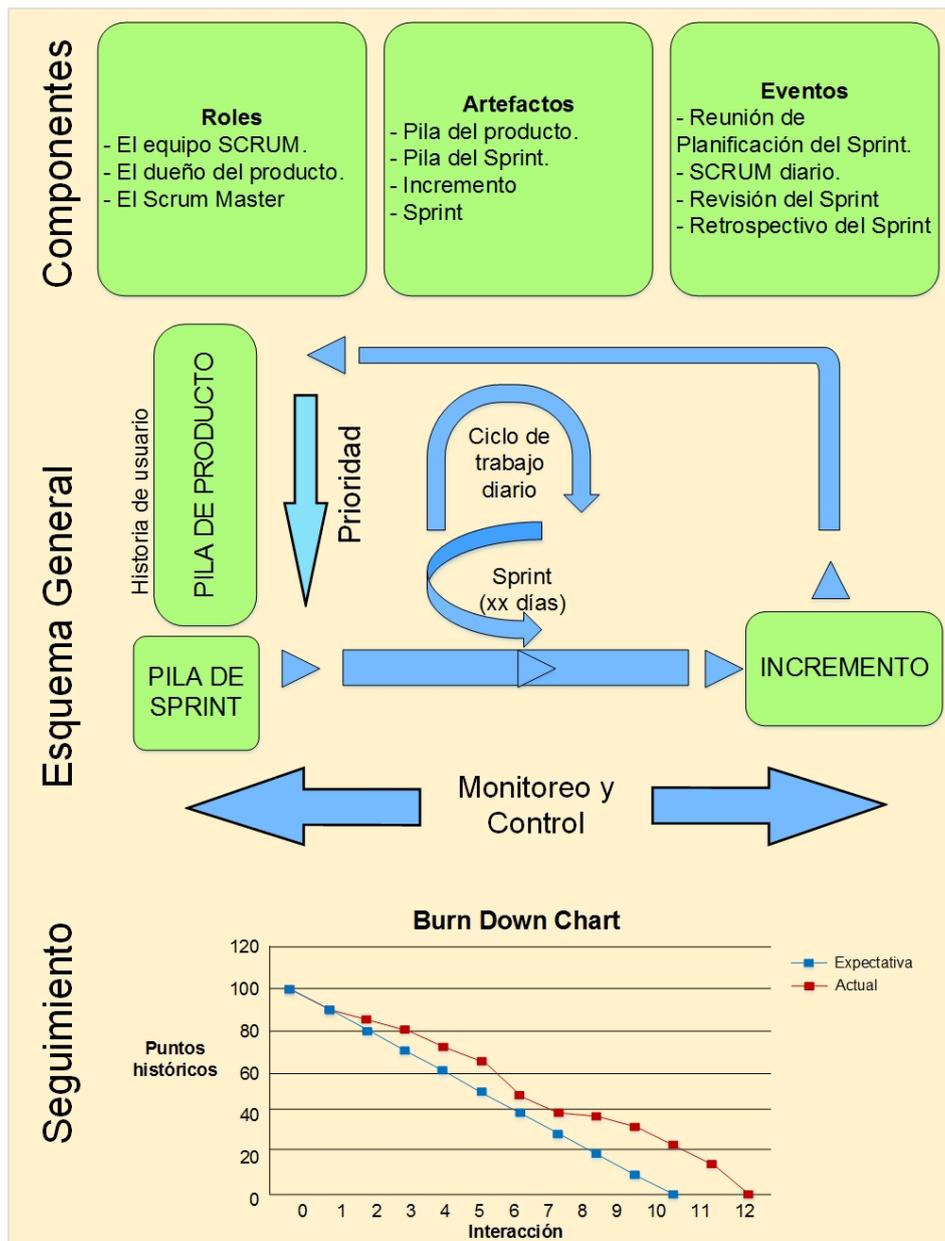


Figura 13. Arquitectura de la metodología SCRUM. Fuente: Elaboración propia adaptado de Palacio (2011).

SCRUM responde a los lineamientos del 'Manifiesto por el desarrollo ágil', que señala que para el desarrollo de software se tiene que valorar a los individuos y su interacción sobre los procesos y las herramientas, el software que funciona por encima de la documentación exhaustiva, la colaboración con el cliente por encima del negocio y la respuesta al cambio sobre el seguimiento de una planificación.

En el mismo sentido, el Product Owner representa la voz del cliente, y el que asegura que el equipo SCRUM trabaje de forma correcta desde la perspectiva del negocio, detalla historias de usuario, que son la representación de un requisito escrito, utiliza el lenguaje de usuario, prioriza y coloca en la pila de producto o Product Backlog.

El Scrum Master elimina obstáculos que retrasan el avance del objetivo de un periodo de trabajo o Sprint. No es un líder del grupo, puesto que este se auto-organiza, sino trabaja como un monitoreo entre el grupo y cualquier influencia que limite las tareas, hace cumplir las normas de la metodología.

Por otra parte, los artefactos que se utilizan en el método son la pila de producto o Product Backlog, pila de Sprint o Sprint Backlog, el Sprint y el incremento que se efectúa en el producto. El Product Backlog constituye los requisitos de usuario, que a partir de la visión inicial del producto evoluciona durante el diseño. El Sprint Backlog conforma las tareas que ejecuta el grupo durante el Sprint para generar el incremento anticipado. El Sprint es la denominación que recibe cada interacción de desarrollo, es decir, el avance en los tiempos establecidos, y finalmente el incremento es el resultado de cada Sprint.

Además, los eventos que manipula SCRUM están enfocados a las reuniones que se establecen para realizar el proyecto, como la reunión de planificación del Sprint, revisión del Sprint o reunión de Scrum diario, etc. La reunión de planificación de Sprint es al inicio de cada Sprint en la que se determina el objetivo del Sprint y las actividades necesarias para conseguir un visto bueno, también se estima el tiempo de cada tarea y se asignan los recursos. El Scrum diario es una breve reunión en la que se analiza que se realizó el día anterior, que se tiene previsto para hoy y que se puede necesitar o eliminar para efectuar las actividades actuales.

Otro punto importante del método es la herramienta para el monitoreo del avance del Sprint que son los gráficos Burn-Down, el mismo que sirve para monitorear el ritmo de avance y desviaciones en las actividades, es actualizado por el grupo en las reuniones diarias de Sprint. Para esto la reunión de revisión de Sprint realiza la entrega del producto o incremento analizando el cumplimiento del objetivo planteado en el Sprint que finaliza. En última instancia,

se complementa con la reunión de retrospectiva del Sprint en donde se revisa la iteración durante las actividades, donde el grupo analiza factores operativos de la manera de trabajo y elabora una planificación de mejoras para aplicarla en el siguiente Sprint (Palacio, 2011).

3.3.2.- Metodología de la solución de BI

La solución no abarca la creación de módulos dentro de los procesos operacionales del CEC o de sistemas existentes, es en base a estas fuentes de información que se obtendrán mejores decisiones para retroalimentar procesos y realizar transacciones diversas. En este sentido la metodología de Ralph Kimball posee todas las características que se desea, por esta razón esta metodología va acorde al TFM.

A continuación, la tabla 2 muestra la comparación de metodologías más conocidas para la implementación de una solución BI, esto con la finalidad de entender el motivo por el cual se seleccionó la metodología de Ralph Kimball que se utilizó para el desarrollo del DataMart. Para el análisis cuantitativo se utiliza una ponderación con los siguientes criterios cuantitativos:

1: No Aplicable;

2: Casi Aplicable;

3: Aplicable

Tabla 2. Cuadro comparativo de las metodologías de BI

Nº	Factores de análisis	Metodología: Ralph Kimball	P	Metodología: Bill Inmon	P	Metodología: Ricardo Bernabéu (Hefesto)	P
1	Arquitectura	Bottom-Up	3	Top-Down	2	Híbrida	1
2	Énfasis	DataMart	3	Data Warehouse	2	Data Warehouse	2
3	Especificaciones	Está basada en experimentos y prototipos, es más flexible	3	Los problemas que se van a solventar ya son conocidos de antemano	1	Enfocada en el análisis de los requerimientos para identificar indicadores, perspectivas y el análisis de datos	2
4	Enfoque	El enfoque es por las áreas del negocio. Formar DataMart para luego construir el Data Warehouse	3	Tiene un enfoque global de todo el CEC. Formar Data Warehouse para obtener DataMart	1	Tiene un enfoque global de toda el CEC. Formar Data Warehouse para obtener DataMart	1

(Continuación)

Nº	Factores de análisis	Metodología: Ralph Kimball	P	Metodología: Bill Inmon	P	Metodología: Ricardo Bernabéu (Hefesto)	P
5	Construcción de Data Warehouse	Toma menos tiempo	3	Toma mucho tiempo	1	Toma mucho tiempo	1
6	Mantenimiento	Sujeta a constantes revisiones	2	Las revisiones son más sencillas	3	Difícil, a menudo redundante y sujeta a revisiones	1
7	Costos	Implementar cada DataMart por área, permite que la solución no genere un alto costo	3	El Data Warehouse permite que la solución genere un alto costo ya que replica muchos datos	1	Implementar el Data Warehouse permite que la solución genere un alto costo	1
8	Tiempo de desarrollo	Tiempo más corto para el desarrollo de esta metodología	3	Toma mayor tiempo para el desarrollo de esta metodología	1	Toma mayor tiempo para el desarrollo de esta metodología	1
9	Adaptabilidad a la propuesta de desarrollo actual	Alto	3	Bajo	1	Bajo	1
10	Requisitos de integración de datos	Áreas del CEC individuales	3	Todo el CEC	1	Todo el CEC	1
11	Afinidad con el sistema actual en desarrollo	Alto	3	Bajo	1	Bajo	1
12	Fácil entendimiento	Medio	2	Alto	3	Medio	2
13	Rápido acceso a reportes	Alto	3	Bajo	1	Medio	2
TOTAL			31		17		15

Fuente: Elaboración propia

La metodología que mejor se ajusta al propósito del TFM, es la metodología de Ralph Kimball la cual plantea como arquitectura de una solución de soporte de decisiones armar primero DataMart para satisfacer las necesidades específicas de un área dentro del negocio, permitiendo así un mejor control de la información que se está abarcando. El mayor puntaje obtenido de los 39 puntos en juego es de la metodología de Ralph Kimball que cumple con características necesarias para el desarrollo del DataMart del Área Administrativa del CEC.

CAPÍTULO 4. DESARROLLO ESPECÍFICO DE LA CONTRIBUCIÓN

4.1.- Identificación de interesados

.Un interesado es una persona, grupo u organización que de forma directa o indirecta se verá afectado por los resultados que se vayan presentando a lo largo del desarrollo del proyecto. Los interesados también pueden tener influencia sobre el proyecto, entregables y el grupo del proyecto a fin de establecer un conjunto de resultados que satisfagan los objetivos estratégicos del negocio.

Los interesados incluyen a todas las personas que se relacionan con los objetivos del proyecto, sean los trabajadores internos, administradores, grupo de trabajo; en la parte interna, así como también entidades externas. El grupo de trabajo del proyecto identifica a los interesados tanto internos como externos, para determinar los requisitos del proyecto y las expectativas de todas las partes interesadas.

4.1.1.- Interesados externos

Entre los interesados externos están las personas u organizaciones que se relacionan indirectamente con el proyecto, como es el caso de:

Clientes y Usuarios: Los clientes o usuarios son los estudiantes porque es con su información con la que se llenan los campos de los registros de la base de datos y son los que permiten la funcionalidad del DataMart para la toma de decisiones en base a su registro, matrícula, desempeño, etc. Además de los estudiantes, los docentes son los interesados externos del DataMart, porque su participación en la formación permanente de los estudiantes permite la generación de información de calificaciones para el CEC.

Entidades gubernamentales: Las entidades gubernamentales que regulan la calidad de la educación superior en el Ecuador, como es el caso del SENESCYT (Secretaría de Educación Superior, Ciencia, Tecnología e Innovación) o el CACES (Consejo de Aseguramiento de la Calidad de la Educación Superior) que tendrían estadísticas descriptivas por parte del trabajo que realiza el CEC de la Universidad Técnica de Machala.

Entidades no gubernamentales: Las entidades no gubernamentales que firmen convenios con el CEC para mejorar la operatividad de ambas entidades, como por ejemplo organizaciones para pasantías o prácticas técnicas profesionales.

Otros interesados: Otros interesados, tales como entidades contratantes de desarrollo de software, instituciones financieras, expertos en la materia, consultores, etc. entidades que tengan interés en el proyecto o en sus resultados y entidades que contribuyan en la solución.

4.1.2.- Interesados internos

Patrocinador: El patrocinador sirve como vía de escalamiento para los asuntos que están fuera del alcance del director del proyecto. También puede participar en asuntos importantes, como la autorización de cambios en el alcance, revisiones de final de Sprint, y cuando los riesgos son particularmente altos, decidir si el proyecto debe continuar o no; además, garantiza una transferencia eficiente de los entregables del proyecto hacia el negocio tras el cierre del proyecto. Para el caso del diseño del DataMart para el CEC, el patrocinador es el Director del Área Administrativa.

Grupo de la organización: Los grupos de la organización se ven afectados por las actividades del equipo del proyecto. Estos grupos pueden contribuir en la especificación de los requisitos, para el caso son los colaboradores del Área Administrativa del CEC.

Gerentes funcionales: Los gerentes funcionales son personas clave que desempeñan el rol de gestores dentro de un área funcional de la entidad, para el caso son los directores departamentales de Recursos Humanos, Finanzas, Contabilidad, Directores de Carrera y Rectoría de la Universidad Técnica de Machala. Estos interesados cuentan con personal permanente propio asignado para la realización de trabajo en curso y gestionan las tareas dentro de su área funcional de responsabilidad, aporta su experiencia en la materia.

4.1.3.- Interesados de desarrollo

Para la consecución del TFM se contará con la participación de tres entidades denominadas partes interesadas de desarrollo porque influyen directa e indirectamente en el diseño del DataMart (BI) orientado a la toma de decisiones del área Administrativa del Centro de Educación Continua (CEC) de la Universidad Técnica de Machala y cuya información se menciona en la tabla 3. El DataMart necesita la participación tanto de los interesados internos,

externos, para especificar mejores resultados en el diseño del DataMart según las configuraciones de la parte interesada de desarrollo.

Tabla 3. *Personas interesadas del diseño del DataMart (BI)*

Persona	Contacto	Rol
Director	jorge.rivas@gmail.com	System Owner
Administrador	eduardo.hidalgo@gmail.com	Administrator DataMart
Técnico	sally.arreaga@gmail.com	Developer

Fuente: Elaboración propia

Para las personas interesadas y su rol en el DataMart (BI) se realizó un levantamiento del sistema físico actual del CEC, realizando un análisis de los distintos roles, Director, Administrador y Técnico; a continuación se detalla las responsabilidades de los actores involucrados en el DataMart y por ende en el Data Warehouse. La figura 14 muestra los actores del DataMart.

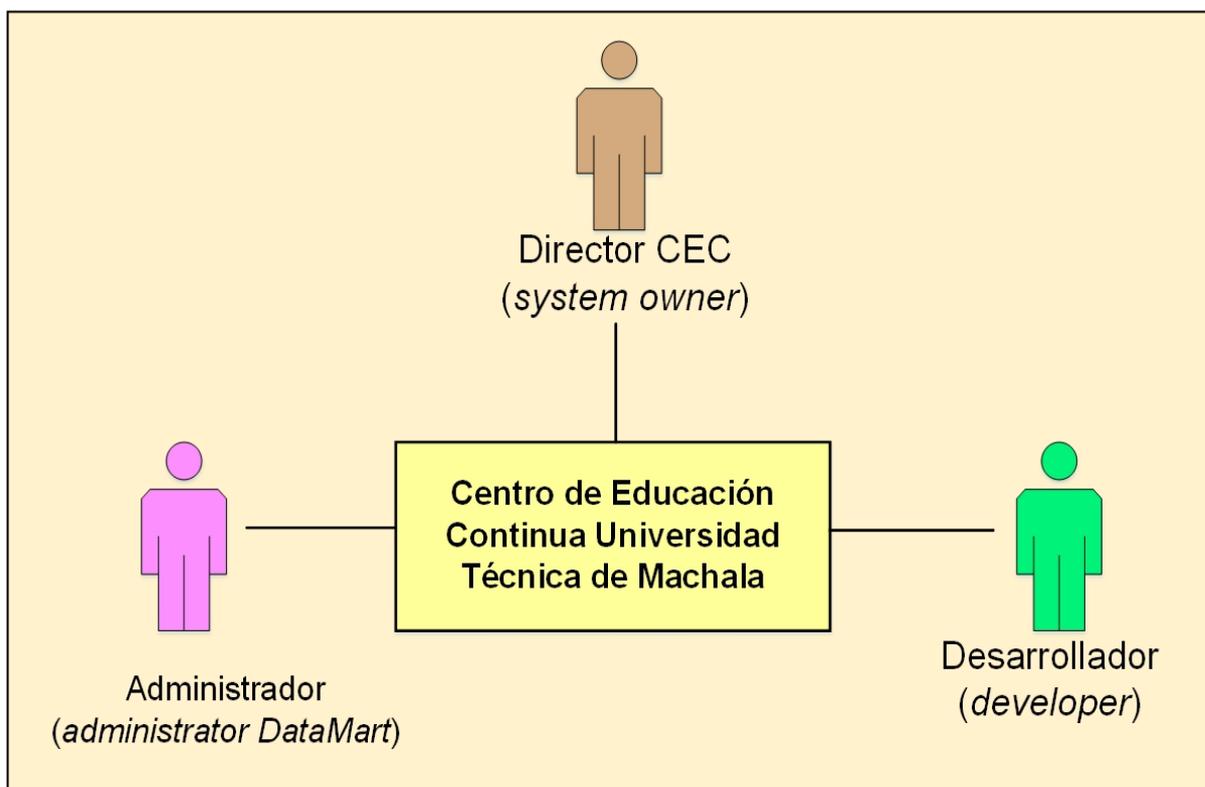


Figura 14. *Personas interesadas del DataMart (BI).* **Fuente:** Elaboración propia

La tabla 4 muestra las funciones del Director, o también denominado para el diseño del DataMart (BI) y por ende del Data Warehouse, como system owner. El Director del CEC es el encargado de gestionar y administrar las funciones para un buen desempeño de este centro

de educación superior, el mismo que procura en sus acciones buscar la mejora continua de la educación de los profesionales que la Universidad Técnica de Machala incorpora a la sociedad ecuatoriana.

Tabla 4. *Funciones system owner*

Actor: System owner
Descripción: Es el encargado de ejecutar las actividades de gestión y administración para el buen funcionamiento del Centro de Educación Continua, realiza las siguientes funciones: <ul style="list-style-type: none">▪ Define las funciones de los puestos de trabajo dentro del CEC.▪ Supervisa el trabajo de los docentes universitarios.▪ Aprueba y regula la inscripción y matrícula de los estudiantes.▪ Supervisa el trabajo pedagógico de los educadores.▪ Elabora junto con el Administrador la malla curricular y planificación del CEC.▪ Administra reuniones con el personal.▪ Administra la parte financiera del CEC.▪ Resuelve problemas en el ámbito académico.▪ Ayuda en la formación de los docentes universitarios.▪ Supervisa la infraestructura del CEC y su buen funcionamiento.▪ Define estrategias a futuro.

Fuente: Elaboración propia

La tabla 5 muestra las funciones del Administrador conocido para propósitos del diseño del DataMart (BI) como administrador DataMart.

Tabla 5. *Funciones Administrator DataMart*

Actor: Administrator DataMart
Descripción: Es el encargado del mantenimiento, control y administración de los recursos informáticos, base de datos, información y conocimiento del CEC para la operatividad de la formación técnica que oferta la entidad, realiza las siguientes funciones: <ul style="list-style-type: none">▪ Supervisa el buen funcionamiento de los servidores de datos.▪ Gestiona la información que almacenan las bases de datos.▪ Controla el sistema informático de inscripción y matrícula.▪ Inspecciona el buen funcionamiento de dispositivos informáticos.

Fuente: Elaboración propia

(Continuación)

Actor: Administrator DataMart
<ul style="list-style-type: none">▪ Organiza la cronología de formación para docentes y estudiantes.▪ Realiza la planificación curricular de la educación continua.▪ Gestiona la promoción de las diferentes capacitaciones que realiza el CEC.▪ Promociona en medios tecnológicos y de comunicación la cronología de formación y capacitaciones en el CEC.▪ Administra lo concerniente al flujo de información y procesamiento de datos.▪ Planifica reuniones de trabajo con el Director y operario para rendición de cuentas y responsabilidades para con el CEC.▪ Interactúa con las Facultades y demás departamentos que conforman el campus de la Universidad Técnica de Machala para la planificación.▪ Socializa los estatutos y estatutos de la Universidad Técnica de Machala.

Fuente: Elaboración propia

La tabla 6 muestra las funciones del Desarrollador, también conocido para propósito del diseño del DataMart (BI) como developer.

Tabla 6. *Funciones Developer*

Actor: Developer
<p>Descripción: Es la persona encargada del diseño, configuración y mantenimiento del DataMart de Inteligencia de negocio, además de la administración general de los metadatos y los Data Warehouse, realizando las siguientes funciones:</p> <ul style="list-style-type: none">▪ Define la configuración de los sistemas DataMart y Data Warehouse.▪ Mantiene operativo las Base de Datos y consultas.▪ Mantiene funcional DataMart para la observación de reportes e informes que permitan la toma de decisiones eficientes.▪ Supervisa el trabajo del administrador del DataMart (BI).▪ Administra los recursos lógicos y reposición de configuraciones del DataMart.▪ Administra la parte de eficiencia en la gestión del conocimiento del CEC.▪ Define estrategias a futuro.▪ Soluciona problemas con respecto a la configuración del DataMart.▪ Proporciona mantenimiento preventivo y correctivo a las interfaces.▪ Realiza mediciones y evaluaciones para la toma de decisiones.▪ Configura nuevos campos, registros y consultas en las Base de Datos.

Fuente: Elaboración propia

Actualmente todo el procesamiento de la información dentro del CEC se lo realiza de manera tradicional, con el control individual de un sistema informático, un control manual que propicia desperdicio de recursos y la existencia de errores operativos, falla de las inscripciones y matriculas, deficiente generación de certificados o pérdidas de información en equipos. En base a este análisis de los roles de las partes interesadas, es preciso establecer dos tipos de usuarios para el diseño del DataMart (BI), los roles se los describe en la tabla 7.

Tabla 7. Tipos y roles de Usuario

TIPO DE USUARIO	ROL
Administrador	<ul style="list-style-type: none">▪ Gestionar el DataMart de Inteligencia de Negocios.▪ Agrega, modifica, elimina y configura información o las configuraciones de entrada, salida, almacenamiento del DataMart (BI).▪ Administra el buen funcionamiento de los Data Warehouse, seguridad informática, gestión de base de datos OLAP y base de datos OLTP.
Usuario	<ul style="list-style-type: none">▪ Mantiene operativo el DataMart y su interfaz.▪ Establece comunicación con el administrador para cualquier dificultad y capacitación.

Fuente: Elaboración propia

4.2.- Definición de requerimientos

Los requerimientos del diseño del DataMart es una especificación de las funciones que tendrá el Data Warehouse. Además, de las necesidades y funciones necesarias, los requisitos describirán con claridad el ambiente operativo en el que se desempeñará el DataMart. Esta es una visión del usuario del instrumento, para el caso el área de Administración del Centro de Educación Continua (CEC) de la Universidad Técnica Machala. El volumen de recolección de requisitos va de acuerdo al método que se utilice para el diseño de la base de datos. Por ejemplo, se utiliza el enfoque de arriba hacia abajo, la actividad de requerimientos es muy extensa, mientras que si se usa el enfoque de abajo hacia arriba, gran parte de los requisitos se definen en base a las expectativas que se tiene del DataMart, y estos requisitos son significativamente menos que el anterior.

Una forma directa de propiciar los requisitos del diseño del DataMart, es por medio de los diferentes puntos de vista de la gente como parte interesada y alta dirección de la Data Warehouse. La figura 15 muestra los requerimientos para el diseño del DataMart, en donde se estipulan las partes importantes que dimensionan la base de datos y sus interesados.

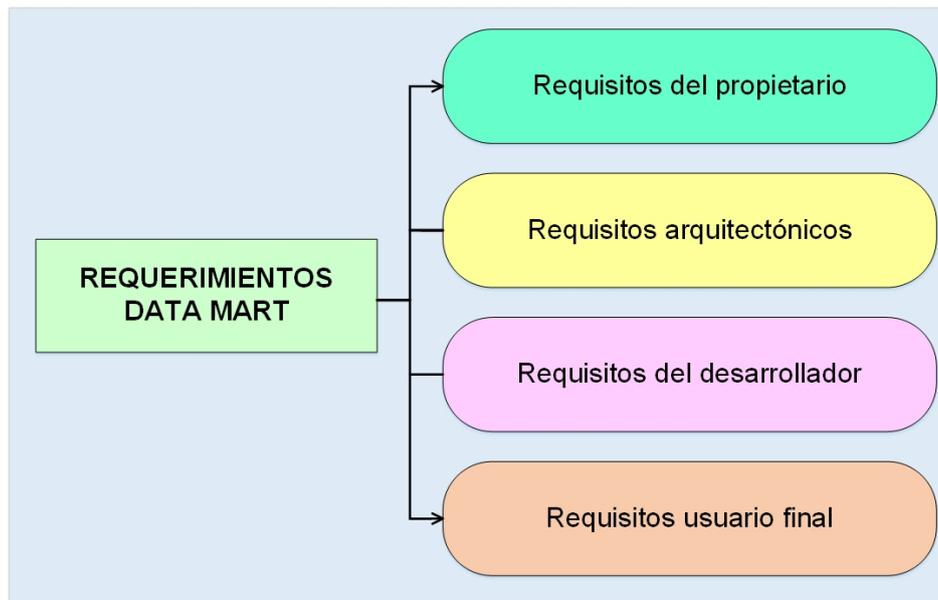


Figura 15. *Requerimientos para el diseño de un DataMart.* Fuente: Ballard (2006)

A continuación se detalla las perspectivas de las personas interesadas en el proyecto de DataMart sobre la definición de sus requisitos:

Requisitos del propietario:

Los propietarios del DataMart estipulan las siguientes interrogantes para su construcción:

- ¿Por qué diseñar un DataMart?
- ¿Qué problema de administración del Centro de Educación Continua CEC abordará?
- ¿Cuáles son los objetivos organizacionales del CEC?
- ¿Qué precio será necesario para poner en práctica el proyecto?
- ¿Qué impacto dispondrá sobre la sociedad?
- ¿Qué impacto tendrá la organización?
- ¿Se dispone de la factibilidad operativa y técnica para realizar el diseño?
- ¿Qué riesgos existen en CEC y el área Administrativa con referencia al diseño del DataMart?

La recolección de requisitos debe permitir contestar las interrogantes, a partir de este resultado, estos requisitos formarán entonces un hallazgo desde la perspectiva del propietario.

En este sentido las áreas potenciales de recolección de requisitos son:

- Objetivos empresariales.

- Ámbito y objetivos del Data Warehouse, DataMart y los usuarios.
- Requerimientos del cliente/servidor
- Requisitos del usuario.
- Fuentes de datos.
- En los distintos planes del CEC, en los calendarios, costes, presupuestos y recursos.
- El impacto sobre las inversiones actuales, como el capital humano, la tecnología y la capacitación.
- Parte de los requisitos organizacionales son también especificaciones del ámbito educativo para la toma de decisiones Data Warehouse, por medio de las siguientes contextualizaciones.

Áreas Tema:

Estos requerimientos son los temas de interés de las distintas funciones organizacionales que cumple con la selección de los departamentos Tema, maximizando la utilidad del DataMart.

Granularidad:

Se refiere al nivel de detalla de la información requerida; la granuralidad tiene una relación directa con las actividades de resumen y adición que deben realizarse sobre los datos fuentes. A menor granularidad mayor cantidad de detalle. Los datos operacionales son considerados como el nivel más bajo de granuralidad. Para incrementar la granuralidad los datos operacionales reunirse y acumularse más con el fin de incrementar su utilidad para la toma de decisiones.

Por lo general, entre mayor sea la granularidad, mayor es el procesamiento para convertir y resumir los datos operacionales. En el mismo sentido los datos con alta granuralidad requieren menos volumen de almacenamiento y pueden ser más rápidamente consultados. El costo del procesamiento es continuo para cada actualización del Data Warehouse con nueva incorporación de información. La figura 16 muestra la pirámide de datos que describe la relación entre:

- El incremento de granuralidad.
- Incremento del computó.
- Disminución del volumen de datos.
- Reporte de la información relevante a la toma de decisiones.
- Gráficas estructuradas y explícitas.
- Estadísticas de avance del proyecto y de tiempo de ejecución.

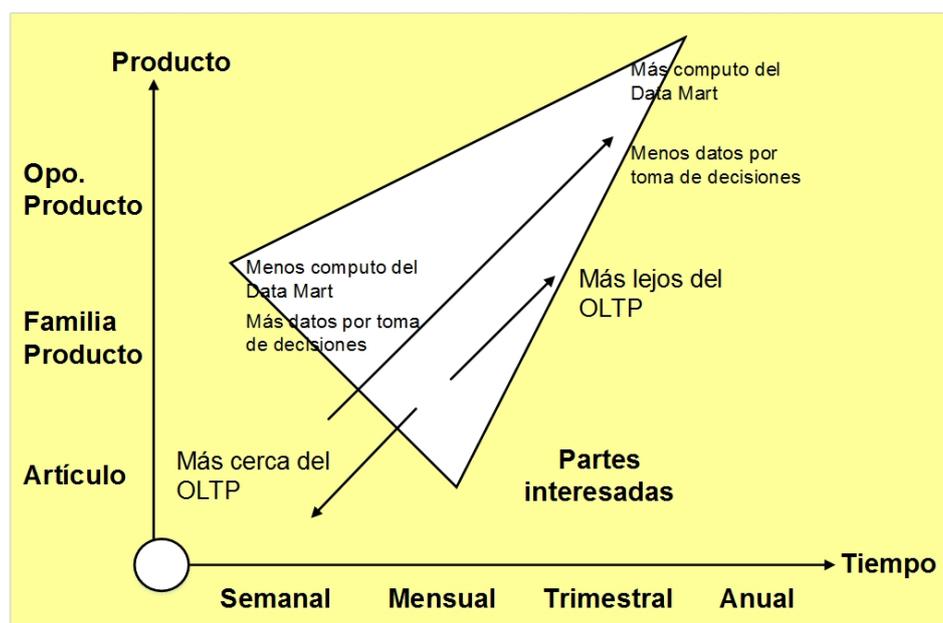


Figura 16. Pirámide de datos que detallan la granularidad. Fuente: Ballard (2006)

Dimensiones:

El Data Warehouse organiza una gran asociación de datos operacionales e históricos mediante múltiples dimensiones de categorización, en este sentido una de las dimensiones más relevantes es el tiempo. A los datos operacionales se les asigna un registro de tiempo en la fuente. Luego el Data Warehouse tiene la capacidad de agrupar todos los datos que sucedieron dentro de un determinado rango de tiempo para responder a una específica solicitud de consulta.

4.3.- Definición de alcance del proyecto

El TFM busca ayudar por medio del diseño de un DataMart de Inteligencia de Negocios para la toma de decisiones en la gestión del área Administrativa del Centro de Educación Continua (CEC) de la Universidad Técnica de Machala, por medio de reportes analíticos con información histórica, veraz y consolidada en tiempo real. Considerando razones como que los informes del CEC están orientados a la administración académica, la estabilidad de los requisitos estudiantiles y la capacidad de entregar informes rápidamente, se ha decidido que la metodología de Kimball sería utilizada para diseñar el DataMart para el CEC.

4.3.1.- Alcance de la solución

Con el rápido crecimiento de la información y datos, con una gran insistencia en comprenderlos para obtener información depurada, los datos históricos deben almacenarse

de forma tal que puedan analizarse rápidamente. También es importante extraer algunas estadísticas importantes y varios KPI del área Administrativa del CEC. Pero esto no se puede implementar de manera efectiva utilizando los modelos de relación entidad de las fuentes de datos operativos (ODS). Los datos deben reorganizarse en un modelo dimensional como lo estipula el método de Kimball, en este sentido la figura 17 muestra un ejemplo de modelo dimensional de un DataMart.

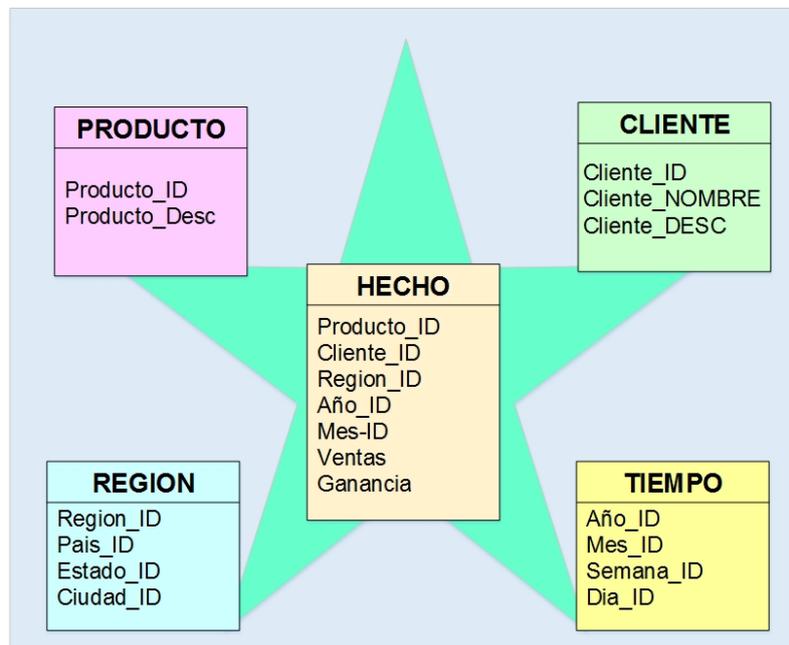


Figura 17. *Ejemplo de modelo dimensional.* Fuente: Ballard (2006)

Según Ballard (2006), el enfoque de modelado dimensional facilita la generación de informes analíticos al mejorar su rendimiento. Los modelos dimensionales proporcionan un rendimiento notablemente mejor que los modelos de relación de entidad (ER), especialmente para consultas grandes. Además, los modelos dimensionales son más fáciles de entender ya que cada modelo está diseñado para un tema definido. Está formado por dos componentes principales, una tabla de registros o también denominada tabla de hechos que contiene las medidas del sujeto y un conjunto de tablas referidas a la tabla de hechos que contienen datos descriptivos sobre las dimensiones de hechos, la figura 16 representa este concepto. La tabla HECHO en el centro, contiene claves foráneas para otras tablas de dimensiones que son PRODUCTO, CLIENTE, REGIÓN y TIEMPO; y también tiene datos de Ventas y Beneficios. Finalmente, la tabla de hechos y las dimensiones forman un esquema de estrella como se puede apreciar en el ejemplo de la figura antes mencionada.

Los beneficios de la técnica de modelado dimensional se pueden resumir de la siguiente manera:

Mejor navegación y presentación de datos: Los datos se modelan de una manera que es fácil de entender incluso para usuarios comerciales. Esto les permite diseñar fácilmente sus propios informes.

Mantenimiento fácil y de bajo costo: Los datos se almacenan de la misma manera en que se presentan, a diferencia del caso de las bases de datos relacionales, donde en la mayoría de los casos se crean vistas para generar informes. Esto aumenta el costo de mantenimiento en el caso de bases de datos relacionales u ODS.

Mejor rendimiento: La mayoría de los informes requieren datos resumidos que resultan en un rendimiento más lento debido a los cálculos sobre la marcha en el caso de vistas no indexadas en ODS. En el caso del modelado dimensional, las tablas resumidas se crean según sea necesario. También permite la capacidad de almacenar el historial de datos de una manera fácil de consultar y generar informes. Dicho diseño ofrece un rendimiento de consulta más rápido y para profundizar en las jerarquías de datos.

Por otro lado, en el modelo ER las tablas se normalizan en una forma donde no hay redundancia de datos llamada Tercera forma Normal (3NF). Las tablas en 3NF tienen atributos donde todas están relacionadas con una clave primaria en la entidad. Una clave primaria puede ser solo un atributo o una clave compuesta que consta de dos o más atributos.

Un tercer modelo ER de forma normal es un buen diseño para manejar datos transaccionales.

Ejemplos de datos transaccionales son los cursos tomados por los estudiantes en cada trimestre, la admisión de los estudiantes, etc. Este diseño es bueno para realizar transacciones rápidas, como inserciones, actualizaciones y eliminaciones, ya que las tablas generalmente tienen una pequeña cantidad de campos con claves foráneas para otras tablas. Este diseño es ineficiente para el propósito de informar, porque para generar un informe simple, uno terminaría uniendo puntajes de tablas en un entorno 3NF que termina siendo muy intensivo en tiempo y recursos (hardware). Aparte de esto, el modelo se vuelve rápidamente muy complejo como resultado de lo cual se hace muy difícil de entender y navegar el modelo incluso para un desarrollador, y mucho menos para un usuario común.

4.3.2.- Exclusiones

Entre las principales restricciones del diseño del DataMart (BI) para el área Administrativa del Centro de Educación Continua (CEC) de la Universidad Técnica de Machala, se encuentra que:

- El costo del diseño del DataMart (BI) no debe exceder los 30.000 € debido a las capacidades financieras del administrador del CEC o system owner.
- El desarrollo del TFM y el diseño del DataMart (BI) no deben exceder los seis meses.
- La planificación para la implantación del DataMart (BI) debe estipularse en un orden cronológico de 12 meses para su cierre técnico y administrativo.
- El proyecto debe ejecutarse entre la vigencia del año 2020 y 2021.
- La plataforma tecnológica utilizada para el DataMart (BI) debe cumplir con todos los requisitos establecidos para el almacenamiento de información y gestión del conocimiento del CEC, debe existir un sistema que resguarde y proporcione seguridad informática para el CEC.

En ramificación a las restricciones del proyecto, se establecen los problemas más comunes con el DataMart (BI) que surgen debido a la falta de atención al flujo y generación de información en el CEC, la falta de mantenimiento y optimización de las base de datos y el sistema informático, la falta de conocimiento entre los directivos y administradores y los problemas técnicos debido a la complejidad de los sistemas de Inteligencia de Negocios (BI), así como las herramientas para gestionar Data Warehouse y los propios DataMart.

4.3.3.- Supuestos

Se cuenta con la disposición de las partes interesadas en el diseño del DataMart (BI), en especial del system owner y el administrador DataMart, para las actividades desarrolladas en la cronología del TFM.

Todas las decisiones técnicas y operativas están sometidas al diálogo entre las personas interesadas en la planificación del TFM y el diseño del DataMart (BI) junto con el Data Warehouse, teniendo en cuenta las recomendaciones internacionales y nacionales que puedan aportar al proyecto.

Las partes interesadas en la planificación del proyecto conocen y están conscientes de los requerimientos para el desarrollo del DataMart (BI), además aprueban la documentación del mismo en favor de conseguir un Centro de Educación Continua (CEC) equilibrado a la calidad de la educación superior.

La empresa o contratista destinada al desarrollo del software, la base de datos y las interfaces de Data Warehouse cuenta con la experiencia necesaria en el desarrollo de proyectos

tecnológicos bajo las normativas y metodologías de gestión planificadas para el manejo de Big Data y herramientas de Inteligencia de Negocios.

La iniciativa de reporte de datos de manera abierta, es una iniciativa estratégica para el cumplimiento de las políticas de la calidad en la educación superior dentro del país, en lo que respecta al derecho del acceso a la información libre y transparente.

4.3.4.- Restricciones

Entre las principales restricciones del diseño del DataMart para la toma de decisiones del CEC de la Universidad Técnica de Machala, se encuentra que:

- La cantidad de tiempo disponible para completar la solución será limitado, debido a que solo una persona desarrolla el proyecto.
- Por motivos de políticas internas de seguridad de información del CEC de la Universidad Técnica de Machala, los datos que se utilizarán han sido tomados desde el servidor, siendo la información necesaria para el diseño de la solución.

4.4.- Plan de trabajo

Este epígrafe se centra en el proceso de desarrollo específico del DataMart para el área Administrativa del Centro de Educación Continua (CEC) de la Universidad Técnica de Machala. Además, expone el enfoque seguido en el proceso de diseño del DataMart. En este sentido, el proceso de almacenamiento de los datos está basado en cuatro fases principales: análisis del sistema, diseño del sistema, desarrollo del sistema, verificación y mantenimiento del sistema. Se utilizará dos tipos de metodologías una de gestión de proyectos PMBOK y otra de inteligencia de negocios como es la metodología Kimball.

4.4.1.- Paquete de trabajo 1: Gestión del proyecto

Se define como el uso del conocimiento, herramientas y habilidades en las tareas del proyecto, para así progresar en el desarrollo final y obtener los objetivos del proyecto. Esto implica identificar los requisitos del proyecto y establecer parámetros medibles para que sean verificados al final del proyecto.

Para poder realizar la gestión del proyecto el Administrador del proyecto tiene que ser capaz de gestionar las restricciones del proyecto, alcance, costo, tiempo, calidad, recursos, riesgos e incluso la satisfacción del cliente; estas pueden estar relacionadas entre sí, ya que al variar

alguna dimensión puede afectar las restricciones, por ejemplo, si se quiere reducir el tiempo que dura el proyecto es muy posible que el costo se incremente. O si desea reducir el costo y tiempo se tiene que reducir la calidad, pero consecuentemente la satisfacción del cliente también se ve afectada.

El Administrador del proyecto debe ser capaz de analizar la situación con respecto a las restricciones y balancearlas a lo largo del proyecto, a fin de conseguir un equilibrio que ayude a llegar a un éxito al final del proyecto. La figura 18 muestra el triple de restricciones en la gestión de proyectos.

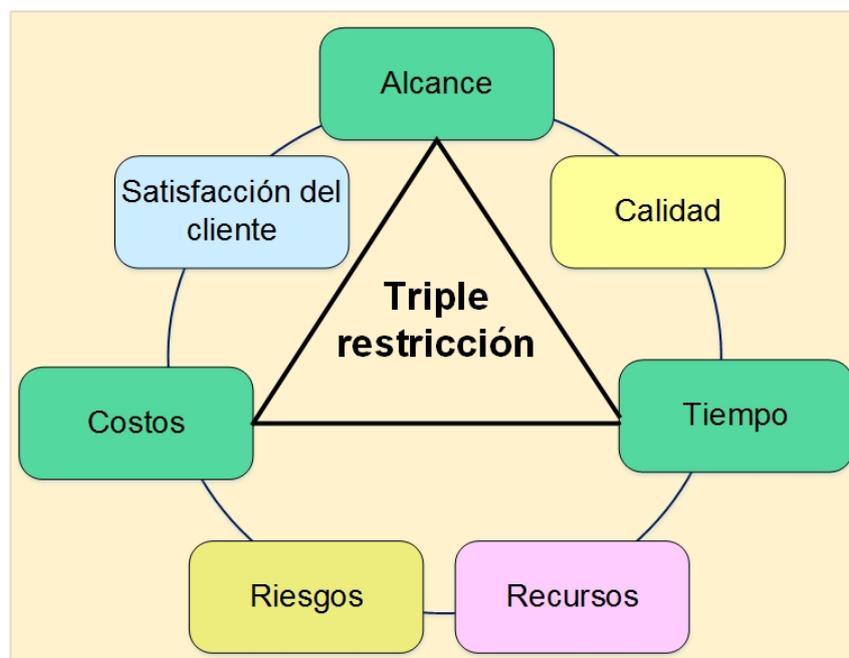


Figura 18. Triple restricción. Fuente: Elaboración propia basado en PMBOK V5.0

Según la metodología PMBOK para la gestión de proyectos, los objetivos deben:

- Los objetivos deben ser recogidos en un documento denominado “Acta de Constitución”.
- Un proyecto se considera completo cuando se han alcanzado los objetivos, razón por la cual tienen que ser bien definidos de manera que sean medibles porque esto permitirá identificar si fueron completos y poder dar el proyecto por cerrado o no.
- Si se ve que los objetivos son imposibles de cumplir, puede ser razón suficiente para terminar un proyecto de manera anticipada, o bien el proyecto pueda necesitar ser redefinido con el efecto de alcanzar dichos objetivos.
- Dicho lo anterior, toda la responsabilidad de que estos objetivos sean cumplidos recae en el Director del Proyecto.

1.- ACTA DE CONSTITUCIÓN DEL PROYECTO

Código de documento:	CEC-UTM-DM-01-2020		
Fecha: _____	Versión: V 1.0		
Proyecto:	Diseño del DataMart		
Unidad:	Área Administrativa del CEC		
Patrocinador:	_____		
Fecha inicio:	_____	Fecha final:	_____

Objetivos del proyecto
<p>Objetivo general:</p> <p>Diseñar un DataMart de inteligencia de negocios, que permita al Área Administrativa del Centro de Educación Continua de la Universidad Técnica de Machala incrementar los niveles de confianza para la toma de decisiones.</p>
<p>Objetivos específicos:</p> <ol style="list-style-type: none"> 1.- Establecer nuevos mecanismos integrales de toma de decisiones. 2.- Generar reportes y alertas con un DataMart intuitivo y de fácil usabilidad. 3.- Aumentar el conocimiento e inteligencia organizacional del CEC.
Justificación del Proyecto
<p>La idea de diseñar un DataMart es para fortalecer el sistema de reportes y toma de decisiones, con que actualmente cuenta el CEC de la Universidad Técnica de Machala, contar con una alerta y control de la información a manera de inducción al Big Data, dado que en algunas ocasiones se presentan dificultades para acceder a la información.</p> <p>Adicionalmente, el patrocinador del proyecto desea que el Área Administrativa cambie su enfoque a uno más orientado a la utilización de Inteligencia de Negocios (BI) de forma funcional, lo que acelera el ritmo de avance de la inteligencia organizacional, la necesidad a resolver por lo tanto, es la forma de gestionar la información, y como esta, se dispone más a las necesidades de la educación superior.</p>
Descripción del entregable final del proyecto
<p>DataMart para la inteligencia de negocios y toma de decisiones que complemente el sistema informático del Área Administrativa del CEC. Dicha herramienta debe ser sin costo de propiedad intelectual o licencia, e intuitiva para el usuario final</p>

(Continuación)

(Continuación)

Principales riesgos	
1.- Ausencia de claridad en la necesidad actual de gestión de la información y la inteligencia organizacional del CEC y su evolución en mediano y largo plazo.	
2.- Deficiente formación en el uso de tecnologías o en la utilidad de estas, por parte de los funcionarios del Área Administrativa del CEC.	
3.- Limitado uso de las herramientas informáticas de inteligencia de negocios, por parte de los usuarios finales para la toma de decisiones y generación de reportes.	
Requisitos de alto nivel	
La solución logre hacer conexión a fuentes de base de datos históricas del CEC.	
Mapa de hitos	
Hito	Fecha de vencimiento
Aprobación del Acta de Constitución	19 noviembre 2019
Aprobación de la Declaración del Alcance	19 noviembre 2019
Aprobación Planes de Gestión complementarios	13 enero 2020
Aprobación Plantilla de Proyecto	21 enero 2020
Presentación de Informes de avance	21 febrero 2020
	06 marzo 2020
Aprobación de Informe de Prueba	14 marzo 2020
Firma de Acta de Cierre de Proyecto	21 marzo 2020
DataMart configurado	21 marzo 2020
Presentación del Informe final de Proyecto	29 marzo 2020
REALIZADO POR:	APROBADO POR:

2.- DECLARACIÓN DEL ALCANCE

Código de documento:	CEC-UTM-DM-02-2020	
Fecha: noviembre 2019	Versión: V 1.0	
Proyecto:	Diseño del DataMart	
Unidad:	Área Administrativa del CEC	
Patrocinador:	_____	

Declaración del alcance	
<p>1.- Implementar nuevas medidas de control, monitoreo y reporte, por medio de la creación de indicadores, índices y reportes de distintas periodicidades y nivel de detalle de la información.</p> <p>2.- Continuar con el proceso de evolución de arquitectura de la inteligencia de negocio del Área Administrativa por medio del modelo operacional de un DataMart para la toma de decisiones.</p> <p>3. Aumentar el conocimiento del Área Administrativa en las dimensiones de la inteligencia de negocios, así como la administración de proyectos tecnológicos.</p>	
Criterios de aceptación del producto	
<p>1.- Técnicos: La solución debe ser vinculante con las distintas fuentes de datos.</p> <p>2.- De calidad: El nivel de satisfacción, debe cubrir por lo menos el 80% a criterio del cliente o usuario.</p> <p>3.- Administrativos: Los entregables debe ser aprobados por el patrocinador del proyecto y los cambios deben discutirse entre el patrocinador, el Administrador del proyecto y algún consultor externo si lo existiese.</p> <p>4.- Comerciales: Se necesita cumplir lo estipulado en algún contrato de licencias de software, si fuese necesario.</p>	
Entregables del proyecto	
Proceso del proyecto	Entregable
Gestión del proyecto	<p>1.- Acta de constitución del proyecto</p> <p>2.- Declaración del alcance</p> <p>3.- Plan de Gestión del proyecto</p> <p>4.- Plantillas del proyecto</p> <p>5.- Acta de cierre del proyecto</p>

(Continuación)

(Continuación)

Informes	1.- Informes de avance 2.- Informe de prueba 3.- Informe final
Exclusiones del proyecto	
El proyecto abarca una parte preliminar que es la presentación de las herramientas de inteligencia de negocios, que cumplen los requisitos, originalmente, solicitados por el patrocinador. También, brinda el plan para la dirección del proyecto desde su fase de inicio y hasta el cierre para el diseño del DataMart, pero no cubre programas de capacitaciones.	

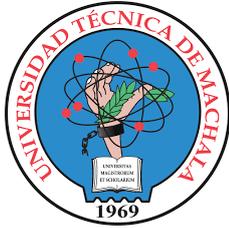
3.- REGISTRO DE INTERESADOS

Código de documento:	CEC-UTM-DM-03-2020	
Fecha: noviembre 2019	Versión: V 1.0	
Proyecto:	Diseño del DataMart	
Unidad:	Área Administrativa del CEC	
Autora:	Sally Astrid Arreaga Santacruz	
Patrocinador:	_____	

Nombre	Rol en el proyecto	Expectativas	Grado de interés e influencia	Tipo
Jorge Rivas	System Owner (Director CEC – Líder de proyecto)	DataMart funcional y que permita un fácil acceso a la información	Bastante favorable - Alto	Interno
Eduardo Hidalgo	Administrator DataMart (Equipo proyecto)	Fácil integración con bases de datos y con las herramientas tecnológicas del CEC	Favorable - Medio	Interno
Sally Arreaga	Developer (Administrador o Director del proyecto)	Se tomen mejores decisiones y generar mayor productividad en el Área y contribuir a la operatividad del CEC	Bastante favorable - Alto	Interno
César Quezada	Rector UTM - Patrocinador	Contribuir en el desarrollo del CEC y la formación continua	Favorable - Alto	Interno
SENESCYT	Asesor - Fiscalizador	Eficiencia y calidad en la educación superior	Neutral - Bajo	Externo
REALIZADO POR:		APROBADO POR:		

4.- PLAN DE GESTIÓN DEL PROYECTO

4.1.- PLAN DE GESTIÓN DE LA INTEGRACIÓN

Código de documento:	CEC-UTM-DM-04-2020	
Fecha: noviembre 2019	Versión: V 1.0	
Proyecto:	Diseño del DataMart	
Unidad:	Área Administrativa del CEC	
Autora:	Sally Astrid Arreaga Santacruz	
Patrocinador:	_____	

Roles de la Gestión de la Integración			
Rol	Perona asignada	Responsabilidades	Nivel autoridad
Director del proyecto	Sally Arreaga	Inspeccionar el funcionamiento de la Gestión d Integración	Autoridad plena sobre el proyecto y sus funciones. Sin embargo, responde al patrocinador
Gestor de la integración	Sally Arreaga	Ejecuta las tareas de la gestión de la integración	Autoridad de ejecutar los procesos necesarios de la integración
Miembros de equipo	Varios	Verificar la información de la gestión de la integración según las labores asignadas en el proyecto	Dependerá de las funciones asignadas a cada uno de los miembros en cada plan de gestión y entregable

Plan de documentación						
Document o	Formato	Acceso	Disponibili dad ampliada	Segurida d de acceso	Recuperaci ón de información	Retención de informaci ón
Acta de constitució n	Electrónico y papel	Intranet	Institucional	Solo lectura	Copia de seguridad electrónica	Duración del proyecto

(Continuación)

(Continuación)

Plan de dirección del proyecto	Electrónico y papel	Intranet	Interesados del proyecto	Solo lectura	Copia de seguridad electrónica	Duración del proyecto
Solicitud de cambios	Electrónico y papel	Intranet	Interesados del proyecto	Solo lectura	Copia de seguridad electrónica	Duración del proyecto
Otros cambios	Electrónico y papel	Intranet	Interesados del proyecto	Solo lectura	Copia de seguridad electrónica	Duración del proyecto
Cierre de proyecto	Electrónico y papel	Intranet	Interesados del proyecto	Solo lectura	Copia de seguridad electrónica	Duración del proyecto

Ítems de integración					
Código	Nombre	Tipo de ítem	Fuente	Formato	Observaciones
	Plan de dirección del proyecto	Documento	Director del Proyecto	Copia impresa y archivo PDF	En su última versión y con toda las aprobaciones necesarias para validar el proyecto
	Informes de sesión bisemanal	Documento	Director del proyecto	Copia impresa y archivo PDF	Con las aprobaciones
	Informe final	Documento	Director del proyecto	Copia impresa y archivo PDF	Firmado y aprobado

Gestión del cambio:

Roles gestión de cambios			
Rol	Persona asignada	Responsabilidades	Nivel de autoridad
Patrocinador	César Quezada	Dar el visto bueno final a las recomendaciones de solicitudes de cambio	Absoluta hacia el proyecto

(Continuación)

(Continuación)

Director del proyecto	Sally Arreaga	Evaluar el impacto de las solicitudes de cambio y recopilar información sobre el impacto del proyecto	Llevar el control sobre las solicitudes de cambio y el impacto de estos en el proyecto
Comité de control de cambios	Dirección del Proyecto Patrocinador Equipo	Decidir, con base en los intereses y beneficios del proyecto, el resultado de las solicitudes de cambio	Aprueban, rechazan o modifican las solicitudes de cambio
Interesados	Los estipulados en el registro de interesados	Solicitar cambios según lo consideren pertinente	Se limita a solicitar cambios en el proyecto sobre cualquier novedad

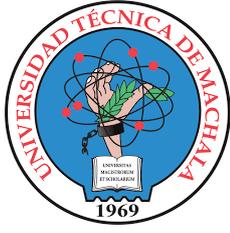
Tipos de cambios	
<p>1.- Cambio preventivo: Puede ser aprobado por el Director del Proyecto sin el proceso de gestión de cambios, siempre y cuando este no afecte las líneas base y quede documentado en el Control de Cambios del Proyecto.</p> <p>2.- Cambio correctivo: Puede ser aprobado por el Director del Proyecto sin seguir el proceso de gestión de cambios, siempre y cuando no afecte las líneas base y que fundamentado en el control de cambios del proyecto.</p> <p>3.- Cambio rectificación: Puede ser aprobado por el inspector de calidad sin seguir el proceso de gestión de cambios, siempre que no afecte las líneas base y quede fundamentado en el control de cambios del proyecto.</p> <p>4.- Cambio al Plan del Proyecto: Debe pasar por el proceso de gestión de cambios, y es el comité de control de cambios el que toma la decisión final sobre la solicitud.</p>	
Proceso de Gestión de cambios	
Solicitud de cambios	<p>1.- Está a cargo del Administrador del proyecto</p> <p>2.- El Director del proyecto es el encargado de recibir la solicitud y verificar que éste completa de la forma correcta</p> <p>3.- De ser necesario puede pedir una reunión con el interesado para conseguir más información</p>
Evaluación de cambios	<p>1.- Se debe verificar el tipo de cambio que se solicita</p> <p>2.- Recopilar información que se amerite para evaluar el impacto integral del cambio hacia el proyecto</p> <p>3.- Describir los resultados posibles en la solicitud de cambio</p> <p>4.- Pasar la solicitud al Comité de Control de Cambios.</p>

(Continuación)

(Continuación)

<p>Replanteamiento del proceso</p>	<p>1.- Corresponde al Comité de Control de Cambios 2.- Evalúa la información proporcionada por el Director del proyecto 3.- Toma la decisión de aprobar, rechazar, o cambiar la Solicitud de cambios.</p>
<p>Implementación del cambio</p>	<p>1.- El Director del proyecto replantea de nuevo el proyecto con la incorporación del cambio aprobado 2.- Se comunica el nuevo planteamiento a los interesados. 3.- Se debe coordinar con el equipo la nueva versión del plan modificado. 4.- De presentarse algún problema con el cambio este se comunica al Director del proyecto</p>
<p>Cierre proceso de cambio</p>	<p>1.- Está a cargo del Director del proyecto 2.- Se verifica la correcta implementación y adaptación del cambio 3.- Actualización a los documentos necesarios tras el cambio</p>
<p align="center">Plan de contingencia ante Solicitudes de cambio catalogadas como urgentes</p>	
<p>De presentarse alguna solicitud de cambio urgente, todo el proceso de Gestión de cambio le corresponde al Director del proyecto, excepto el replanteamiento del proceso, pues todo cambio debe ser aprobado por el Comité de Control de Cambios siempre deben tomar la decisión final.</p>	
<p align="center">Verificación y auditorías de configuración</p>	
<p>Las auditorías a la gestión de integración se realizarán cada semana por parte del Patrocinador y este debe garantizar la entereza de la información que otorga sustento al ítem de integración.</p>	
<p>REALIZADO POR:</p>	<p>APROBADO POR:</p>

Estado de cambios del proyecto:

Código de documento:	CEC-UTM-DM-05-2020	
Fecha: noviembre 2019	Versión: V 1.0	
Proyecto:	Diseño del DataMart	
Unidad:	Área Administrativa del CEC	
Autora:	Sally Astrid Arreaga Santacruz	
Patrocinador:	_____	

ID del cambio:	
Tipo de cambio	
Solicitante	
Fecha de solicitud	
Estado	
Fecha revisión	
Fecha próximo seguimiento	
Observaciones adicionales	
REALIZADO POR:	APROBADO POR:

4.2.- PLAN DE GESTIÓN DEL ALCANCE

Código de documento:	CEC-UTM-DM-06-2020	
Fecha: noviembre 2019	Versión: V 1.0	
Proyecto:	Diseño del DataMart	
Unidad:	Área Administrativa del CEC	
Autora:	Sally Astrid Arreaga Santacruz	
Patrocinador:	_____	

Proceso de definición del Alcance
<p>La definición del proceso del Alcance del proyecto de diseño del DataMart para la toma de decisiones y la inteligencia de negocios, se basa en reuniones y análisis de las necesidades de mejora en los procesos de monitoreo, control y reporteo; lo que partirá de la visión de cambio del Patrocinador de proyecto y del soporte del Área Administrativa. La declaración del alcance es lo que sienta la base de lo que proporcionará este proyecto.</p>
Proceso de elaboración de la EDT
<p>La EDT del proyecto se realiza por medio de la técnica de descomposición, a partir de los entregables del proyecto, los cuales marcan las etapas del mismo.</p>
Proceso para la elaboración del diccionario EDT
<p>Una vez aprobada la EDT, se elabora su diccionario, el cual toma como referencia los paquetes de trabajo y, se centra en el objetivo de cada uno de ellos; describiéndose brevemente y, asignando las responsabilidades.</p>
Proceso de verificación del Alcance
<p>Cada uno de los entregables tiene que ser presentados al Patrocinador, en una reunión con el director del proyecto. Una vez que este sea aprobado, se hará el comunicado oficial a los interesados.</p>
Proceso para control del Alcance
<p>La aceptación de cada uno de los entregables depende del Patrocinador y del Director del proyecto, por lo que corroborarán que se cumpla con la línea base del alcance. Si no se acepta, se deben de enviar las observaciones, correcciones y mejoras junto con la Solicitud de cambios. Por el contrario, si el entregable es aceptado se debe constar en el Acta de aceptación de entregables.</p>

Acta de aceptación de entregables:

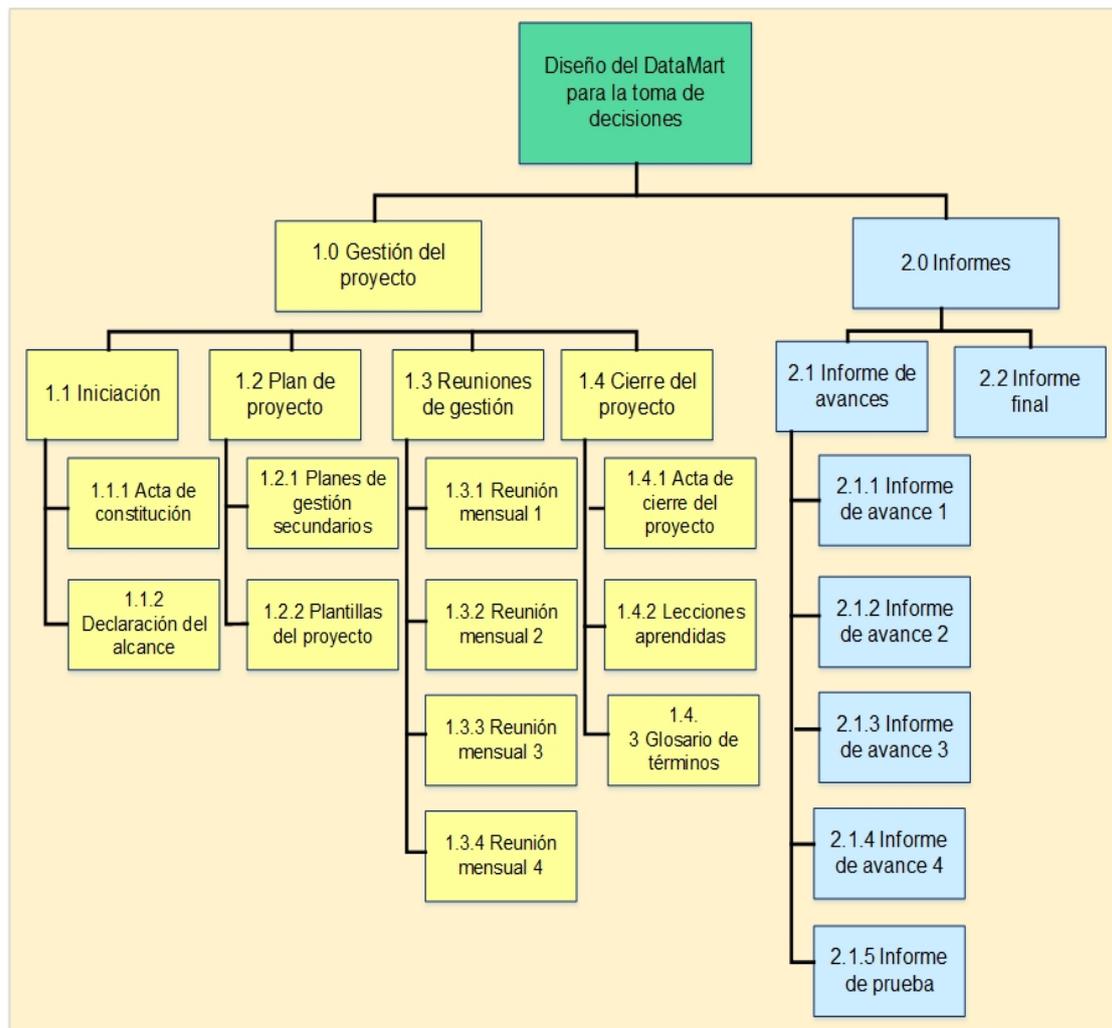
Código de documento:	CEC-UTM-DM-07-2020	
Fecha: noviembre 2019	Versión: V 1.0	
Proyecto:	Diseño del DataMart	
Unidad:	Área Administrativa del CEC	
Autora:	Sally Astrid Arreaga Santacruz	
Patrocinador:	_____	

Generalidades del entregable		
Entregable		
Versión		
Fecha cronograma		
Fecha de verificación		
Fecha de entrega real		
Descripción del entregable		
Observaciones		
	Nombre	Firma
ENTREGADO POR:		
ACEPTADO POR:		

Estructura de desglose de trabajo EDT:

Código de documento:	CEC-UTM-DM-08-2020	
Fecha: noviembre 2019	Versión: V 1.0	
Proyecto:	Diseño del DataMart	
Unidad:	Área Administrativa del CEC	
Autora:	Sally Astrid Arreaga Santacruz	
Patrocinador:	_____	

ESTRUCTURA DE DESGLOSE DE TRABAJO (EDT)



Diccionario de la Estructura de desglose de trabajo - EDT:

Código de documento:	CEC-UTM-DM-09-2020	
Fecha: noviembre 2019	Versión: V 1.0	
Proyecto:	Diseño del DataMart	
Unidad:	Área Administrativa del CEC	
Autora:	Sally Astrid Arreaga Santacruz	
Patrocinador:	_____	

Especificaciones de paquetes de trabajo			
1.0 Gestión del proyecto	1.1 Iniciación	1.1.1 Acta de constitución	Documento emitido por el Patrocinador que formaliza el inicio del proyecto, que contiene los objetivo, justificación, descripción del entregable, riesgos, hitos e interesados
		1.1.2 Declaración del alcance	Documento que contiene la descripción minuciosa del entregable final. Comprende la declaración, criterios de aceptación, entregables, exclusiones y restricciones.
	1.2 Plan del proyecto	1.2.1 Planes de gestión	Herramientas para iniciar, planificar, ejecutar, monitorear y controlar, además de cerrar el proyecto. En este proyecto se desarrollarán los planes de integración, alcance, requisitos, tiempo, costos, calidad, recursos humanos, comunicaciones, riesgos, adquisiciones e interesados.
		1.2.2 Plantillas de proyecto	Documentos que contienen un formato o esquema preestablecido de información que se debe recopilar, para el caso se presentan las plantillas de control de cambios, hoja de correcciones de entregables, acta de aceptación de entregables, solicitud de cambios, roles, monitoreo de riesgos.
	1.3 Reuniones de gestión bisemanal		Reuniones cada dos semanas en las que deben presentarse el Patrocinador, Director del proyecto y el Equipo. En estas se ven los avances del proyecto, los cambios solicitados, y los acontecimientos presentados que necesiten ser resueltos por consenso.

(Continuación)

(Continuación)

Especificaciones de paquetes de trabajo			
	1.4 Cierre de proyecto	1.4.1 Acta de entrega de proyecto	Es un documento que dispone la conclusión del proyecto y la conformidad con los resultados.
		1.4.2 Lecciones aprendidas	Conocimiento que se ha adquirido gracias a la ejecución del proyecto.
		1.4.3 Glosario de términos	Definiciones de conceptos clave y técnicos que se utilizan en la trayectoria del proyecto
2.0 Informes	2.1 Informes de avance	2.1.1 Informes de avance bisemanales	En estos documentos deben presentarse todas las medidas de monitoreo y control de las líneas base, los cambios y cualquier detalle del proyecto que haya sido visto en las reuniones bisemanales
		2.1.5 Informe de prueba	Documento que presenta información sobre el rendimiento del DataMart, las opiniones de los usuarios finales, alguna dificultad, etc.
	2.2 Informe final		Documento que incluye un breve resumen de las actividades desarrolladas, resultados obtenidos, así como el origen del proyecto.

Las demás planificaciones que estipula la metodología PMBOK, se encuentran en los epígrafes correspondientes a los tratados en la estructura del TFM.

4.4.2.- Paquete de trabajo 2: Análisis y especificación de requisitos

4.3.- PLAN DE GESTIÓN DE LOS REQUISITOS

Código de documento:	CEC-UTM-DM-10-2020	
Fecha: noviembre 2019	Versión: V 1.0	
Proyecto:	Diseño del DataMart	
Unidad:	Área Administrativa del CEC	
Autora:	Sally Astrid Arreaga Santacruz	
Patrocinador:	_____	

Actividades de los requisitos
Los requisitos se identificaron por medio de reuniones, entre los interesados, el Patrocinador; a partir de los intereses, las necesidades y el conocimiento previo en inteligencia de negocios.
Actividades de gestión de configuración
En caso de necesitarse cambiar los requisitos del proyecto, el interesado planteará en cambio formalmente, por medio de la Solicitud de Cambios, donde se detalla toda la información que debe proporcionarse. Queda a criterio del equipo, la evaluación del impacto de la solicitud, así como su aprobación o rechazo. Si el cambio es acogido, este se debe comunicar a los interesados por el medio pertinente, y se incorporarán las modificaciones.
Priorización de requisitos
En la matriz de trazabilidad se encuentra la priorización de los requisitos, la cual se basa en el análisis de riesgos, las consecuencias y repercusiones que se tendría sobre el proyecto, el incumplimiento de cada uno de los requisitos; así como, en la dificultad de corregir la ausencia de este requerimiento en el proyecto. Una vez hecha la priorización, es necesaria la aprobación del Director del proyecto.
Métricas del producto
Para efectos del grado de satisfacción de la solución, esta debe ser igual o superior al 80%, dando una puntuación a la priorización de los requisitos de: 1.- Muy alta: (5) 2.- Alta: (4) 3.- Media: (3) 4.- Baja: (2) 5.- Muy baja: (1)
Estructura de trazabilidad
La matriz de trazabilidad recoge información en dos sentidos: 1.- Información de requisitos: código, descripción, estado actual, nivel de estabilidad, grado de complejidad, prioridad. 2.- Relación de trazabilidad: oportunidad, objetivo del proyecto, entregable de la EDT.

Documentación de los requisitos:

Código de documento:	CEC-UTM-DM-11-2020	
Fecha: noviembre 2019	Versión: V 1.0	
Proyecto:	Diseño del DataMart	
Unidad:	Área Administrativa del CEC	
Autora:	Sally Astrid Arreaga Santacruz	
Patrocinador:	_____	

Necesidad de la organización			
1.- Establecer sistemas de control y monitoreo de las actividades del proyecto			
2.- Disponer de mayor información relevante, relacionada al DataMart			
3.- Mejorar los tiempos y la calidad de respuesta ante la solicitud de información por parte del Patrocinador.			
Objetivo del proyecto			
Implementar un DataMart para mejorar la toma de decisiones y la inteligencia de negocios del CEC de la Universidad Técnica de Machala.			
Requisitos funcionales			
Interesado	Prioridad interesado	Requisitos	
		Códigos	Descripción
CEC de la UTM	Muy alta	RF01	DataMart debe ser de libre de licencias
	Media	RF02	El DataMart debe ser de código libre, que pueda modificarse y adaptarse a las necesidades del área.
	Muy alta	RF03	Capaz de integrarse con distintas fuentes de datos.
	Alta	RF04	Contar con un sistema de programación de alertas con parámetros y reglas internas
	Alta	RF05	Proveer la facilidad de crear cuadros de mando con diferentes tipos de información
	Media	RF06	Visualización de detalles en informes y reportes
	Muy alta	RF07	Plan de trabajo para lograr el cumplimiento con el diseño del DataMart
	Alta	RF08	Mediciones e informes periódicos del avance del proyecto
	Alta	RF09	Capacitación sobre inteligencia de negocios

(Continuación)

(Continuación)

Requisitos no funcionales			
Interesado	Prioridad interesado	Requisitos	
		Códigos	Descripción
CEC de la UTM	Alta	RN01	Disponible en idioma español, tanto en su operatividad como en la generación de información final
	Muy alta	RN02	Gestionar la toma de decisiones en el área
	Media	RN03	Intuitivo y fácil uso para los usuarios finales
	Media	RN04	Transversalidad en la información
	Media	RN05	Capacidad para trabajar con cualquier herramienta de interfaz de reportes
	Alta	RN06	Brindar asesoría en todo el ciclo de vida del proyecto
Requisitos de calidad			
Interesado	Prioridad interesado	Requisitos	
		Código	Descripción
CEC de la UTM	Muy alta	RC01	Aseguramiento de la calidad en los datos procesados
Criterios de aceptación			
Concepto		Criterio de aceptación	
Técnicos		Compatibilidad de la base de datos con el DataMart	
De calidad		Lograr satisfacer los requisitos del Patrocinador en un 80%	
Administrativos		Aprobación del Patrocinador de los entregables del proyecto	

Matriz de trazabilidad requisitos:

Código de documento:	CEC-UTM-DM-12-2020	
Fecha: noviembre 2019	Versión: V 1.0	
Proyecto:	Diseño del DataMart	
Unidad:	Área Administrativa del CEC	
Autora:	Sally Astrid Arreaga Santacruz	
Patrocinador:	_____	

Código	Interés	Estado actual	Nivel de estabilidad	Grado de complejidad	Necesidad/ Oportunidad/ Observación	Objetivo del proyecto
RF01	Muy alto	AC	M	A	Menor costo para la entidad	Cumplir el presupuesto
RF02	Medio	AC	M	A	Satisfacción del cliente	Cumplir el alcance
RF03	Muy alto	AC	A	M	Necesidad interna	Cumplir el alcance
RF04	Alto	AC	A	M	Necesidad interna	Cumplir el alcance
RF05	Alto	AC	A	B	Oportunidad de mejora en la calidad	Cumplir con los niveles de calidad
RF06	Medio	AC	A	B	Oportunidad de mejora en la calidad	Cumplir con los niveles de calidad
RN01	Bajo	AC	A	B	Satisfacción del cliente	Cumplir con los niveles de calidad
RN02	Muy alto	AC	A	B	Necesidad interna	Cumplir con el alcance
RN03	Medio	AC	M	M	Satisfacción cliente	Cumplir el cronograma
RN04	Medio	AC	M	M	Necesidad interna	Cumplir el alcance
RN05	Medio	AC	A	M	Necesidad interna	Cumplir el alcance

(Continuación)

(Continuación)

Código	Interés	Estado actual	Nivel de estabilidad	Grado de complejidad	Necesidad/ Oportunidad/ Observación	Objetivo del proyecto
RC01	Muy alto	AC	A	M	Requisito de alto nivel	Cumplir con el alcance
RF07	Muy alto	AC	A	A	Necesidad interna	Cumplir el proyecto
RF08	Alto	AC	A	B	Minimizar riesgos	Cumplir el cronograma
RF09	Alto	AC	A	M	Necesidad interna	Cumplir el alcance
RN06	Muy alto	AC	M	M	Mejorar el rendimiento del proyecto	Cumplir con el alcance
Leyenda						
Estado actual		Nivel de estabilidad		Grado de complejidad		
Estado	Abreviatura	Estado	Abreviatura	Estado	Abreviatura	
Aceptado	AC	Alto	A	Alto	A	
Diferido	DF	Medio	M	Medio	M	
Cancelado	CN	Bajo	B	Bajo	B	

Solicitud de cambios:

Código de documento:	CEC-UTM-DM-13-2020	
Fecha: noviembre 2019	Versión: V 1.0	
Proyecto:	Diseño del DataMart	
Unidad:	Área Administrativa del CEC	
Autora:	Sally Astrid Arreaga Santacruz	
Patrocinador:	_____	

Generalidades del cambio	
ID. Cambio	
Tipo de cambio	
Área del cambio	
Solicitante	
Fecha solicitud	

Causa del cambio	
Descripción cambio propuesto	
Justificación del cambio propuesto	
Impacto en las líneas base	
Alcance	
Cronograma	
Costo	
Calidad	

(Continuación)

(Continuación)

Implicaciones de recursos
Implicaciones para los interesados
Implicaciones en la documentación
Riesgos
Comentarios

4.4.- PLAN DE GESTIÓN DEL TIEMPO

Código de documento:	CEC-UTM-DM-14-2020	
Fecha: noviembre 2019	Versión: V 1.0	
Proyecto:	Diseño del DataMart	
Unidad:	Área Administrativa del CEC	
Autora:	Sally Astrid Arreaga Santacruz	
Patrocinador:	_____	

Proceso de definición de las actividades
<p>La identificación y secuenciación de las actividades necesita que la declaración del alcance, la EDT, estén definidos y aprobados, en este caso se continua con:</p> <ol style="list-style-type: none"> 1.- Por los entregables de EDT se postulan las actividades necesarias para elaborarlo de manera totalitaria, para cada actividad se debe incluir un código de secuencia según la EDT. 2.- Se realiza la secuencia de manera preliminar.
Proceso de secuencia de actividades
<ol style="list-style-type: none"> 1.- Se establecerá una red del proyecto por medio de sus entregables. 2.- Se deben jerarquizar y secuenciar los entregables y sus actividades consecuentes.
Proceso de estimación de recursos de las actividades
<p>Los recursos del proyecto como personal, equipo, insumos, materiales u otros servicios, toma como base los entregables con actividades secuenciadas:</p> <ol style="list-style-type: none"> 1.- Determinar la duración de cada actividad y el tipo de recurso que requiera según la opinión de expertos, cada una con su tiempo de respuesta. 2.- Para el recurso humano se especifica el nombre del integrante del equipo, duración en horas, o supuestos si fuese el caso. 3.- Para los recursos materiales va incorporado el nombre, cantidad en unidades, costos o supuestos si fuese el caso. 4.- EL recurso que tiene que ver con equipos debe incorporar el nombre, cantidad en unidades, costo, o supuestos si fuese el caso. 5.- Si se requiere un servicio adicional, se adhiere la descripción del servicio, cantidad en horas de servicio, costo, proveedor, o supuestos si fuese el caso.
Proceso de estimación de duración en las actividades
<p>El tiempo del proyecto se expresará en horas y se defina por medio de los entregables, además de las actividades que ya fueron definidas, adjuntando la cantidad de personal y la</p>

(Continuación)

(Continuación)

estimación de tiempo para cumplir con la actividad, según la opinión de expertos o plazos acordados. Sin embargo, si fuese el caso de atrasos, se calcularán los nuevos plazos a través de la Estimación por tres valores con distribución triangular según la fórmula:

$$\text{Tiempo estimado}(t_E) = \frac{[\text{Más Probable}(t_M) + \text{Optimista}(t_O) + \text{Pesimista}(t_P)]}{3}$$

Los tres tiempos que se estima son probabilidades de rangos de tiempo que resultan de causantes del atraso en la actividad del proyecto.

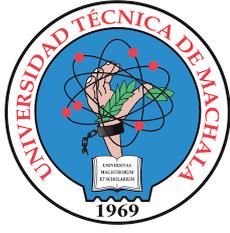
Proceso de desarrollo del cronograma

El desarrollo del cronograma inicia con la identificación y secuenciación de las actividades más relevantes para el diseño del DataMart, así como el tiempo estimable para cumplir con los entregables. Es necesario que exista la aprobación del Patrocinador. Para más detalle el cronograma se desarrolla en el epígrafe 4.4.6 del presente TFM.

Proceso de control del calendario

El proceso de control del cronograma se ejecutará por medio del índice de desempeño del cronograma (SPI) que se calculará semanalmente, además que se analizará en las reuniones bisemanales del proyecto. En el caso de desviaciones, es indispensable replantear el cronograma y comunicarle a las partes interesadas.

4.5.- PLAN DE GESTIÓN DE COSTOS

Código de documento:	CEC-UTM-DM-15-2020	
Fecha: noviembre 2019	Versión: V 1.0	
Proyecto:	Diseño del DataMart	
Unidad:	Área Administrativa del CEC	
Autora:	Sally Astrid Arreaga Santacruz	
Patrocinador:	_____	

Nivel de precisión	Unidades de medida	Umbrales de control
Decimales: No se utilizan, por lo tanto si: $x < 0,50 = 0$ y $x \geq 0,50 = 1$ Decenas: si se utiliza, por lo tanto si: $x < 75 = 75$ y $x \geq 75 = 100$ Nota: Se aplica para los resultados finales y no para los respectivos cálculos.	Horas	Óptimo: +/- 5% costo planeado Aceptable: +/- 10% costo planeado Crítico: > +/- 10% costo planeado
Reglas para la medición del desempeño		
Para establecer los costos del proyecto se requiere tomar en cuenta, los siguientes aspectos: <ol style="list-style-type: none"> 1.- Costo de elaboración de los entregables. 2.- Costo de los requerimientos de la calidad. 3.- Costo vinculado a los riesgos del proyecto 4.- Costos de desarrollo e implantación 5.- Costos de imprevistos El detalle del presupuesto se encuentra ubicado en el epígrafe 4.10 del presente TFM. Los reportes de desempeño de las actividades del proyecto se estipulan con la herramienta del método de valor ganado, en particular los indicadores de Gestión de costos (CPI) y de Gestión del cronograma (SPI).		
a.- Índice de rendimiento de costos (CPI): Expresa la eficiencia en el costo del proyecto, para determinar la eficiencia de uso de los recursos, y se expresa según la fórmula:		
$CPI = \frac{\text{Valor ganado (EV)}}{\text{Costo real (AC)}}$		
Para consolidar el resultado, es importante determinar el trabajo útil por dólar gastado:		
$\text{Trabajo útil por dólar gastado} = X * CPI$		

(Continuación)

(Continuación)

Donde, X = dólares gastados.

Interpretación:

CPI < 1 = sobrecosto con referencia a las estimaciones

CPI > 1 = costo inferior a las estimaciones

CPI = 1 = costo de acuerdo con el presupuesto

b.- Índice de desempeño del cronograma (SPI): Expresa la eficiencia del equipo de trabajo en la cronología del tiempo establecido para el desarrollo del proyecto. Además, ayuda a determinar junto con el CPI las estimaciones de finalización del proyecto, se calcula mediante la fórmula:

$$SPI = \frac{\text{Valor ganado (EV)}}{\text{valor planificado (PV)}}$$

Para el complemento del resultado, se calcula las horas útiles invertidas en el proyecto:

$$\text{Horas útiles} = X * SPI$$

Donde, X = horas trabajadas

Interpretación:

SPI < 1 = retraso, el tiempo está sobre lo programado

SPI > 1 = adelanto, el tiempo está debajo de lo programado

SPI = 1 = acorde, el tiempo es el mismo programado

4.4.3.- Paquete de trabajo 3: Diseño y desarrollo de la arquitectura

En base a la metodología adoptada, se plantea una mezcla de varios artefactos y herramientas que ayudan a la construcción de un sistema DataMart para la toma de decisiones, como son el modelamiento para la elaboración del repositorio central de información aplicando la metodología Kimball, desarrollo del proyecto con un método ágil como SCRUM, que no solo tienen que ver con la parte relacionada a la construcción de sistemas de información sino que engloba el conocimiento profundo de los procesos que se realiza en el CEC y la visión estratégica que tiene, en la figura 19 se visualiza la arquitectura con la metodología planteada para el desarrollo del DataMart para la toma de decisiones.

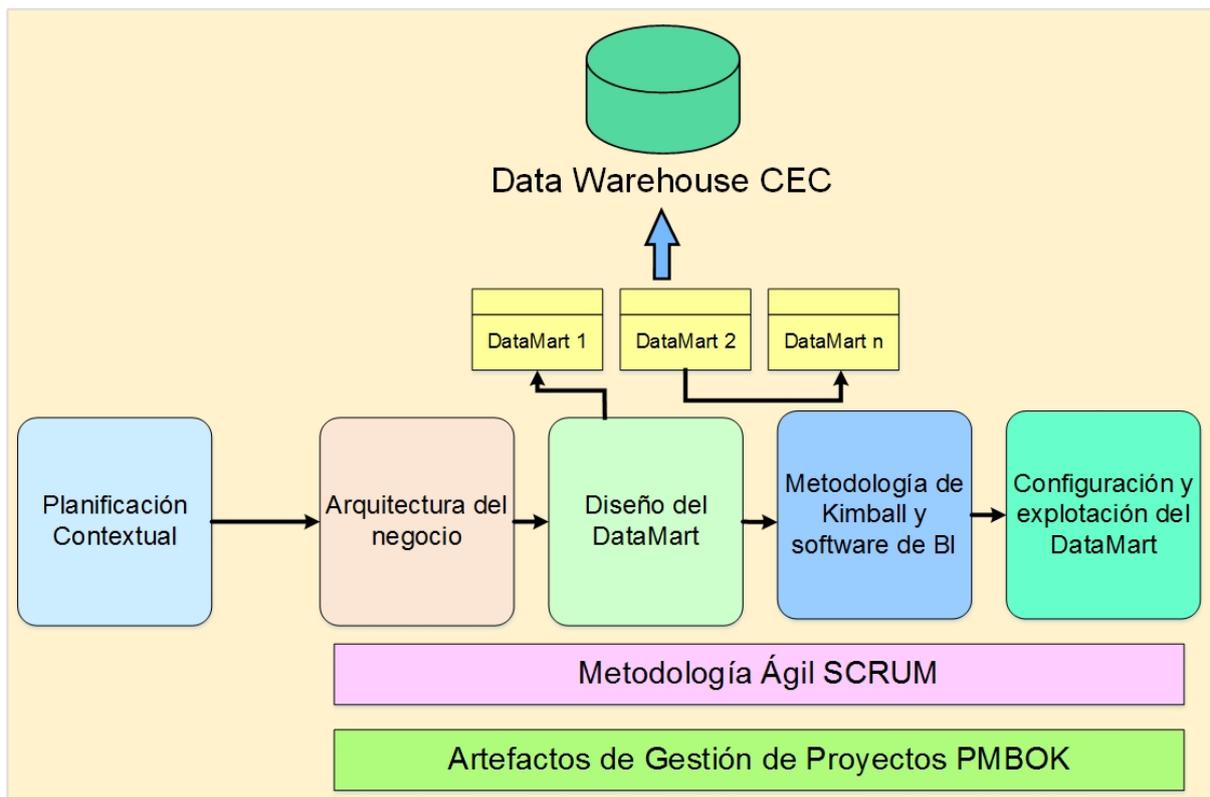


Figura 19. Arquitectura de la solución. Fuente: Elaboración propia

Esta arquitectura está enfocada en conceptos de Data Warehouse especialmente por el hecho de obtener un DataMart, parte fundamental del Data Warehouse como repositorio central de información, validada e íntegra que permite al CEC explotarlo para la toma de decisiones y para conseguir resultados para la inteligencia de negocios.

Se han considerado cinco etapas, las mismas que se plantean inicialmente una forma de planificación del proyecto completa y ágil, un conocimiento profundo de la entidad que ayude a estipular adecuadamente los requerimientos de información, una integración de las distintas

fuentes de información que se tienen dentro de la institución respecto a los servicios y actividades que se realizan para la academia y la formación permanente, una propuesta tecnológica en donde va soportado la integración y explotación de la información por medio de la metodología de Kimball, y finalmente una descripción de la explotación del DataMart consolidado para apoyar al Área Administrativa en la toma de decisiones.

4.4.4.- Paquete de trabajo 4: Desarrollo de un prototipo demostrador

En este epígrafe, se analiza los datos transaccionales y la información procesada en el CEC que se han acumulado a lo largo de los años. El objetivo general de este epígrafe es comprender los módulos principales de la plataforma del CEC e identificar a todos los usuarios comerciales que necesitan acceder al DataMart.

Además, el análisis comercial en esta fase ayuda a comprender las relaciones entre los usuarios y las operaciones comerciales. En este epígrafe, son identificados los datos que están en uso actualmente por los usuarios y cómo les gustaría mejorar. El análisis del sistema depende de la retroalimentación de los usuarios comerciales que dará como resultado entidades comerciales, jerarquías y relaciones entre las entidades.

Para el caso del presente TFM, el análisis incluyó un estudio detallado de los procedimientos actuales de presentación de informes en el CEC, lo que condujo a nuevas especificaciones para el DataMart. Durante el análisis, se examinó los procedimientos de admisión y registro, financiación estudiantil y finanzas para diseñar los respectivos DataMart.

4.4.4.1.- Datos del CEC

Los datos disponibles dentro del CEC de la Universidad Técnica de Machala son muy grandes ya que se acumularon durante años. El CEC necesita hacer un mejor uso de estos datos para tomar decisiones rápidas y correctas en el momento adecuado. El TFM tiene el propósito de proporcionar una solución de DataMart, que ayudará principalmente en el proceso de toma de decisiones relacionadas con los estudiantes. Dentro del alcance del TFM, se analiza el desarrollo del DataMart para una parte del área Administrativa que es esencial para los estudiantes, como son Admisión y Registro, Financiamiento Estudiantil y Finanzas.

4.4.4.2.- Requisitos del sistema

Para el propósito del DataMart, se analizó el uso del sistema operativo Microsoft Windows. La aplicación principal, que se utilizará es ORACLE 11g como un motor de base de datos, que a

su vez, es un sistema de gestión de bases de datos relacionales comercializado por Oracle Corporation. El software tiene el sistema de gestión de bases de datos relacionales que es capaz de almacenar todos los datos necesarios para el Data Warehouse (DW) y lógicamente el DataMart. El proceso de ETL se construirá utilizando Talend Open Studio, que es un proveedor de integración de software. Talend Open Studio fue el primer proveedor comercial de software de código abierto de software de integración de datos.

Para la implementación del front-end se utilizará la herramienta BI Publisher de Oracle Corporation. Utilizada para diseñar informes interactivos de acuerdo con la solicitud del usuario y también para diseñar diferentes medidores y paneles. Tiene la capacidad de diseñar diferentes tablas y gráficos. Se puede integrar fácilmente con la plataforma del CEC utilizando los servicios web de BI Publisher.

Para ejecutar el software pertinente al diseño del DW, el sistema operativo sería Windows 2012 y superior, con un mínimo de 8 GB de memoria, incluidos los datos, se requieren 100 GB de espacio en el disco duro. La velocidad del procesador debe ser de 3.0 MHz en adelante; los otros componentes del computador son necesarios para admitir el sistema operativo y el resto de aplicaciones necesarias para diseñar el DataMart.

4.4.4.3.- Diseño del sistema

En esta fase, el objetivo es obtener las especificaciones del sistema a partir de los requisitos del usuario, que se recopilan de las partes interesadas. Algunos de los componentes de Data Warehousing se basan en estas especificaciones, como el extractor de datos, la transformación de datos y las herramientas de integración de datos. Las especificaciones del sistema se basan en el diseño lógico y físico de los DataMarts en el Data Warehouse.

Como se discutió anteriormente en la revisión de la literatura del TFM, hay dos enfoques para el Data Warehousing que son los enfoques de Bill Inmon y Ralph Kimball. En el diseño del DataMart se adoptó el enfoque de Ralph Kimball. Los requisitos recopilados de cada dimensión de la plataforma del CEC se traducen en DataMarts, y los DataMarts resultantes presentan la versión final del Data Warehouse.

Una metodología de desarrollo describe la evolución y gestión esperada de la ingeniería del sistema. Uno de los principios más importantes de la Ingeniería de Sistemas es evaluar un sistema desde la perspectiva del Ciclo de Vida. El establecimiento de una metodología también proporcionará una estrategia para el gerente del proyecto y el equipo del proyecto a

medida que ejecutan el proyecto del DataMart a lo largo de todas las fases de desarrollo (Burton & Green, 2016).

Modelo en espiral:

El modelo en espiral es una secuencia de modelos en cascada que corresponde a una mejora iterativa orientada al riesgo, y reconoce que los requisitos no siempre están disponibles y claros cuando el sistema se implementa por primera vez.

Dado que diseñar y construir un DW es un proceso iterativo, el método espiral es la mejor metodología de desarrollo para el propósito del TFM. La figura 20 muestra una serie de cascada en un modelo espiral recomendado de un ciclo de vida del DW.

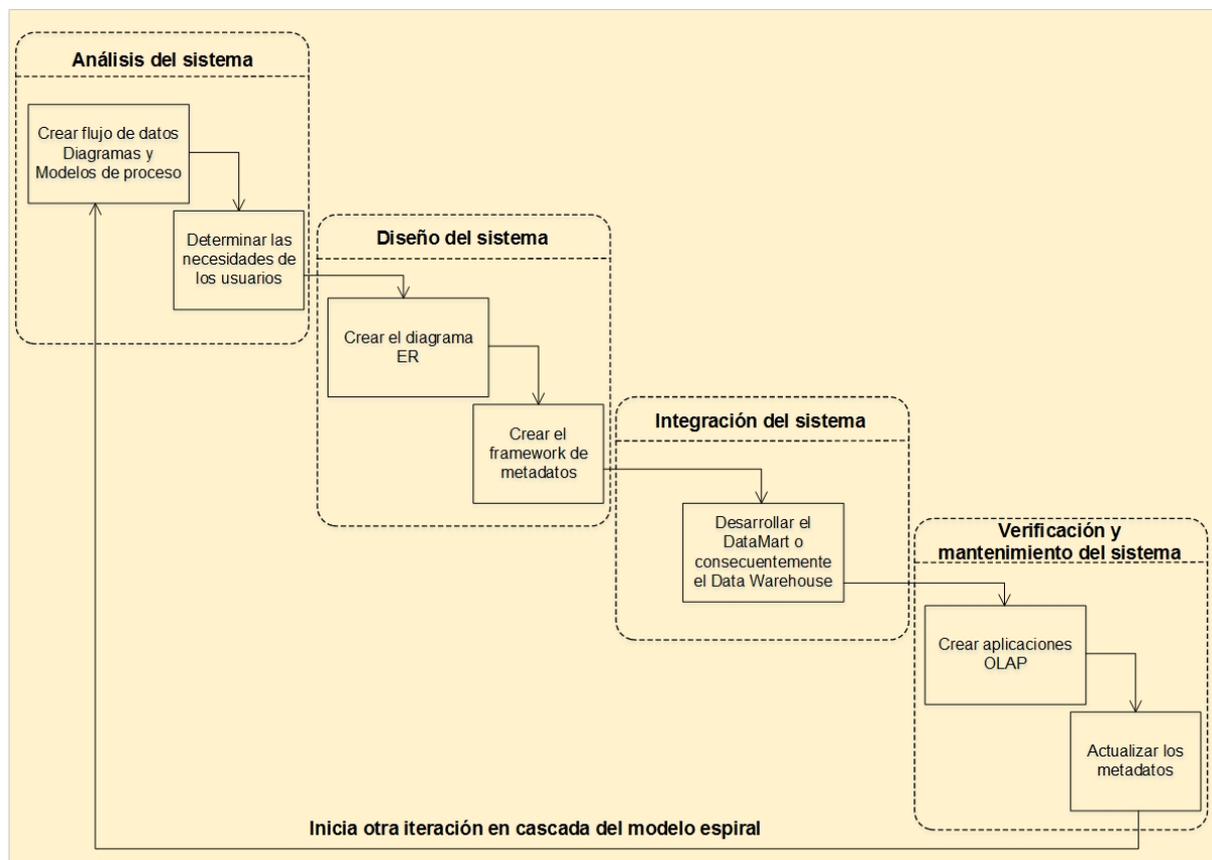


Figura 20. Modelo en espiral del ciclo de vida del almacén de datos. **Fuente:** Elaboración propia adaptado de Burton & Green (2016).

Se inicia el diseño del DataMart especificando la medida; las medidas son la base y la información de retroalimentación que requieren los tomadores de decisiones. Para ello se concilia los requisitos con lo que está disponible en el sistema fuente (OLTP). Para el propósito del TFM, se utilizará el esquema en estrella para el diseño del DataMart. El esquema en estrella es un esquema de base de datos relacional que se utiliza para contener medidas y

dimensiones en un DataMart. Las medidas se almacenan en una tabla de hechos y las dimensiones se almacenan en tablas de dimensiones. Para cada DataMart, solo hay una tabla de hechos rodeada por las tablas de dimensiones, de ahí el nombre de esquema de estrella.

Como se mencionó en la literatura, el centro de la estrella está formado por la tabla de hechos. La tabla de hechos tiene una columna o la medida y la columna para cada dimensión que contiene la clave foránea para un registro de esas dimensiones. La clave para esta tabla se forma concatenando todos los campos de clave foránea. La clave principal para la tabla de hechos generalmente se conoce como clave compuesta; contiene las medidas, de ahí el nombre "Hecho".

Las dimensiones se almacenan en la tabla de dimensiones. La tabla de dimensiones tiene una columna para el identificador único de un miembro de la dimensión, generalmente un número entero de un valor de carácter corto; además, tiene otra columna para la descripción. Una de las partes más importantes del DW es la extracción, transformación y carga de datos desde las bases de datos transaccionales operativas al propio DW para aprovechar al máximo los datos. Los dos procesos principales del DW son la carga de datos y el acceso a los datos. El diseño de un DW es muy robusto sobre todo en la carga del propio DW, por esta razón se espera realizarlo mediante el uso del proceso ETL.

En base a lo argumentado el diseño de las bases de datos comenzó con las teorías del diseño de base de datos y la regla que respalda las necesidades comerciales; por lo tanto, se inició el proceso de diseño del DataMart con el diseño lógico.

Modelos lógicos:

El modelo lógico es una representación de los datos de una manera que se puede presentar a la entidad, para el caso el CEC, para que sirva como una hoja de ruta para la implementación física. Los elementos principales de un modelo lógico son entidades, atributos y relaciones. De esta forma, se inició el diseño de los DataMarts a través de las tablas de hechos y dimensiones; es preciso recordar que todo el diseño de la base de datos comienza con un diseño lógico.

Tablas de hechos y dimensiones:

La tabla de hechos contiene las medidas asociadas con un proceso comercial específico. Un registro en una tabla de hechos es una medida, y un evento de medición siempre puede

producir un registro de tabla de hechos. Estos eventos generalmente tienen medidas numéricas que cuantifican la magnitud de los eventos; a estos números se les llaman hechos.

Las dimensiones son la base del modelo dimensional, que describe los objetos del negocio, como el estudiante, la colegiatura, el semestre y otras tablas de dimensiones que se utilizarán en este diseño del DataMart.

Según Ralph Kimball, la dimensión sirve como los sustantivos del sistema DW/BI. Describen los eventos de medición circundantes junto con los procesos de negocio (hechos) que son la acción del negocio en el que participa la dimensión. Cada tabla de dimensiones enlaza con todos los procesos de negocio en los que participa. En este sentido, los DataMarts representan una unidad o proceso departamental dentro de una organización, por esta razón, se puede afirmar que un DataMart es la colección de tablas de hechos y sus tablas de dimensiones; usando el diseño de un DW de abajo hacia arriba, la combinación del DataMart formaría el DW.

Según la etapa de análisis del sistema, el diseño final del DataMart tiene 11 tablas de hechos y 15 tablas de dimensiones. Muchas tablas de dimensiones se comparten entre diferentes tablas de hechos. En este epígrafe es pertinente describir primero la tabla de dimensiones, luego ilustrar cada uno de los DataMarts en el DW. En todas las tablas siguientes se muestra los nombres de las columnas, el tipo de datos de cada columna y la descripción y el propósito de cada una.

Dimensión de los años académicos (DIM_ANO_ACADEMICO_ACD):

Esta dimensión contiene datos simples sobre todos los años académicos del CEC, las columnas se describen en la tabla 8. El número al lado del tipo de dato VARCHAR denota el tamaño de esa columna.

Tabla 8. *Dimensión de años académicos*

Nombre Columna	Tipo de dato	Descripción
ANO_NUMERO	NUMERIC	Número del año académico formado por 4 dígitos, por ejemplo: 2019
ANO_TITULO	VARCHAR(60)	Título descriptivo del año académico
FECHA_INICIO	DATE	Primer día del año académico
FECHA_FINAL	DATE	Ultimo día del año académico

Fuente: Elaboración propia

Dimensión de los semestres académicos (DIM_SEMETRE_ACADEMICO_ACD):

Los semestres académicos son cargados en la tabla de dimensiones DIM_SEMETRE_ACADEMICO_ACD; la tabla se describe en la tabla 9.

Tabla 9. *Dimensión de los semestres académicos*

Nombre Columna	Tipo de dato	Descripción
SEMESTRE_NUMERO	NUMERIC	Número de semestre académico identificado por el sistema heredado del CEC. Contiene 5 dígitos, los primeros 4 dígitos representan el año académico y el último dígito representa el número de semestre (1 para el primero, 2 para el segundo, 3 para el tercero)
SEMESTRE_TITULO	VARCHAR(100)	Título descriptivo del semestre
ANO_NUMERO	NUMERIC	Número del año académico formado por 4 dígitos, por ejemplo: 2019
ANO_TITULO	VARCHAR(60)	Título descriptivo del año académico
FECHA_INICIO	DATE	Primer día del año académico
FECHA_FINAL	DATE	Ultimo día del año académico

Fuente: Elaboración propia

Dimensión de la fecha (DIM_FECHA_GEN):

La dimensión de fecha proporciona datos adicionales sobre una fecha. Incluye información como la fecha completa, el día de la semana, el nombre del día y el nombre del mes, y es una dimensión general que debería existir en cada DataMart. Esta dimensión es proporcionada por el grupo Kimball (The Microsoft Data Warehouse Toolkit, 2017) y se describe en la tabla 10 a continuación.

Tabla 10. *Dimensión de la fecha*

Nombre Columna	Tipo de dato	Descripción
FECHA_KEY	NUMERIC	Clave primaria que contiene 8 dígitos en formato "aaaammdd" de la fecha, por ejemplo: 20191119

Fuente: Elaboración propia

(Continuación)

Nombre Columna	Tipo de dato	Descripción
SEMESTRE_NUMERO	NUMERIC	Clave externa de la dimensión de semestres
FECHA_FULL	DATE	Fecha completa como objeto de fecha
NOMBRE_FECHA	VARCHAR(11)	Fecha en formato simple
DIA_SEMANA	NUMERIC	Día de la semana. A continuación se muestra el día que corresponde a la semana: 1 - domingo 2 - lunes 3 - martes 4 - miércoles 5 - jueves 6 - viernes 7 - sábado
DIA_NOMBRE_SEMANA	VARCHAR(10)	Nombre del día de la semana, por ejemplo: "Sábado", "Viernes"
DIA_MES	NUMERIC	Día del mes. Puede ser un número entre 1-31
DIA_AÑO	NUMERIC	Día del año. Puede ser un número entre 1-365
DIALABORABLE_FINDESEMANA	VARCHAR(10)	Clasifica el día como "Día de la semana" o "Fin de semana"
SEMANA_AÑO	NUMERIC	Número de la semana en el año
NOMBRE_MES	VARCHAR(10)	Nombre del mes, por ejemplo, "enero"
MES_AÑO	NUMERIC	Número del mes en el año
ULTIMO_DIA_MES	NUMERIC	Clasifica el día si es el último día del mes ("Y") o no ("N")
TRIMESTRE_CALENDARIO	NUMERIC	Número trimestral en el año
AÑO_CALENDARIO	NUMERIC	Numero de año

Fuente: Elaboración propia

(Continuación)

Nombre Columna	Tipo de dato	Descripción
ANO_MES_CALENDARIO	VARCHAR(10)	Número de año y mes separados por un "-" (es decir, "1990-01")
ANO_QTR_CALENDARIO	VARCHAR(10)	Número de año y trimestre separados por un "-" (es decir, "1990-Q1")
MES_ACADEMICO_SEMESTRE	NUMERIC	Número del mes académico en el semestre
SEMANA_ACADEMICA_SEMESTRE	NUMERIC	Número de la semana académica en el semestre
DIA_ACADEMICO_SEMESTRE	NUMERIC	Número del día académico en el semestre

Fuente: Elaboración propia

Dimensión de programas de estudio (DIM_PROGRAMAS_ESTUDIO_ACD):

Esta dimensión es muy simple ya que solo almacena el número de clave del programa académico y su título como se describe en la tabla 11.

Tabla 11. Dimensión de los programas de estudio

Nombre Columna	Tipo de dato	Descripción
NUMERO_PROGRAMA	VARCHAR(1)	Clave primaria del programa académico 'B': Bachiller 'M': Máster 'D': Doctorado
NOMBRE_PROGRAMA	VARCHAR(60)	Título del programa

Fuente: Elaboración propia

Dimensión de las colegiaturas (DIM_COLEGIATURA _ ACD):

Esta dimensión almacena el nombre de la colegiatura y para qué programa académico está relacionado. Además, tiene un atributo para indicar si la colegiatura está activa o cerrada. La tabla 12 muestra las columnas. La dimensión de la colegiatura permitirá establecer si el estudiante permanece en formación continua o no; este atributo vinculado al sistema de la Universidad podrá depurar información de estudiantes rezagados.

Tabla 12. *Dimensión colegiatura*

Nombre Columna	Tipo de dato	Descripción
NUMERO_COLEGIATURA	NUMERIC	Clave principal
ES_ACTIVIA	VARCHAR(1)	'Y' significa que está activa y 'N' en el caso de que la colegiatura no está disponible

Fuente: Elaboración propia

Dimensión departamentos (DIM_DEPARTAMENTOS _ ACD):

Esta tabla contiene detalles del departamento. Un departamento debe estar relacionado con una colegiatura como se muestra en la tabla 13.

Tabla 13. *Dimensión de departamentos*

Nombre Columna	Tipo de dato	Descripción
NUMERO_DEPARTAMENTO	NUMERIC	Clave principal
NUMERO_COLEGIATURA	NUMERIC	Clave foránea para la colegiatura
ES_ACTIVIA	VARCHAR(1)	'Y' significa que está activa y 'N' en el caso de que la colegiatura no está disponible

Fuente: Elaboración propia

Dimensión de cursos (DIM_CURSOS _ ACD):

La tabla 14 muestra la descripción de cada columna en la tabla Dimensión de Cursos, esta tabla contiene información relacionada con los cursos referentes a la educación continua que promueve la Universidad Técnica de Machala.

Tabla 14. *Dimensión de cursos*

Nombre Columna	Tipo de dato	Descripción
NUMERO_ASIGNATURA	NUMERIC	Clave principal
CODIGO_ASIGNATURA	VARCHAR(10)	Manualmente identificado el código de la asignatura contiene 5 letras y 5 dígitos
TIPO_ASIGNATURA	NUMERIC	Identifica si la asignatura es obligatoria u opcional
NUMERO_DEPARTAMENTO	NUMERIC	Clave foránea de Departamento

Fuente: Elaboración propia

(Continuación)

Nombre Columna	Tipo de dato	Descripción
HORAS_ACADEMICAS	NUMERIC	Las horas académicas en el plan
HORAS_FINANCIERAS	NUMERIC	Horas financieras a pagar

Fuente: Elaboración propia

Dimensión de años de colegiatura (DIM_ANOS_COLEGIATURA_ACD):

En esta tabla, se define los años de colegiatura. Contiene la clave para cada año en formato de fecha "aaaa" y un título descriptivo para este año como se muestra en la tabla 15.

Tabla 15. Dimensión de los años de colegiatura

Nombre Columna	Tipo de dato	Descripción
NUMERO_ANO	NUMERIC	Clave primaria que se formó con 4 dígitos del año educativo
NOMBRE_ANO	VARCHAR(60)	Título descriptivo del año

Fuente: Elaboración propia

Dimensión de los grados de colegiatura (DIM_GRADOS_COLEGIATURA_ACD):

La tabla de calificaciones de colegiatura almacena información adicional sobre el rango de calificaciones de los estudiantes. Esta dimensión se describe en detalle en la tabla 16.

Tabla 16. Dimensión de los grados de colegiatura

Nombre Columna	Tipo de dato	Descripción
NUMERO_GRADO	VARCHAR(2)	Clave primaria que debe ser una de las siguientes: 'A +' para calificaciones entre 95 y 100 'A' para grados entre 90 y 94.9 'B +' para calificaciones entre 85 y 89.9 'B' para grados entre 80 y 84.9 'C +' para grados entre 75 y 79.9 'C' para grados entre 70 y 74.9 'D +' para grados entre 65 y 69.9 'D' para grados entre 60 y 64.9 'E +' para grados entre 55 y 59.9 'E' para grados entre 50 y 54.9 'F' para calificaciones inferiores a 50

Fuente: Elaboración propia

(Continuación)

Nombre Columna	Tipo de dato	Descripción
TITULO_GRADO	VARCHAR(60)	Título descriptivo de la nota
GRADO_DESDE	NUMERIC	Calificación mínima como se describe en la columna NUMERO_GRADO
GRADO_A	NUMERIC	Calificación máxima como se describe en la columna NUMERO_GRADO

Fuente: Elaboración propia

Dimensión de las secciones de colegiatura (DIM_SECCIONES_COLEGIATURA_ACD):

Esta es una tabla de dimensiones simple, tabla 17, que almacena una clave de serie de la sección de colegiatura predefinida en el sistema heredado del CEC, y un título descriptivo, es decir, "Sección científica".

Tabla 17. Dimensión de las secciones de colegiatura

Nombre Columna	Tipo de dato	Descripción
NUMERO_RAMA	NUMERIC	Clave principal/número de serie de la sección
NOMBRE_RAMA	VARCHAR(60)	Título descriptivo de la sección

Fuente: Elaboración propia

Dimensión de estados estudiantiles (DIM_ESTADOS_ESTUDIANTILES_ACD):

En esta tabla de dimensiones, un código de estado único y un nombre para el estado se definen, como se describe en la tabla 18.

Tabla 18. Dimensión de estados estudiantiles

Nombre Columna	Tipo de dato	Descripción
CODIGO_ESTADO	VARCHAR(1)	Clave principal/una letra que representa el estado académico del estudiante, es decir, 'R' para estudiantes regulares
NOMBRE_ESTADO	VARCHAR(60)	Título descriptivo del estado

Fuente: Elaboración propia

Para poder corroborar el estado del estudiante, es necesario una dimensión de niveles.

Dimensión de niveles de los estudiantes (DIM_NIVEL_ESTUDIANTE_ACD):

La tabla DIM_NIVEL_ESTUDIANTE_ACD, tabla 19, contiene detalles sobre los niveles académicos en el CEC de la Universidad Técnica de Machala. Un estudiante universitario comienza su vida académica en el nivel 1, y si pasa ciertas horas académicas, pasa al siguiente nivel.

Tabla 19. *Dimensión de niveles de los estudiantes*

Nombre Columna	Tipo de dato	Descripción
NUMERO_NIVEL	NUMERIC	Clave principal y un número que representa la secuencia del nivel académico del estudiante, es decir, 1 para el primer nivel, 2 para el segundo nivel, etc.
NOMBRE_NIVEL	VARCHAR(60)	Título descriptivo del nivel

Fuente: Elaboración propia

Dimensión de ubicaciones geográficas (DIM_UBICACION_GEOGRAFICA_ACD):

La tabla DIM_UBICACION_GEOGRAFICA_ACD como se describe en la tabla 20, enumera todos los vecindarios predefinidos en el sistema CEC. Además, vincula los vecindarios con sus ciudades y provincias como se muestra a continuación.

Tabla 20. *Dimensión de ubicaciones geográficas*

Nombre Columna	Tipo de dato	Descripción
ID_VECINDARIO	NUMERIC	Clave principal y un número de serie de ubicaciones predefinidas en el sistema heredado CEC
NOMBRE_VECINDARIO	VARCHAR(60)	Nombre del vecindario
ID_CIUDAD	NUMERIC	ID de la ciudad en la que está relacionado este vecindario
NOMBRE_CIUDAD	VARCHAR(60)	Nombre de la ciudad
ID_PROVINCIA	NUMERIC	ID de la provincia que se identificó en el sistema CEC
NOMBRE_PROVINCIA	VARCHAR(60)	Nombre de la provincia

Fuente: Elaboración propia

Dimensión de los estudiantes (DIM_ESTUDIANTES_ACD):

En esta dimensión se almacenan instancias para cada estudiante en el CEC de la Universidad Técnica de Machala, algunos metadatos almacenados en esta dimensión se describen en detalle en la tabla 21.

Tabla 21. Dimensión de los estudiantes

Nombre Columna	Tipo de dato	Descripción
NUMERO_ESTUDIANTE	NUMERIC	Clave principal y un número que representa el identificador académico del alumno. Contiene 9 dígitos
NOMBRE_ESTUDIANTE	VARCHAR(100)	Nombre completo del alumno
ID_ESTUDIANTE	VARCHAR(10)	Identificación ciudadana del estudiante
NACIONALIDAD_ESTUDIANTE	VARCHAR(30)	Identificación nacional del estudiante
GENERO_ESTUDIANTE	VARCHAR(1)	Género del estudiante: 'M' si es Masculino y 'F' para Femenino.
ID_VECINDARIO	NUMERIC	Ubicación del estudiante, clave foránea de Ubicación geográfica.
SEMESTRE_ADMISION	NUMERIC	El semestre cuando el estudiante fue admitido
TIPO_ADMISION	VARCHAR(1)	El tipo de admisión especifica el tipo de estudio previo en el que el estudiante fue aceptado
PROGRAMA_ESTUDIO	VARCHAR(1)	Programa académico del alumno, clave foránea para la dimensión de los Estados
ESTADO_ACADEMICO	VARCHAR(1)	Estado del estudiante que especifica, por ejemplo, si es un estudiante regular o graduado
NIVEL_ESTUDIANTE	NUMERIC	Nivel académico para estudiantes, un valor entre 1 y 6
PUNTAJE_SECUNDARIA	NUMERIC	Puntaje del estudiante en el colegio
RAMA_SECUNDARIA	VARCHAR(2)	Rama de la secundaria
ES_GRADUADO	NUMERIC	Una bandera para el estado de graduación del estudiante, 1: graduado, 0: aún no graduado
FECHA_GRADUACION	DATE	Fecha de graduación del alumno si existe

Fuente: Elaboración propia

Dimensión de becas (DIM_BECAS_ACD):

En esta dimensión, el DataMart almacena información básica sobre las becas o pagos del CEC, que son el número de beca, el título de beca y su tipo, como se muestra en la tabla 22. Las becas CEC son una de las herramientas para ayudar a los estudiantes a completar sus estudios financieramente.

Tabla 22. *Dimensión de las becas*

Nombre Columna	Tipo de dato	Descripción
NUMERO_BECA	NUMERIC	Clave primaria y un número que representa un identificador único de subvención o beca
NOMBRE_BECA	VARCHAR(60)	Nombre de la beca como se identifica en el CEC
TIPO_BECA	NUMERIC	Tipo de becas de la siguiente forma: 1: Beca interna, 2: Beca externa, 3: Préstamo.

Fuente: Elaboración propia

Dimensión de horas de registro de estudiantes (DIM_HORAS_REGISTRO_ESTUDIANTES_ACD):

La dimensión horas de registro de estudiantes contiene detalles sobre los rangos de horas de registro, que solían clasificar el registro de los estudiantes en función de cuántas horas ya se registraron en un semestre específico. Esta tabla de dimensiones se describe en la tabla 23.

Tabla 23. *Dimensión de horas de registro de estudiantes*

Nombre Columna	Tipo de dato	Descripción
NUMERO_SERIAL	NUMERIC	Clave primaria
ETIQUETA_RANGO	VARCHAR(60)	Etiqueta descriptiva para el rango de horas, es decir, "3-6 horas de registro"
VALOR_MINIMO	NUMERIC	Metadatos que representan el valor mínimo de horas de registro de este rango.
VALOR_MAXIMO	NUMERIC	Metadatos que representan el valor máximo de horas de registro de este rango

Fuente: Elaboración propia

DataMart de Admisión y Registro:

En esta sección, se describe cada DataMart diseñado para manejar los procesos operativos del área Administrativa en el evento de Admisión y Registro del CEC.

DataMart de resultados de colegiatura

La figura 21 muestra la tabla de datos de resultados de colegiatura y sus dimensiones relacionadas como un esquema en estrella. La tabla de hechos contiene dos medidas en cada año para cada sección de la colegiatura y dividida en 9 grados, las dos medidas son el conteo de estudiantes y su promedio en la colegiatura. Como se muestra en la figura 21, la tabla de hechos tiene tres claves foráneas para las tablas de dimensiones Años de colegiatura, Rama de colegiatura y Grado de colegiatura que identifica su granularidad.

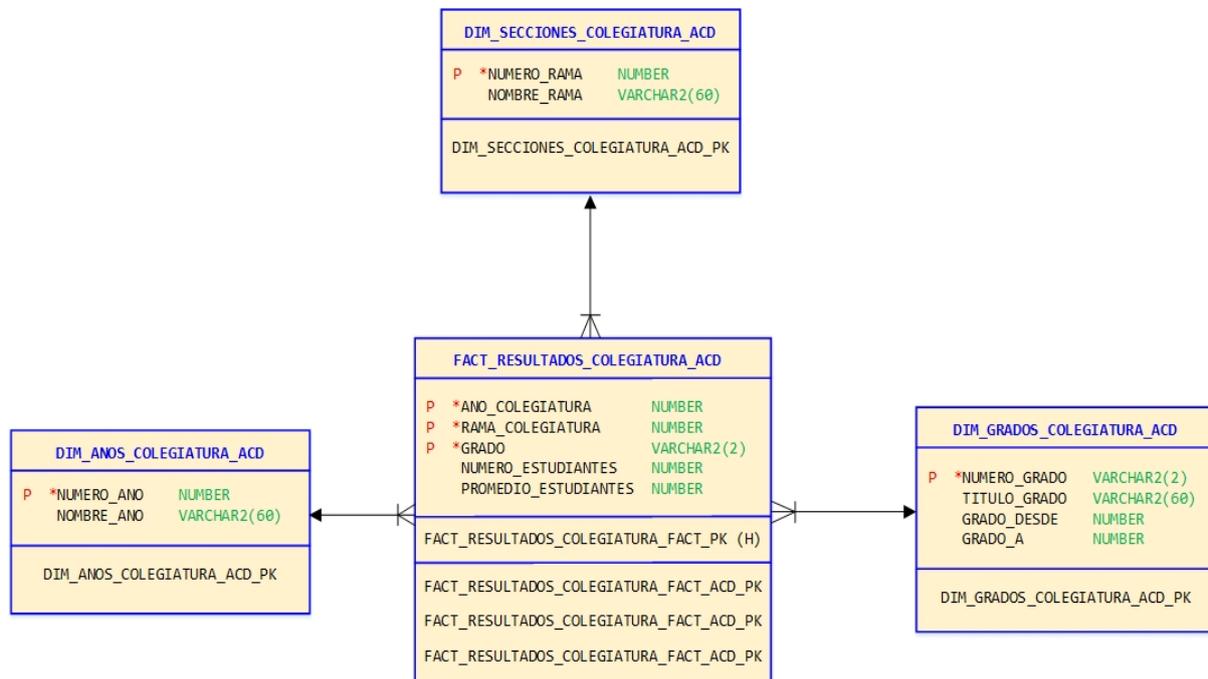


Figura 21. DataMart resultados de colegiatura. Fuente: Elaboración propia

DataMart admisión de estudiantes

DataMart de admisión de estudiantes está diseñado para calcular dos medidas, que son admitidas por los estudiantes y su promedio de calificaciones en la colegiatura. La granularidad de la tabla de hechos identificada como fila 1 por año, por sección de colegiatura, por grado de colegiatura, por departamento, como se muestra en la figura 22. La tabla de hechos se refiere por año académico, sección de colegiatura, grado de colegiatura y dimensiones del departamento.

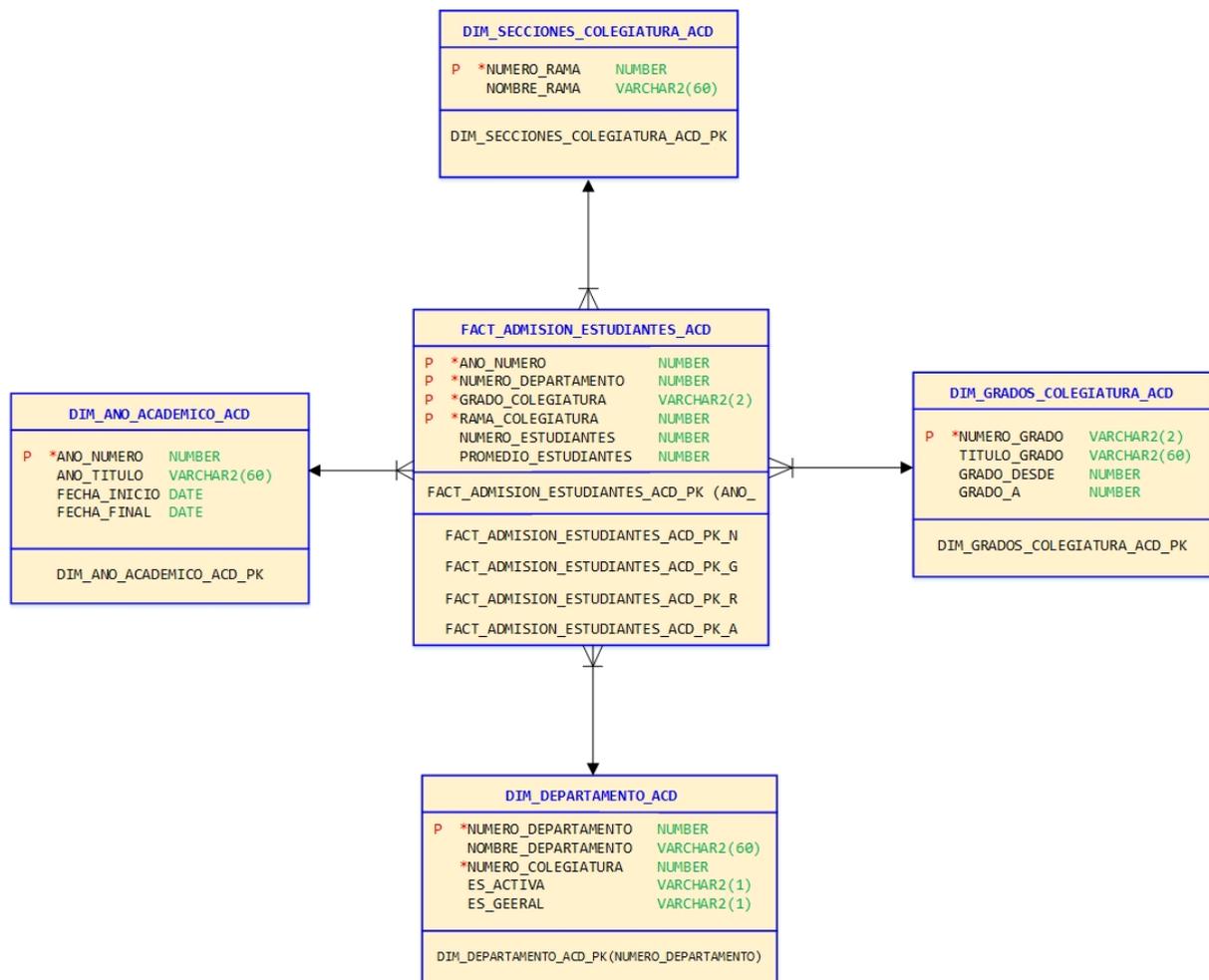


Figura 22. DataMart admisión de estudiantes. Fuente: Elaboración propia

Este DataMart utilizado para calcular el recuento de estudiantes y el promedio de grado en tres casos:

- El total de estudiantes admitidos cuenta con un promedio calculado para cada sección y grado.
- El total de estudiantes admitidos cuenta con un promedio calculado para cada departamento.
- El total de estudiantes admitidos cuenta con un promedio calculado para cada colegiatura.

Se predefine las consultas de cálculo para estos tres casos en vistas lógicas como se muestra en la figura 23. Primero se crea una vista maestra para este DataMart titulado "V_FACT_ADMISION_ACD", esta vista une la tabla de hechos con todas las tablas de dimensiones y sus jerarquías. A continuación, se crea vistas de sección y departamento, llamadas "V_ADMISION_RAMA_ACD" y "V_ADMISION_DEPARTAMENTO_ACD"

respectivamente, basado en "V_FACT_ADMISION_ACD". Finalmente, la vista de la colegiatura es compilada basada en "V_ADMISION_DEPARTAMENTO_ACD".

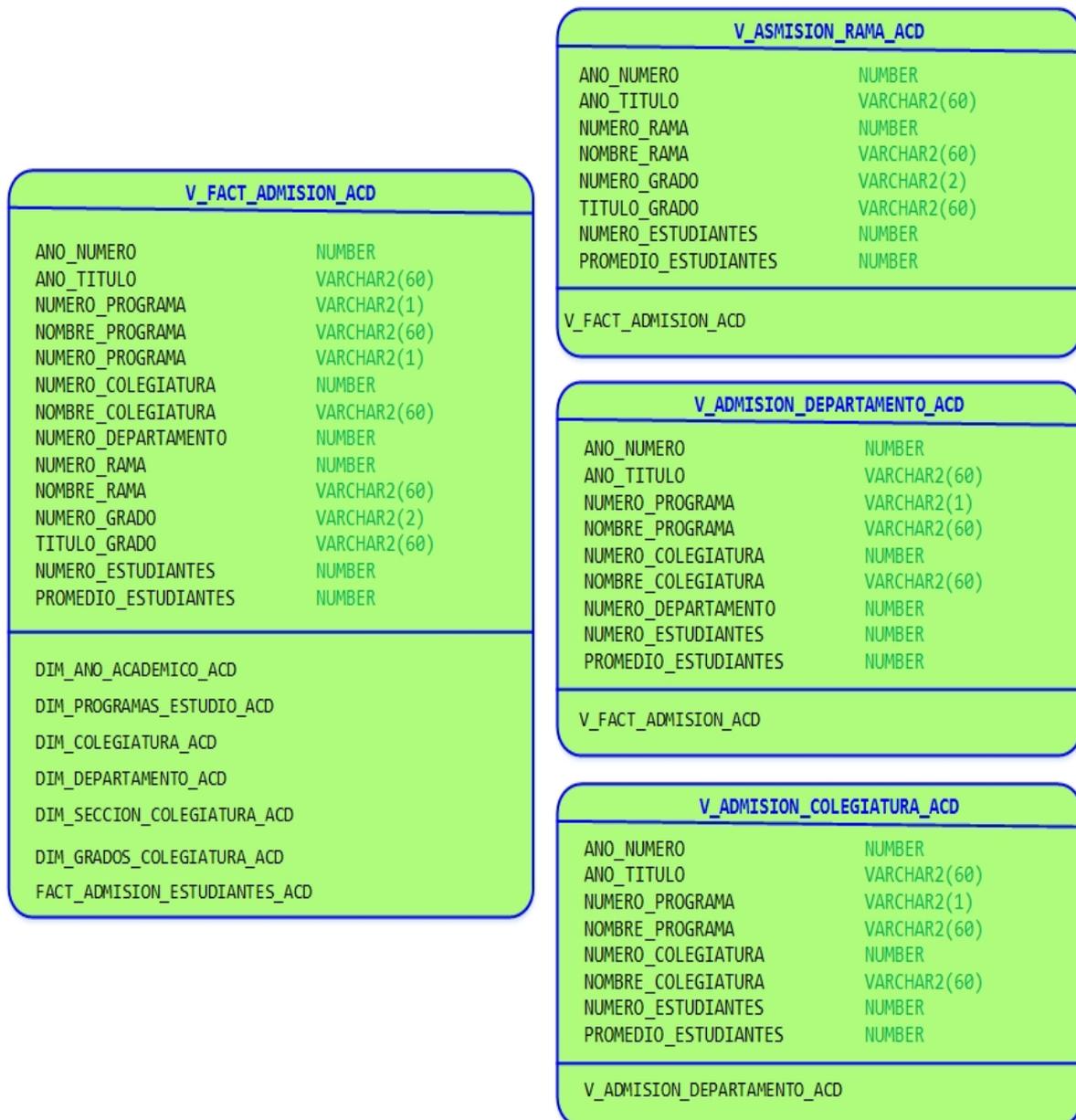


Figura 23. *Vistas de admisión de estudiantes.* **Fuente:** Elaboración propia

DataMart inscripción de estudiantes

Este DataMart tiene una medida simple que los estudiantes registrados cuentan en un semestre. Esta medida se dividió en cinco dimensiones enumeradas como: Semestre, Estado de los estudiantes, Género, Programa académico y Horas de registro, así como Rango Dimensiones como se muestra en la figura 24, que identifican la granularidad de hechos.

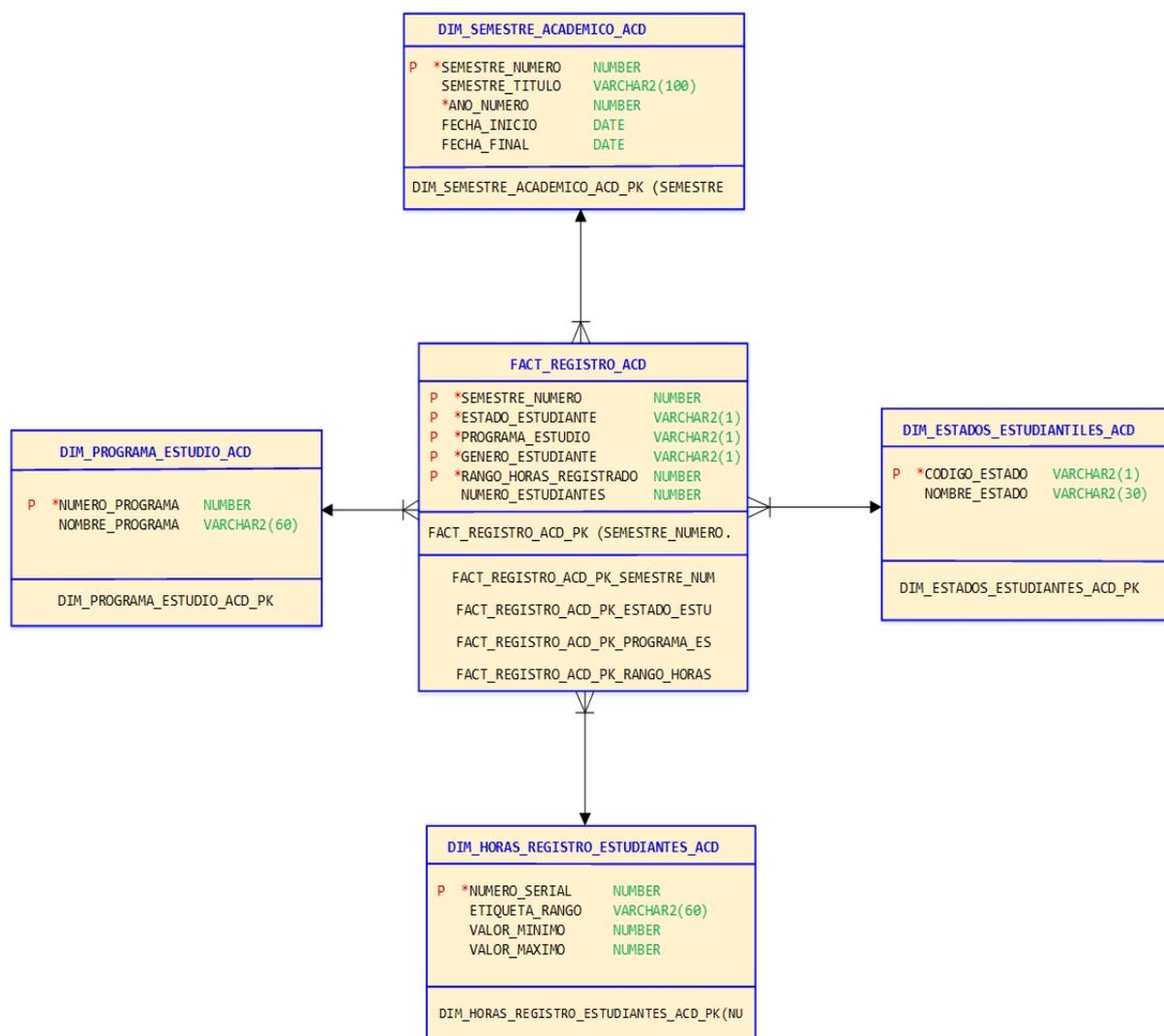


Figura 24. DataMart de registro de estudiantes. Fuente: Elaboración propia

Según los requisitos comerciales, los usuarios deben calcular el recuento total de estudiantes registrados en semestres académicos de dos maneras:

- El total de estudiantes registrados cuenta para cada Programa Académico y para cada Género, finalmente categorizados según el Rango de Horas de Registro disponible.
- El total de estudiantes registrados cuenta para cada programa académico y género.

Por lo tanto, de acuerdo con los requisitos comerciales, se crea vistas lógicas que se muestran en la figura 25. "V_FACT_REGISTRO_ACD" es la vista maestra que une la tabla de hechos con las cuatro tablas de dimensiones. A continuación, en base a esta vista maestra, se crea "V_HORAS_REGISTRO_ACD" que abordan el primer cálculo requerido. Luego, finalmente, la vista "V_SEMESTRES_REGISTRO_ACD" calcula el total de estudiantes registrados en semestres académicos para cada programa académico y género.

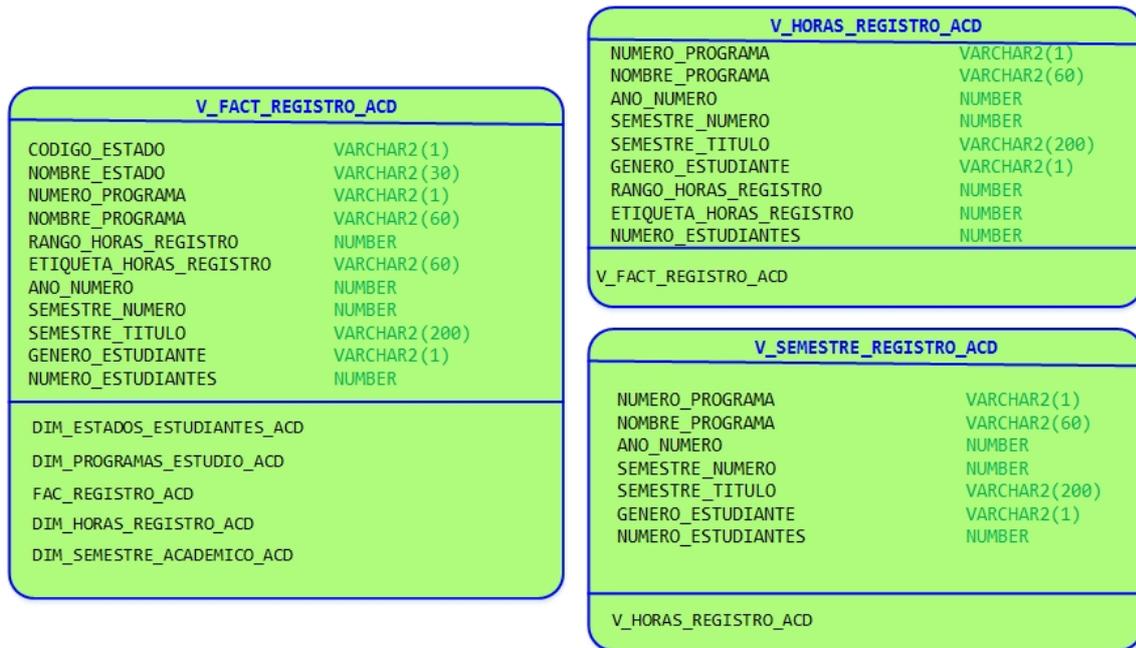


Figura 25. Vistas de registro de estudiantes. Fuente: Elaboración propia

DataMart Ubicación Geográfica de estudiantes registrados

El CEC de la Universidad Técnica de Machala necesita investigar la ubicación de los estudiantes cada semestre para poder facilitar el transporte de los estudiantes. Por lo tanto, esta tabla de hechos contiene el recuento total de estudiantes registrados en cada semestre para cada colegiatura con respecto a sus ubicaciones. La tabla de hechos que se muestra en la figura 26 tiene tres claves foráneas para las dimensiones Vecindario, Semestre y Colegiatura.

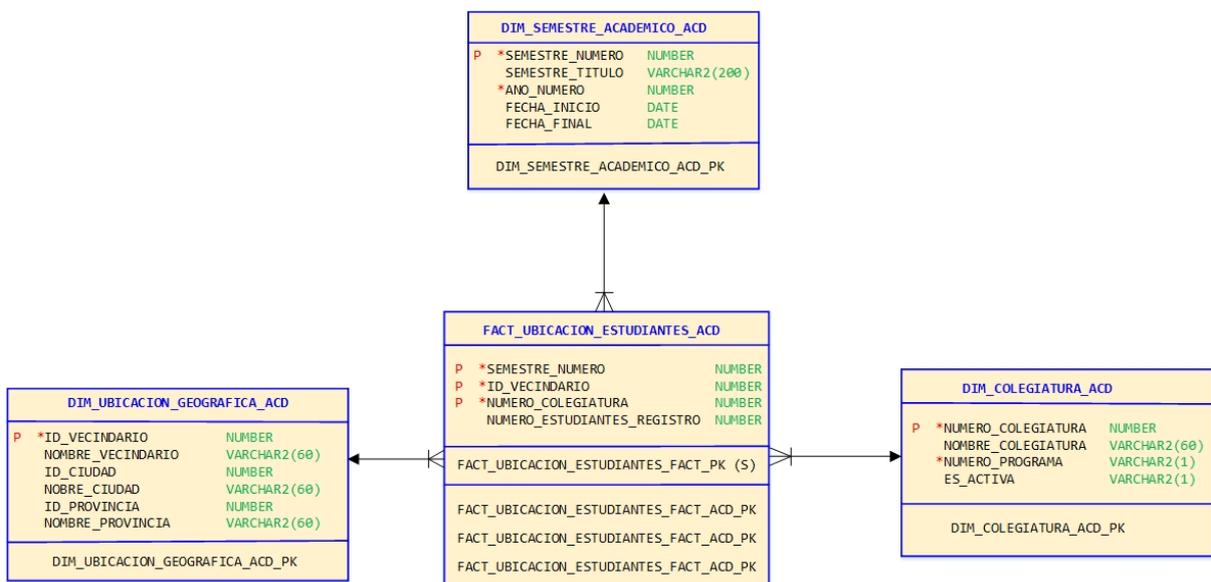


Figura 26. DataMart ubicaciones de los estudiantes. Fuente: Elaboración propia

Una vista maestra que une la tabla de hechos con las tablas de dimensiones se crea y se considera como una base para la vista lógica "V_VECINDARIO_ACD" como se muestra en la figura 27. Esta última vista calcula el recuento total de estudiantes registrados independientemente de todas las colegiaturas. Esta vista fue la base para calcular el recuento total de estudiantes registrados para cada ciudad que se presentó en la vista "V_REGISTRO_CIUADAD_ACD". Finalmente, "V_REGISTRO_PROVINCIA_ACD" calcula el recuento total de estudiantes registrados categorizados por las provincias en el CEC de la Universidad Técnica de Machala. Todas estas tres vistas presentan una jerarquía.

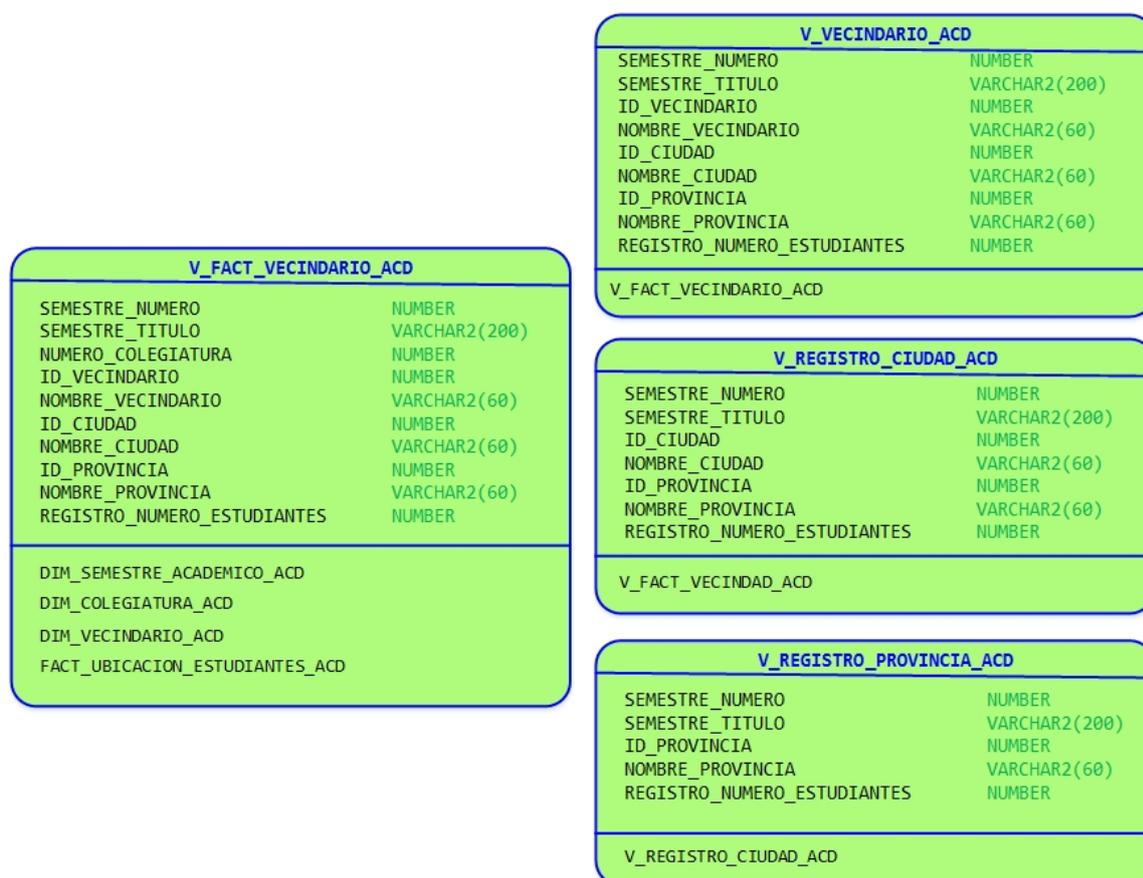


Figura 27. Vistas de ubicaciones de estudiantes. Fuente: Elaboración propia

DataMart promedio de estudiantes universitarios

Cada colegiatura en el CEC usa el promedio de los estudiantes como indicador principal de su desempeño. Sin embargo, este DataMart almacena cinco medidas, que son el recuento de estudiantes, su promedio de calificaciones en el semestre, sus horas promedio de éxito del semestre, las horas promedio de éxito para todos los semestres y, finalmente, el promedio de calificaciones de los estudiantes como se muestra en la figura 28. La tabla de hechos se refiere

por cuatro dimensiones que son año académico, semestre, departamento y nivel académico. En este sentido, se dispone de medidas HORAS_EXITO y SEMESTRE_HORAS_EXITO para escalar las medidas SEMESTRE_PROMEDIO en caso de cálculos jerárquicos. Se debe tener en cuenta que el nivel académico agregado a este DataMart es solo para uso futuro, actualmente no se considera un requisito comercial.

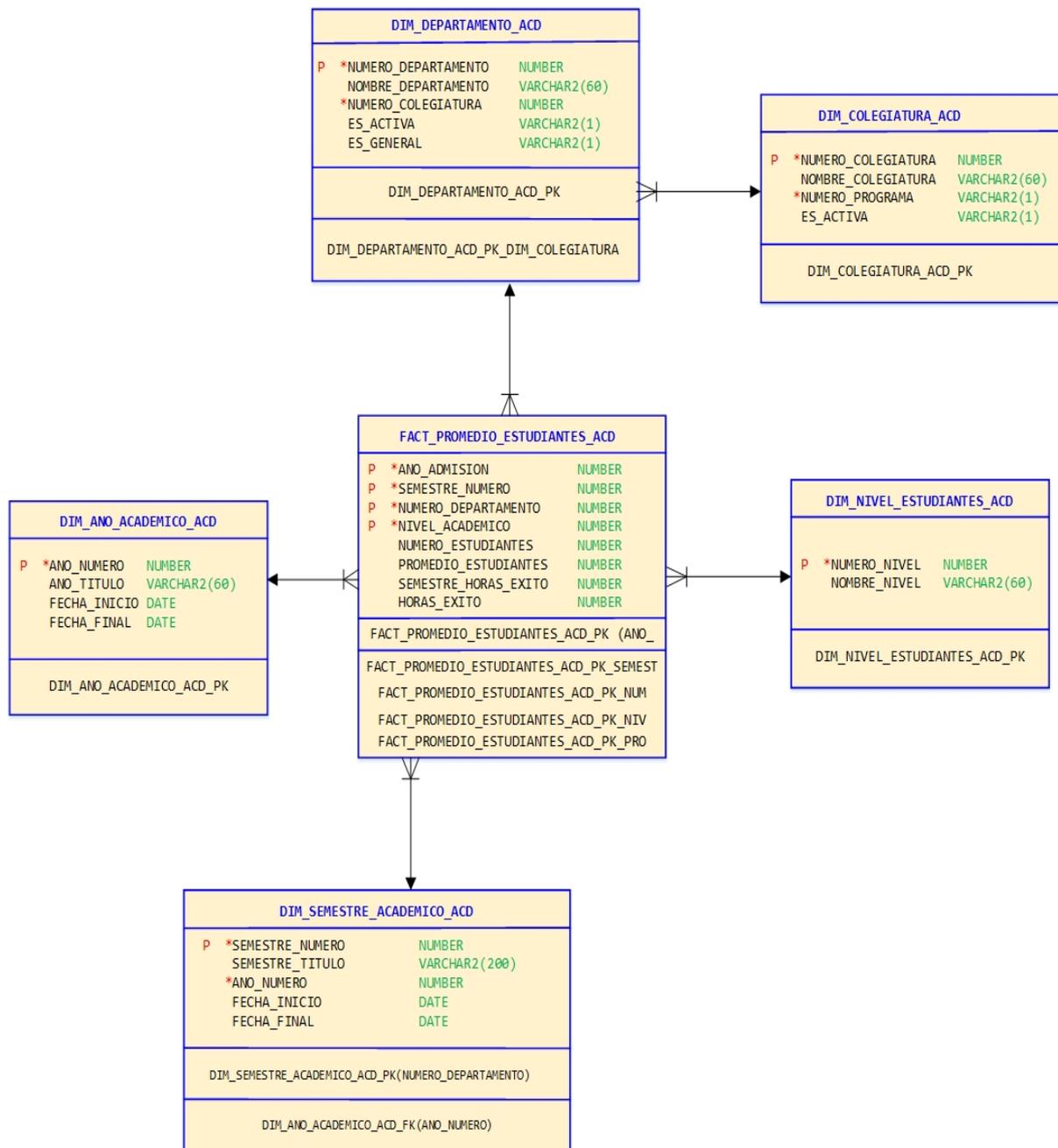


Figura 28. DataMart promedio del estudiante. Fuente: Elaboración propia

De este DataMart, se extrajo cuatro vistas lógicas que se calculan de antemano como se muestra en la figura 29. Estos resultados se pueden resumir de la siguiente manera:

- "V_FACT_PROMEDIO_ESTUDIANTE_ACD" muestra el recuento total de estudiantes y su promedio para cada año de admisión y semestre separados por departamentos.
- "V_PROMEDIO_COLEGIATURA_ACD" calcula el recuento total de estudiantes y su promedio para cada semestre y cada colegiatura.

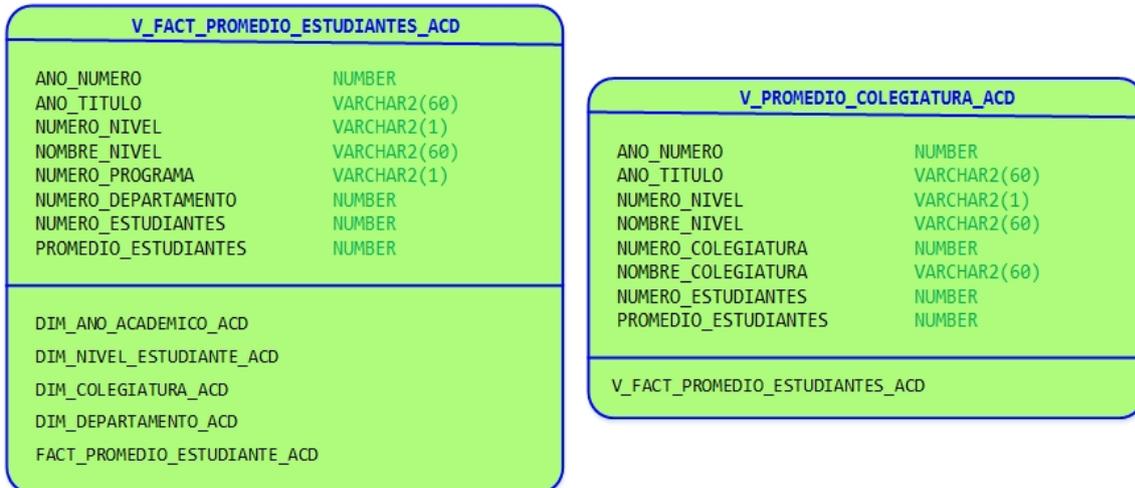


Figura 29. Vistas de promedios de colegiatura estudiantes. Fuente: Elaboración propia

4.4.5.- Paquete de trabajo 5: Divulgación y explotación

Al margen de que el equipo de trabajo junto con el Patrocinador tengan la intención de publicar los resultados de la solución en medios de comunicación, de impacto para el área de estudio, es importante que los resultados sean útiles para usuarios, profesionales o empresas, en este sentido el equipo de trabajo debe contemplar la difusión de los mismos a distintos colectivos.

Plan de divulgación:

A continuación, la tabla 24 muestra una lista de los mecanismos de divulgación posibles en general para la solución del DataMart para la toma de decisiones y la inteligencia de negocios en la Universidad Técnica de Machala, proyecto que permitirá mejorar la administración del CEC para realizar procesos de matriculación, registro, consulta de calificaciones e información financiera de los estudiantes.

Tabla 24. *Divulgación de la solución*

Nº	Medio, acción o actividad
1	Publicación mimetizada en revistas científicas
2	Contribución para congresos científicos
3	Talleres o conferencias de formación
4	Participación en congresos o foros profesionales
5	Participación en ferias de nuevas tecnologías nacionales e internacionales
6	Publicar en revistas sectoriales, empresariales o profesionales
7	Elaborar metodologías, normativas o reglamentos
8	Folleto explicativos de la solución o también notas de prensa
9	Convenios con entidades privadas y públicas para la divulgación
10	Divulgación en la web

Fuente: Elaboración propia

Plan de explotación:

Los aspectos más relevantes que se establecen en este tipo de planificación son los que muestra la tabla 25, en referencia a las actividades, medios o acciones de explotación.

Tabla 25. *Explotación de la solución*

Nº	Medio, acción o actividad	Recurso
1	Identificación de conocimientos técnicos y científicos derivados de la solución	Relevancia de innovación tecnológica con el DataMart y la inteligencia de negocios para la toma de decisiones, metodología SCRUM, PMBOK. Ventajas en la toma de decisiones y la administración de la información generada por el CEC Desarrollo en la inteligencia y memoria del CEC
2	Interés del equipo de trabajo para la explotación de la solución, congruencia en su aporte con el proyecto	Transferencia del conocimiento con respecto a la solución del DataMart y la inteligencia de negocios para el administración de la información del CEC
3	Previsión para el resguardo de la información derivados del desarrollo y consecución de la solución	Gestión de la propiedad intelectual, patentes, marcas, modelos de utilidad, know-how, entre otros. Titularidad de la solución para el equipo de trabajo, derechos de explotación y seguridad del proyecto
4	Identificación del mercado	Usuarios potenciales, así como entidades que quieran replicar la explotación del proyecto, además de otros sectores interesados en la tecnología como los institutos educativos de la provincia y las universidades del Ecuador
5	Acuerdos con entidades privadas y públicas	Negociación de convenios para la transferencia o cooperación en el ámbito de la inteligencia de negocios
6	Planteamiento de una expansión o versión, o también comercialización	Desarrollos y tiempos estimados hasta la implantación definitiva de la solución. Elaboración de ofertas para la explotación del proyecto a sectores potenciales. Elaboración de propuestas a entidades financiadoras para establecer procesos ulteriores de comercialización del proyecto

Fuente: Elaboración propia

4.4.6.- Cronograma

La tabla 26 muestra el cronograma del diseño del DataMart para el CEC de la Universidad Técnica de Machala.

Tabla 26. Cronograma de actividades

Fase (Actividad)	Duración (Días)	Mes 1				Mes 2				Mes 3				Mes 4			
		Semana				Semana				Semana				Semana			
		1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Recolección de Información	20	█	█	█													
Digitalización de registros		█	█														
Digitalización de información				█													
Planeación	30				█	█	█	█									
Análisis de la metodología de Kimball (iteraciones)					█	█	█	█									
Desarrollo del plan de trabajo						█											
Selección de roles							█										
Asignación de presupuesto								█	█								
Diseño	40								█	█	█	█	█				
Determinar los requerimientos del DataMart									█								
Definición y diseño del modelo dimensional										█							
Abstracción atributos medibles – Tabla de hechos											█	█					
Asignación de las dimensiones al modelo dimensional												█	█				

(Continuación)

4.5.- Estructura de gestión

La estructura de gestión del diseño del DataMart para la toma de decisiones y la inteligencia de negocios del Centro de Educación Continua (CEC) de la Universidad Técnica de Machala, está constituida por los ítems que muestra la tabla 27.

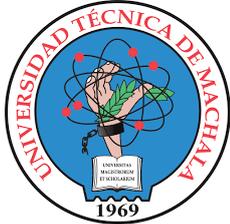
Tabla 27. Estructura de gestión del proyecto

Estructura de gestión	
1.-	Acta de constitución del proyecto
2.-	Declaración del alcance del proyecto
3.-	Registro de interesados
4.-	Plan de gestión del proyecto
4.1.-	Plan de gestión de la integración
4.2.-	Plan de gestión del alcance
4.3.-	Plan de gestión de requisitos
4.4.-	Plan de gestión del tiempo
4.5.-	Plan de gestión de costos
4.6.-	Plan de gestión de la calidad
4.7.-	Plan de gestión de las comunicaciones
4.8.-	Plan de gestión de riesgos
5.-	Cierre del proyecto
6.-	Plantillas del proyecto
REALIZADO POR:	APROBADO POR:

Fuente: Elaboración propia

4.6.- Plan de calidad

4.6.- PLAN DE CALIDAD

Código de documento:	CEC-UTM-DM-16-2020	
Fecha: noviembre 2019	Versión: V 1.0	
Proyecto:	Diseño del DataMart	
Unidad:	Área Administrativa del CEC	
Autora:	Sally Astrid Arreaga Santacruz	
Patrocinador:	_____	

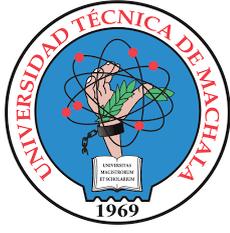
Política de calidad de la solución				
La calidad de la solución va direccionada a cumplir los plazos establecidos en el cronograma, el presupuesto dado e, implantar la solución de inteligencia de negocio que sirva para la toma de decisiones del CEC.				
Línea base de calidad de la solución				
Factor de calidad	Métrica de uso	Objetivo de calidad	Frecuencia/Tiempo de medición	Frecuencia/Tiempo de reporte
Rendimiento de la solución en términos de costo	Índice rendimiento de costo	CPI \geq 95%	Semanal/Viernes	Mensual/Febrero
Rendimiento de la solución en términos de tiempo	Índice rendimiento de tiempo	SPI \geq 95%	Semanal/Viernes	Mensual/Febrero
Satisfacción de la herramienta seleccionada	Nivel de satisfacción = Cumplimiento de al menos el 80% de los requisitos	Satisfacción \geq 95%	Bisemanal / (1.-Acta de constitución; 2.- Implantación de prueba))	Bisemanal / (1.- Informe de avance; 2.- Informe de prueba)
Plan de mejora de los procesos				
Si algún indicador de la calidad no se ubica en el nivel adecuado, es necesario mejorar el proceso; en este sentido se debe realizar las siguientes acciones:				
1.- Definir el proceso				
2.- Identificar la oportunidad para mejora				
3.- Analizar la información del proceso que haya influido en el rendimiento.				

(Continuación)

(Continuación)

4.- Plantear las medidas que permitan mejorar el proceso 5.- Implementar las acciones de mejora del proceso 6.- Controlar las acciones y sus resultados	
Roles para la gestión de calidad	
Patrocinador	Aprueba y controla las medidas correctivas frente a desvíos de las normas de calidad al Director del proyecto. Es importante su capacidad por depurar los indicadores que más beneficien a la solución y la inteligencia de negocios.
Líder del proyecto	Ejecuta el plan de calidad del proyecto y vigila su cumplimiento; ejecuta las acciones de mejora cuando las desviaciones de la calidad lo ameriten; revisa los entregables de calidad.
Equipo del proyecto	Mide y desarrolla los entregables de la calidad; con las especificaciones del Director del proyecto y los recursos correspondientes.
Organigrama de la calidad de la solución	
<pre> graph TD Patrocinador[Patrocinador] --- CC[Comisión control calidad] Patrocinador --- LP[Líder del proyecto] LP --- EQ[Equipo del proyecto] </pre>	
Procesos de gestión de la calidad	
Enfoque de aseguramiento de la calidad	En la trayectoria del desarrollo de la solución existirá un control constante de la calidad y rendimiento del equipo de trabajo, a través de los indicadores correspondientes. El monitoreo constante erradica desviaciones y ayuda aplicar el plan de mejora de procesos, para ello se utiliza la hoja de solicitud y control de cambios
Enfoque de control de calidad	Recibe la aceptación del Director del proyecto a la revisión de entregables, y tras la aprobación del Patrocinador. Las desviaciones de la calidad serán corregidas y notificadas en los entregables efectuando el aseguramiento de la calidad.
Enfoque de mejora de procesos	Se debe aplicar el plan de mejora de procesos de la presente planificación.

Matriz de actividades de control:

Código de documento:	CEC-UTM-DM-17-2020	
Fecha: noviembre 2019	Versión: V 1.0	
Proyecto:	Diseño del DataMart	
Unidad:	Área Administrativa del CEC	
Autora:	Sally Astrid Arreaga Santacruz	
Patrocinador:	_____	

Entregable	Estándar de calidad aplicable	Acciones de prevención	Acciones de control
1.1.1.- Acta de constitución del proyecto	Uso de la plantilla	Asesoría interna y entidad de supervisión	Aprobación del Patrocinador, visto bueno del entregable
1.1.2.- Declaración del alcance	Uso de la plantilla	Asesoría interna y entidad de supervisión	Aprobación del Patrocinador, visto bueno del entregable
1.2.1.- Planes de gestión secundarios	Estándares de PMI	Asesoría interna y entidad de supervisión	Aprobación del Patrocinador, visto bueno del entregable
1.2.2.- Plantillas del proyecto	Estándares de PMI	Asesoría interna y entidad de supervisión	Aprobación del Director del proyecto, visto bueno del entregable
1.4.1.- Acta de cierre del proyecto	Uso de la plantilla	Asesoría interna y entidad de supervisión	Aprobación del Patrocinador, visto bueno del entregable
2.1.1.- Informes de avance	Estándares de PMI	Asesoría interna y entidad de supervisión	Aprobación del Patrocinador, visto bueno del entregable
2.1.5.- Informe de prueba	Información requerida por el Patrocinador	Asesoría interna por el Equipo de trabajo	Aprobación por el Director del proyecto
2.2.- Informe final	Información requerida por el Director del proyecto	Asesoría interna y entidad de supervisión	Aprobación del Patrocinador

4.6.1.- Comunicación y colaboración

La información genera y básica que genera el diseño del DataMart para el CEC y la toma de decisiones de inteligencia empresarial. El mayor interés es detallar aquella información periódica que resulta del avance, y eventos especiales de la solución. La tabla 28 muestra los lineamientos para la comunicación y colaboración de la solución tecnológica para la toma de decisiones del CEC de la Universidad Técnica de Machala.

Tabla 28. *Comunicación y colaboración en el proyecto*

Estrategias de comunicación	Medio	Acción
1.- Todos los miembros del equipo de trabajo deben ser proactivos y asumir su responsabilidad compartida de la efectividad de la solución	Prácticas de clima laboral e inteligencia emocional	Elaborar talleres de formación y capacitación en aspectos de inteligencia emocional y clima laboral.
2.- Se deben marcar objetivos concretos, con tiempos asignados de manera realista	Entregables del proyecto	Informar y comunicar de la planificación estratégica de la entidad
3.- Se debe tomar una decisión de cada tema u objetivo marcado con fecha límite de ejecución y responsable	Plan de gestión de proyectos	Socializar la planificación principal y planificaciones secundarias.
4.- Todo el equipo de trabajo debe escuchar y participar de forma equilibrada con sus opiniones e ideas	Informes de avances	Comunicar la planificación de riesgos, la planificación de calidad y la gestión del proyecto
5.- Rigor absoluto en cuanto a al cronograma y la celebración de reuniones de trabajo	Informes de avances	Socializar de los avances y acciones de cambio que desvíen la calidad de la solución

(Continuación)

(Continuación)

6.- Se deben planificar reuniones de trabajo operativas y tácticas, así como reuniones estratégicas y de evaluación de la solución; así como también la cohesión del equipo	Informes de actas de reuniones	Cumplir con el cronograma de la solución y con las actividades que se resuelvan en las reuniones del equipo de trabajo
7.- Establecer espacios de integración y colaboración del equipo de trabajo para cada planificación de la solución	Espacios de integración y convivencia	Potenciar e impulsar reuniones destinadas únicamente para la convivencia
8.- Información de actas de reuniones periódicas y comités	Informe de avances	Reconocer y gratificar los avances del proyecto de manera individual

Fuente: Elaboración propia

4.6.2.- Aseguramiento de la calidad

El modelo de aseguramiento de la calidad está basado en la metodología denominada PHVA (Planificar-Hacer-Verificar-Actuar):

Planificar: Establecer los procedimientos idóneos para conseguir resultados de acuerdo con los requerimientos del Patrocinador del proyecto, y trasladar esas necesidades a especificaciones. Es necesario determinar las desviaciones de las actividades de desarrollo de la solución.

Hacer: Implementar los procedimientos y la capacidad de los mismos para cumplir los requerimientos del Patrocinador de la solución.

Verificar: Realizar el seguimiento y la medición de procedimiento respecto a los requerimientos para una solución eficaz e informar sobre los resultados.

Actuar: Tomar acciones para mejorar continuamente el desempeño de los procesos e institucionalizar las mejoras logradas.

El aseguramiento de la calidad pretende dar confianza en que la solución reúne las condiciones necesarias para satisfacer los requerimientos de la inteligencia de negocios, por

tanto; para asegura la calidad el Comité de calidad deberá realizar un conjunto de actividades que sirven para:

- Reducir, eliminar y prevenir las deficiencias de calidad de la solución tecnológica.
- Alcanzar una razonable confianza en que las características de la solución tecnológica sirvan para satisfacer los requerimientos de las partes interesadas.

Las actividades necesarias por el Comité de aseguramiento de la calidad, están estipuladas por las funciones:

- Identificar las posibles desviaciones en los estándares aplicados, así como en los requisitos y procedimientos especificados.
- Comprobar que se haya llevado a cabo las medidas preventivas o correctivas necesarias.

En este sentido, la tabla 29 muestra las diferentes etapas que constituyen la planificación del aseguramiento de la calidad.

Tabla 29. *Planificación del aseguramiento de la calidad*

Tarea	Medio	Técnicas y prácticas	Participantes
1.- Identificación de las propiedades de la calidad			
Constitución del Comité de aseguramiento de la calidad	Equipo de trabajo de la solución Planificación de aseguramiento de la calidad	Planificación de aseguramiento de la calidad Reuniones de trabajo	Director del proyecto y Comité de aseguramiento de la calidad
Determinación del proceso y solución de aseguramiento de la calidad	Proceso y solución objeto del aseguramiento de la calidad	Reuniones de trabajo	Director del proyecto y Comité de aseguramiento de la calidad
Identificación de las propiedades de la calidad	Proceso y solución objeto del aseguramiento de la calidad o propiedades de la calidad	Reuniones de trabajo	Director del proyecto y Comité de aseguramiento de la calidad

(Continuación)

(Continuación)

Tarea	Medio	Técnicas y prácticas	Participantes
2.- Establecimiento de la planificación del aseguramiento de la calidad			
Necesidad de la planificación de la calidad con alternativas propuestas	Solución objeto del aseguramiento de la calidad	Reuniones de trabajo	Comité de aseguramiento de la calidad
Alcance del aseguramiento de la calidad	Planificación del aseguramiento de la calidad para cada alternativa	Reuniones de trabajo	Comité de aseguramiento de la calidad
Impacto en el costo de la solución	Valoración de alternativas o costo de la planificación de aseguramiento de la calidad	Análisis costo-beneficio	Comité de aseguramiento de la calidad y Director del proyecto
3.- Adecuación de la planificación de aseguramiento de la calidad			
Ajuste de la planificación de aseguramiento de la calidad	Plan de calidad y aseguramiento de la calidad de la alternativa elegida	Reuniones de trabajo	Comité de aseguramiento de la calidad y Director del proyecto
Aprobación de la planificación de aseguramiento de la calidad	Registro de aprobación o rechazo	Reuniones de trabajo	Comité de aseguramiento de la calidad
4.- Especificación inicial de la planificación de aseguramiento de la calidad			
Definición de las estrategias de aseguramiento de la calidad para la solución	Estrategias de aseguramiento de la calidad o aspectos generales	Reuniones de trabajo	Comité de aseguramiento de la calidad y Director del proyecto

(Continuación)

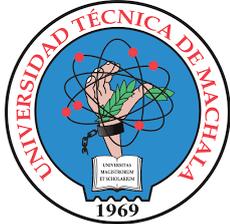
(Continuación)

Tarea	Medio	Técnicas y prácticas	Participantes
5.- Especificación detallada de la planificación de aseguramiento de la calidad			
Contenido de la plan de aseguramiento de la calidad para la solución	Plan de aseguramiento de la calidad	Reuniones de trabajo	Comité de aseguramiento de la calidad y Director del proyecto
6.- Revisión del análisis de consistencias			
Revisión de las especificaciones de requisitos	Plan de calidad de Revisión de requerimientos	Revisión de carácter técnico	Comité de aseguramiento de la calidad
Revisión de la consistencia entre procesos del desarrollo de la solución	Plan de calidad de Revisión de consistencia entre los procesos	Revisión de carácter técnico	Comité de aseguramiento de la calidad
7.- Revisión de la planificación de pruebas			
Revisión de la planificación de pruebas	Planificación de aseguramiento de la calidad, planificación de pruebas	Revisión de carácter técnico	Comité de aseguramiento de la calidad

Fuente: Elaboración propia

4.7.- Plan de gestión del riesgo

4.7.- PLAN DE GESTIÓN DE RIESGOS

Código de documento:	CEC-UTM-DM-18-2020	
Fecha: noviembre 2019	Versión: V 1.0	
Proyecto:	Diseño del DataMart	
Unidad:	Área Administrativa del CEC	
Autora:	Sally Astrid Arreaga Santacruz	
Patrocinador:	_____	

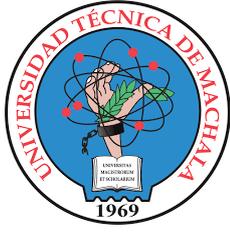
Metodología de gestión de riesgos			
Proceso	Descripción	Herramientas	Fuentes de información
Planificación de gestión de riesgos	Consiste en planificar la gestión de los riesgos	Guía de PMBOK	Patrocinador, Director del proyecto, Equipo del proyecto
Identificación de los riesgos	Determinar los riesgos que afecten al proyecto positiva o negativamente	Bitácora de riesgos	Patrocinador, Director del proyecto, Equipo del proyecto
Análisis cualitativo del riesgo	Evaluar el impacto de la materialización del riesgo y jerarquía de importancia	Matriz de probabilidad e impacto	Patrocinador, Director del proyecto, Equipo del proyecto
Planificación de respuesta al riesgo	Establecer respuestas a la materialización de los riesgos	Guía de PMBOK	Patrocinador, Director del proyecto, Equipo del proyecto
Seguimiento y control del riesgo	Corroborar la ocurrencia de un riesgo y supervisar la ejecución de respuestas	Guía de PMBOK	Patrocinador, Director del proyecto, Equipo del proyecto

(Continuación)

(Continuación)

Periodicidad de la Gestión de riesgos			
Proceso	Momento de ejecución	Entregable de la EDT	Periodicidad de ejecución
Planificación de gestión de riesgo	Al comenzar el proyecto	Plan de gestión de riesgo	Una sola vez
Identificación de los riesgos	Al empezar el proyecto con revisiones frecuentes en las reuniones del equipo de trabajo	Reuniones al ajustar el mes	Bisemanal
Análisis cualitativo de los riesgos	Al empezar el proyecto con revisiones frecuentes en las reuniones del equipo de trabajo	Informe en las reuniones ajustadas a cada mes	Bisemanal
Planificación de respuesta al riesgos	Al empezar el proyecto con revisiones frecuentes en las reuniones del equipo de trabajo	Informe en las reuniones ajustadas a cada mes	Bisemanal
Monitoreo y control de los riesgos	Periódicamente pasadas las dos semanas	Informe en las reuniones ajustadas a cada mes	Bisemanal
Formatos de la Gestión de riesgos			
Planificación de la gestión de riesgos	Plan de gestión de riesgos		
Identificación de riesgos	Caracterización y evaluación cualitativa de los riesgos		
Análisis cualitativo de los riesgos	Caracterización y evaluación cualitativa de los riesgos		
Planificación de la respuesta ante el riesgo	Plan de respuestas frente a los riesgos		
Monitoreo y control de los riesgos	1.- Informe de monitoreo de los riesgos. 2.- Solicitud de los cambios 3.- Cambios preventivos y correctivos		

Identificación de Riesgos:

Código de documento:	CEC-UTM-DM-19-2020	
Fecha: noviembre 2019	Versión: V 1.0	
Proyecto:	Diseño del DataMart	
Unidad:	Área Administrativa del CEC	
Autora:	Sally Astrid Arreaga Santacruz	
Patrocinador:	_____	

Código	Descripción	Causa	Proceso afectado	Probabilidad de ocurrencia	Grado de impacto
PGR01	Implantación del material	Las características institucionales favorables al Patrocinador	2.0	No frecuente	Crítico
PGR02	Resistencia de manera interna por el cambio	No tomar en cuenta las normativas internacionales	2.1	Ocasional	Menor
PGR03	Capacitación deficiente	Escaso tiempo para realizar las capacitaciones del proyecto	1.2	Frecuente	Despreciable
PGR04	Expectativas erróneas de resultados	Deficiente comunicación entre las partes interesadas del proyecto	1.2	Poco frecuente	Serio
PGR05	Alta demanda de los supervisores	Excelente admisión de la solución en su periodo de prueba	1.1	Frecuente	Moderado

(Continuación)

(Continuación)

Código	Descripción	Causa	Proceso afectado	Probabilidad de ocurrencia	Grado de impacto
PGR06	Aplazamiento de algún entregable	Responsabilidad de las partes interesadas	Toda la solución	Ocasional	Serio
PGR07	Retraso en la divulgación y explotación de la solución	Proceso interno en las políticas de la entidad	2.1 y 2.2	Ocasional	Medio
PGR08	Deficiente ejecución de la solución	Deficiente planificación y comprensión de los procesos	Todo el proyecto	Frecuente	Crítico
PGR09	Limitada usabilidad de la solución	Escasa formación en el uso y beneficios de la solución	2.2	Posible	Serio

Plan de respuesta a los riesgos:

Código de documento:	CEC-UTM-DM-20-2020	
Fecha: noviembre 2019	Versión: V 1.0	
Proyecto:	Diseño del DataMart	
Unidad:	Área Administrativa del CEC	
Autora:	Sally Astrid Arreaga Santacruz	
Patrocinador:	_____	

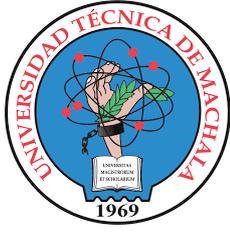
Código	Tipo de riesgo	Acción y requerimiento	Nivel de riesgo	Línea base afectada	Responsable
PGR01	Amenaza	Análisis de las características internas y la oportunidad de uso de la solución	10	Alcance	Entidad supervisora
PGR02	Amenaza	Comunicación interna del alcance en la operatividad y beneficios del CEC	6	Tiempo	Líder del proyecto
PGR03	Amenaza	Plan de seguimiento de las actividades de la solución	4	Calidad	Entidad supervisora
PGR04	Amenaza	Disponibilidad para resolver dudas y detalles en los avances de la solución	8	Alcance	Director del proyecto
PGR05	Oportunidad	Ampliación del periodo de usabilidad y prueba de la solución	12	Tiempo y alcance	Director del proyecto

(Continuación)

(Continuación)

Código	Tipo de riesgo	Acción y requerimiento	Nivel de riesgo	Línea base afectada	Responsable
PGR06	Amenaza	Adecuaciones en el cronograma por cambios en el equipo u peticiones del Director o Patrocinador del proyecto	12	Tiempo	Director del proyecto
PGR07	Amenaza	Apertura de iniciativas en tecnología de forma oportuna y programada	9	Tiempo	Equipo del proyecto
PGR08	Amenaza	Apoyo de la entidad supervisora al proceso de avance y ejecución	25	Todas	Entidad supervisora
PGR09	Amenaza	Empoderamiento del equipo de trabajo en consenso a la entidad supervisora por los beneficios de la solución	12	Alcance	Patrocinador

Matriz de probabilidad e impacto:

Código de documento:	CEC-UTM-DM-21-2020	
Fecha: noviembre 2019	Versión: V 1.0	
Proyecto:	Diseño del DataMart	
Unidad:	Área Administrativa del CEC	
Autora:	Sally Astrid Arreaga Santacruz	

Medidas cualitativas de probabilidad				
Nivel	Rango	Descriptor	Escala	Cualidad
5	ALTO	Muy frecuente	Semanal	Se espera que suceda en casi todas las circunstancias
4		Frecuente	Bisemanal	Probablemente suceda en casi todas las circunstancias
3	MEDIO	Ocasional	Mensual	Puede suceder en algún momento de la solución
2	BAJO	Poco frecuente	Trimestral	Puede suceder en algún momento de la solución
1		No frecuente	Anual	Puede suceder únicamente en circunstancias excepcionales

Medidas cualitativas del impacto							
Nivel	Rango	Descriptor	Efectos	Variables			
				Costo	Tiempo	Alcance	Calidad
5	ALTO	Crítico	No reparables	Incremento del costo > 20%	Desviación de la solución > 20%	Solución 70% inservible	Solución 70% inservible
4		Serio	Efectos extensivos	Incremento del costo > 11% y < 20%	Desviación de la solución 11 - 20%	Reducción del alcance <70% - 50%	Reducción de la calidad <70% - 50%
3	MEDIO	Moderado	Efectos considerables	Incremento del costo > 5% y < 11%	Desviación de la solución <5%	Reducción del alcance <50% - 25%	Reducción de la calidad <50% - 25%
2	BAJO	Menor	Efectos mínimos	Incremento del costo < 5%	Variación del calendario <5%	Reducción del alcance <25% - 5%	Reducción de la calidad <25% - 5%
1		Despreciable	Efectos exigüos	Insignificante incremento	Insignificante variación	Reducción del alcance <5%	Reducción de la calidad <5%

Matriz de análisis de riesgos cualitativos:

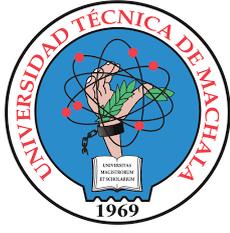
Código de documento:	CEC-UTM-DM-22-2020	
Fecha: noviembre 2019	Versión: V 1.0	
Proyecto:	Diseño del DataMart	
Unidad:	Área Administrativa del CEC	
Autora:	Sally Astrid Arreaga Santacruz	

Impacto				
Despreciable	Menor	Moderado	Serio	Crítico

Probabilidad	Muy frecuente	5	10	15	20	25
	Frecuente	4	8	12	16	20
	Ocasional	3	6	9	12	15
	Poco frecuente	2	4	6	8	12
	No frecuente	1	2	3	4	5

Leyenda	
	Riesgo grave: Necesita medidas preventivas urgentes, no se debería iniciar la solución tecnológica sin mitigar el riesgo
	Riesgo importante: Medidas preventivas obligatorias, se deben mitigar los riesgos durante el desarrollo de la solución
	Riesgo apreciable: Introducir medidas preventivas para reducir posibles niveles de riesgo.

Hoja de monitoreo de los riesgos:

Código de documento:	CEC-UTM-DM-23-2020	
Fecha: noviembre 2019	Versión: V 1.0	
Proyecto:	Diseño del DataMart	
Unidad:	Área Administrativa del CEC	
Autora:	Sally Astrid Arreaga Santacruz	

Código	Tipo de riesgo	Estado	Acción	Seguimiento	Indicador
ELABORADO POR:			APROBADO POR:		

4.8.- Estrategia de divulgación

El diseño del DataMart para la toma de decisiones en el Área Administrativa del CEC de la Universidad Técnica de Machala, tiene como población objetivo a docentes, estudiantes y la comunidad en general, esto con el propósito de informar y generar interactividad con los diferentes medios de divulgación que permiten conocer sobre la solución y proyecto tecnológico, en este sentido la tabla 30 permite conocer la estrategia de divulgación. La divulgación es muy importante para transferir el conocimiento aprendido sobre los DataMart como estrategia para una gestión empresarial inteligente y la administración de proyectos tecnológicos de forma planificada y organizada.

Tabla 30. Estrategia de divulgación

Estrategia	Medio	Acción
Promoción informativa digital	Página web institucional, correo electrónicos	Noticias e información sobre el proyecto tecnología del DataMart
Programa de Radio y televisión	Radios y televisoras locales	Información en espacios publicitarios sobre el proyecto tecnológico
Gestión de redes sociales	Redes sociales Facebook, Twitter, Instagram, Youtube.	Publicación de información y mensajes de interés a la comunidad sobre el proyecto tecnológico
Artículos de prensa	Presa escrita y digital con información del proyecto tecnológico	Divulgación de estrategias realizadas en referente al proyecto tecnológico
Reuniones presenciales	Equipo de trabajo del proyecto tecnológico	Socialización de actividades y soluciones referentes al DataMart
Bitácora y blogs en la web	Diseño de blogs y portales para divulgación de información	Blog de docentes y blog de estudiantes. Página web de la Universidad
Otros medios	Memorandos Escritos Afiches y publicidad personalizada	Evidencias el impacto del proyecto tecnológico.

Fuente: Elaboración propia

4.9.- Estrategia de explotación de resultados

La explotación de resultados del diseño del DataMart para la toma de decisiones en el Área Administrativa del CEC de la Universidad Técnica de Machala, enmarca beneficios de la innovación tecnológica del CEC, porque actualmente la entidad mantiene la toma de decisiones de manera improvisada e imprevista, lo que no ocurre con los reportes y la información que deriva de las herramientas de inteligencia de negocios y del DataMart, para ello la explotación de resultados mantiene la estrategia que detalla la tabla 31 con su respectivo medio y acción.

Tabla 31. Estrategia de explotación de resultados

Estrategia	Medio	Acción
Valoración de prestaciones	Generación de conocimiento organizacional	El proyecto tecnológico del DataMart supone una mejora en la gestión de información e innovación hacia la inteligencia de negocios
Estimación de la calidad	Perfil profesional del estudiante	La educación superior y formación permanente mantiene garantía en la excelencia y mayor seguridad en la gestión de la información
Innovación y capacidad tecnológica	Sistema informático y base de datos eficientes.	Incremento del capital tecnológico y los canales de absorción de la tecnología
Reducción de costes	Colegiaturas de inscripciones y matriculas efectivos	El DataMart propone un manejo del ámbito financiero con mayor precisión para el cobro de rubros y colegiaturas
Impacto laboral y ambiental	Oportunidades de mejora profesional y laboral	Generación de empleo y estrategias de conservación del medio ambiente
Relación institucional	Convenios institucionales	Realización de prácticas profesionales en el campo real de trabajo
Posición competitiva en el ámbito académico	Aulas virtuales, evaluaciones digitales	Construcción de recursos tecnológicos ligados al DataMart
Desarrollo regional	Calidad institucional	Mejora de infraestructura tecnológica, empleo cualificado, preparación docente, etc.
Medio ambiente y calidad de vida	Universidad ecológica y cero desperdicio de recursos	Incorporación de una filosofía respetuoso del medio ambiente y el uso de la tecnología para cuidar los ecosistemas y la vida

Fuente: Elaboración propia

4.9.1.- Plan de internacionalización de los resultados

El diseño del DataMart para la toma de decisiones en el Área Administrativa del CEC de la Universidad Técnica de Machala, además de las consecuentes planificaciones, debe tener un plan de internacionalización de los resultados, para ello la tabla 32 muestra las líneas de actuación con su respectivo medio y acción a ejecutar.

Tabla 32. *Internacionalización de los resultados*

Línea de actuación	Medio	Acción
Gestión de la información	Prácticas de gestión del conocimiento	Participación activa en redes relativas al posicionamiento y visibilidad internacional de los Data Warehouse y DataMart para el negocio.
Movilidad internacional	Redes internacionales de investigadores	Formación internacional de docentes y estudiantes para la transferencia de conocimiento
Alianzas y convenios	Redes de universidades y centros de investigación.	Publicar resultados de proyectos tecnológicos y nuevas investigaciones en el área de la inteligencia de negocios en alianzas internacionales.
Visibilidad internacional	World-class University	Divulgación del proyecto tecnológico sobre la inteligencia de negocio con el DataMart para formarse e investigar en el tema, de manera internacional
Capacidades de idiomas	Dominio de idiomas extranjeros	Mejorar las capacidades interculturales de los estudiantes y docentes para originar una conexión internacional en avances de proyectos tecnológicos
Transferencia del conocimiento e innovación	Hábitat de tecnología y ciencia	Potenciar e impulsar la formación internacional, apoyando a investigadores en trabajos de la inteligencia de negocios DataMart
Conferencias, seminarios, publicaciones de artículos y papers internacionales	Jornadas internacionales de divulgación de conocimiento	Reconocimiento del proyecto tecnológico involucrado en los Data Warehouse y DataMart internacionalmente, además de otros programas y proyectos internacionales para el desarrollo de la investigación y formación profesional.

Fuente: Elaboración propia

4.9.2.- Plan de gestión de la propiedad intelectual

El plan de gestión de la propiedad intelectual es una planificación para conseguir valor del conocimiento, tecnología y solución de las actividades innovadoras. En este sentido, existen dos directrices importantes en la gestión de la propiedad intelectual, que son:

- Generar el conocimiento para poder ser transmitido.
- Utilizar el conocimiento para consultar la productividad y competitividad.

En congruencia a estas directrices, se plantean objetivos estratégicos que coadyuvan con la inteligencia de negocios en la ciencia, tecnología e innovación, como detalla la matriz de objetivos estratégicos en la tabla 33.

Tabla 33. *Objetivos estratégicos de gestión de la propiedad intelectual*

Objetivo a obtener	Acciones y elementos de PI	Situación identificada
Fortalecer el conocimiento de la ciencia, tecnología e innovación, apoyando iniciativas tendientes a su consolidación	Entidades coordinadoras en materia de PI, que pueda adoptar políticas y lineamientos. Promoción de la PI para consolidar los recursos en favor de la innovación	Carencia de políticas eficaces para facilitar los derechos en materia de PI
Fortalecer el capital humano en tecnología y ciencia para generar conocimiento e innovación	Capacitación en el uso de patentes y marcas como fuente de información tecnológica	Ausencia de recurso humano capacitado en propiedad intelectual, ciencia y tecnología.
Facilitar la apropiación social de la tecnología, ciencia e innovación para su divulgación	Automatización de procesos y establecimiento de sistemas de información referente a la PI	La información utilizable en materia de PI rezaga el valor agregado
Fomentar un sistema de gestión de los activos intangibles producto de la PI	Promover el uso de contratos de participación o procedimientos y trámites de licencias para monitorear las actividades de PI	Escasa utilización de las normativas de PI cohibiendo el desarrollo económico, tecnológico e inventivo

Fuente: Elaboración propia

En base a los objetivos estratégicos de la planificación de PI, es necesario seguir cuatro pasos importantes, como son:

1.- Identificar activos: Es importante la identificación del conocimiento protegible durante los procesos de investigación, así como en los resultados para propiciar una transferencia tecnológica relacionada con la productividad y competitividad.

2.- Protección: Es necesario asimilar en la protección de la PI, que lo que se transfiere es el derecho, no el conocimiento. El autor de una invención adquiere el derecho exclusivo para realizar la explotación, en el territorio que le confiere el derecho. Entre las principales acciones de la protección se identifican: el ingreso a nuevos mercados, evitar la copia o reproducción, la creación de una marca e identidad, además de agregar valor a una organización.

3.- Valorizar: En la valoración interviene el término de la regalía que es el pago realizado al titular de los derechos de PI, por disponer de su innovación intelectual.

4.- Comercialización: Para este proceso de negociación se requiere comprender los asuntos relevantes de la tecnología y su potencial comercial por medio de la elaboración de una propuesta de valor, en el que se define el mecanismo de transferencia como licencia, franquicia o venta para negociar.

4.10.- Presupuesto

El presupuesto del diseño del DataMart para la inteligencia de negocios en el CEC de la Universidad Técnica de Machala matiza el uso de los recursos tecnológicos del entorno de desarrollo, así como la caracterización de los recursos de software y hardware, el recurso humano con el propósito de determinar el costo total de la solución. El rendimiento del proyecto, así como la cronología de las actividades que se ejecutan en cada proceso de diseño del DataMart, necesitan de los aspectos financieros para poder monitorear y controlar su cumplimiento, un proyecto tecnológico, al igual que la mayoría de proyectos de innovación requiere analizar la parte administrativa y la financiera para poder sustentarlo.

Recurso Humano:

Se considera como recurso primario al capital humano que interviene en cada una de las fases de la solución. La tabla 34 muestra el recurso humano que interviene en el diseño del DataMart para el CEC y la toma de decisiones.

Tabla 34. *Recurso Humano*

CARGO	NOMBRE	SUELDO TOTAL (€)
Director del proyecto	Jorge Rivas	4.000,00
Desarrollador	Sally Arreaga	2.900,00
TOTAL	2 personas	6.900,00

Fuente: Elaboración Propia

Caracterización del Hardware para el Desarrollo:

La caracterización del hardware para el desarrollo, se refiere al equipo que esta disponible para el diseño del DataMart para la toma de decisiones en el CEC. La tabla 35 muestra la caracterización del hardware para el desarrollo.

Tabla 35. *Caracterización del Hardware para el desarrollo*

HARDWARE	CARACTERÍSTICAS
Computador portátil HP Pavilion R2200s	
Disco Duro	1 TB HDD
Memoria RAM	4 GB DDR3
Procesador	Intel® Core™ i5
Display	14.0" 16:9 HD LED LCD
Otros	Memoria Gráfica: Intel® HD Graphics DVD-Super Multi DL drive Protocolo: 802.11b/g/n Batería: 18-cell Li-ion
PRECIO	€ 780,00

Fuente: Elaboración Propia

La tabla 36 muestra el hardware adicional para el desarrollo.

Tabla 36. Hardware adicional para el desarrollo

HARDWARE ADICIONAL PARA EL DESARROLLO	
DESCRIPCIÓN	VALOR (€)
Impresora EPSON TX 1200NT	90,00
TOTAL	90,00

Fuente: Elaboración Propia

La tabla 37 muestra el costo total del hardware para el desarrollo.

Tabla 37. Costo total del hardware para el desarrollo

COSTO TOTAL DEL HARDWARE PARA EL DESARROLLO	
DESCRIPCIÓN	COSTO TOTAL (€)
Hardware para el desarrollo	780,00
Hardware adicional para el desarrollo	90,00
TOTAL	870,00

Fuente: Elaboración Propia

Caracterización del Hardware para la Implantación:

La tabla 38 muestra la caracterización del hardware para la implantación

Tabla 38. Caracterización del Hardware para la implantación

HARDWARE	CARACTERÍSTICAS
Computador de escritorio HP TN5500s	
Disco Duro	1 TB HDD
Memoria RAM	2 GB DDR3
Procesador	Intel® Core™ i3
Monitor	17.0" 32:16 LCD
Otros	DVD-Super Multi DL drive Teclado, mouse, parlantes Protocolo: 802.11b/g/n
PRECIO	€ 670,00

Fuente: Elaboración Propia

La tabla 39 muestra el hardware adicional para la implantación.

Tabla 39. *Hardware adicional para la implantación*

HARDWARE ADICIONAL IMPLANTACIÓN	
DESCRIPCIÓN	VALOR (€)
Cable eléctrico de 2 pares 220V 1,5 A	40,00
Periféricos y almacenamiento	100,00
TOTAL	140,00

Fuente: Elaboración Propia

La tabla 40 muestra el costo total del hardware para la implantación.

Tabla 40. *Costo total del hardware para la implantación*

COSTO TOTAL DEL HARDWARE PARA LA IMPLANTACIÓN	
DESCRIPCIÓN	COSTO TOTAL (€)
Hardware para la implantación	670,00
Hardware adicional implantación	140,00
TOTAL	810,00

Fuente: Elaboración Propia

Caracterización del Software para el Desarrollo:

La tabla 41 muestra la caracterización del software para el desarrollo.

Tabla 41. *Caracterización del Software para el desarrollo*

SOFTWARE	DESCRIPCIÓN	VALOR (€)
Windows 8.1 Win. 32 Professional	Sistema Operativo	180,00
Oracle 11g Express Edition	Gestión de base de datos	FREE
Microsoft Office 2016:		
Word	Procesamiento de texto.	60,00
Visio	Diseño de gráficos de ingeniería.	
Access		
TOTAL		240,00

Fuente: Elaboración Propia

Caracterización del Software para la Implantación:

La tabla 42 muestra la caracterización del software para la implantación.

Tabla 42. *Caracterización del Software para la implantación*

SFTWARE	CARACTERÍSTICAS	VALOR (€)
Windows 10 Win. 64 Professional	Sistema Operativo	180,00
Pentaho 7.0 Suite	Herramienta de integración de datos y reportes	FREE
TOTAL		180,00

Fuente: Elaboración Propia

Capacitación del Recurso Humano:

La tabla 43 muestra la capacitación del recurso humano.

Tabla 43. *Capacitación del Recurso Humano*

RECURSO	DESCRIPCIÓN	DURACIÓN	TEMAS	VALOR (€)
1	Capacitación para la administración del proyecto	10 horas	- Inteligencia de negocios - Proyectos tecnológicos	500,00
2	Capacitación para operación del DataMart	10 horas	- Explotación de DataMart - Diseño de base datos y gestión de información	600,00
TOTAL				1.100,00

Fuente: Elaboración Propia

Para el diseño del DataMart para la toma de decisiones del CEC, se necesita dar capacitación al Capital Humano ya que al ser un proyecto de innovación tecnológica requiere de competitividad y horas de experiencia en ámbitos técnicos como la gestión de base de datos y software de inteligencia de negocios. La capacitación del capital humano es muy importante para mejorar la preparación del equipo de trabajo en las tareas que tienen bajo su responsabilidad, estas tareas son la base del desarrollo de la solución.

Costo del diseño de la base de datos y modelos multidimensionales:

El presupuesto contempla el coste total del diseño del DataMart así como un desglose de los conceptos económicos. Para el diseño de la base de datos y los modelos multidimensionales se ha tenido que realizar un estudio previo de algoritmos bosquejo antes de llegar a los modelos definitivos. La tabla 44 muestra el costo del diseño del DataMart en aristas como la base de datos y los modelos multidimensionales.

Tabla 44. Costo del diseño del DataMart

DESCRIPCIÓN	VALOR (€)
Líneas de Códigos (LCD)	690,00
Pago de sueldo del Desarrollador por 2 meses	5.800,00
Depreciaciones de los equipos (30% anual)	43,50
Depreciación por utilización de las licencias del software	13,75
TOTAL	6.547,25

Fuente: Elaboración Propia

Costo total del diseño del DataMart:

La tabla 45 muestra el costo total del diseño del DataMart para la toma de decisiones en el CEC de la Universidad Técnica de Machala.

Tabla 45. Costo total del diseño del DataMart

DESCRIPCIÓN	VALOR (\$)
Costo del diseño del DataMart	6.547,25
Software para la implantación	180,00
Hardware para la implantación	810,00
Capacitación del Capital Humano	1.100,00
Mobiliario	70,00
Imprevistos (10%)	870,72
TOTAL GENERAL	9.577,97

Fuente: Elaboración Propia

El presupuesto total del diseño del DataMart asciende a la cantidad de **NUEVE MIL QUINIENTOS SETENTA Y SIETE 97/100 DÓLARES**. Dinero financiado por la Administradora del proyecto y patrocinado por el CEC de la Universidad Técnica de Machala. Es necesario destacar que la inversión en el proyecto tecnológico ayudará a mejorar la capacidad del CEC.

4.11.- Impacto socioeconómico y medio ambiental

El TFM de proyecto tecnológico como solución al diseño de un DataMart se respalda en un análisis técnico de los impactos en el ámbito socioeconómico y medio ambiental; este análisis utiliza los siguientes criterios:

- Se selecciona un nivel de impacto positivo o negativo de acuerdo a un criterio de ponderación como se aprecia en la tabla 46.

Tabla 46. Rango de niveles de impacto positivo y negativo

Nivel de impacto	Estado
-3	Impacto ALTO negativo
-2	Impacto MEDIO negativo
-1	Impacto BAJO negativo
0	NO HAY IMPACTO
1	Impacto BAJO positivo
2	Impacto MEDIO positivo
3	Impacto ALTO positivo

Fuente: Elaboración propia

- Se construye para cada uno de los ámbitos de análisis una matriz en la que horizontalmente se ubica el nivel de impacto establecido, mientras que verticalmente se determina una serie de indicadores que permitirá tener información específica y puntual del área analizada.

Impacto socioeconómico:

La tabla 47 muestra las métricas del impacto socioeconómico.

Tabla 47. Métricas del impacto socioeconómico

Indicador	Nivel de impacto	-3	-2	-1	0	1	2	3	TOTAL
Inversión en conocimiento								x	3
Gestión de la información							x		2
Generación de productividad						x			1
Inteligencia de negocio								x	3
	TOTAL					1	2	6	9

Fuente: Elaboración propia

Total de impacto socioeconómico = 9/3

Total de impacto socioeconómico = 3

Nivel de impacto socioeconómico = MEDIO positivo.

Análisis

El diseño de un DataMart permite gestionar la información que se genera en el CEC de la Universidad Técnica de Machala como solución de la inteligencia de negocio para la toma de decisiones, este producto permite plantear estrategias importantes en la inversión en conocimiento que necesita la entidad para generar mayor y mejor productividad con respecto a la formación permanente de los estudiantes y futuros profesionales de la sociedad ecuatoriana; con una información más organizada y de rápido acceso los estudiantes pueden acceder a una educación de calidad, mejorar su perfil profesional y ser más competitivos para sus lugares de trabajo o emprendimientos, es beneficioso el impacto a nivel socioeconómico que incita la innovación tecnológica y la administración eficiente de la información para la toma de decisiones. En base a estos lineamientos el nivel de impacto socioeconómico es medio positivo; es decir, es favorable para la economía y para el desarrollo de la sociedad.

Impacto medioambiental

La tabla 48 muestra las métricas del impacto medioambiental.

Tabla 48. Métricas del impacto medioambiental

Indicador	Nivel de impacto	-3	-2	-1	0	1	2	3	TOTAL
Conservación del medio ambiente								x	3
Uso de documentación electrónica								x	3
Negocios inteligentes y saludables							x		2
Agricultura de precisión								x	3
	TOTAL						2	9	11

Fuente: Elaboración propia

Total de impacto medioambiental = 11/3

Total de impacto medioambiental = 3,66

Nivel de impacto medioambiental = ALTO positivo.

Análisis

El diseño de proyectos tecnológicos y entre ellos el DataMart o las Data Warehouse sirven para mejorar la gestión de la información que genera una entidad, en este caso el CEC de la Universidad Técnica de Machala; misma que de manera tradicional utiliza de manera

desmesurada el papel impreso para la generación de reportes o documentación, con los repositorios digitales esta agresiva contaminación ambiental se limita. La conservación del medio ambiente es una información en la que trabajan los proyectos tecnológicos, tales como la agricultura de precisión, las energías renovables, la reutilización y el reciclaje; esto se ha podido generar por el cruce de información en la red, por la negociación de empresas ecológicas con la industria y la manufactura tecnológica, los negocios se vuelven inteligentes y saludables, no se desperdicia y se mantiene amistad con los lineamientos del reciclaje y la conservación de recursos. En base a estos argumentos el nivel de impacto medioambiental dispone de una cualificación al alto positivo; es decir, es muy conveniente para la conservación del medio ambiente, prevención y protección del entorno laboral y ecológico.

Impacto en forma general:

La tabla 49 muestra las métricas del impacto en forma general

Tabla 49. Métricas del impacto en forma general

Indicador	Nivel de impacto	-3	-2	-1	0	1	2	3	TOTAL
Impacto socioeconómico								x	3,00
Impacto medioambiental								x	3,66
	TOTAL							6	6,66

Fuente: Elaboración propia

$$\sum \text{Nivel de impacto} / \text{No Indicadores}$$

$$6,66/2 = 3,33$$

Análisis

El impacto de forma general en estos ámbitos de análisis como es el socioeconómico y medioambiental permiten reflejar que la tecnología, y en específico el diseño de DataMart y herramientas de inteligencia de negocios, ayudan a la conservación del medio ambiente por medio de prácticas de ahorro de recursos, gestión de la información, y sobre todo con la inteligencia de las organizaciones que ayudan a consolidar industrias enmarcadas en la práctica de procesos tecnificados y de precisión con tecnología de punta que no afecta a la naturaleza; proyectos como los coches eléctricos, el reciclaje de material para la elaboración de fundas o productos reutilizables; son estrategias favorables para los ecosistemas y que se miden con reportes de Big Data, repositorios de información, Data Warehouse, entre otros; el

control de la información es muy valioso para la inteligencia y la memoria de las corporaciones y multinacionales. En este sentido el impacto de la solución DataMart es ALTO positivo.

4.11.1.- Creación de empleo

La inteligencia de negocios y los proyectos tecnológicos, entre ellos los DataMart y Data Warehouse han evolucionada la capacidad de crear empleo, hasta la línea de inventar nuevas ramas profesionales como:

- 1.- Profesores en línea: Este empleo permite a los educadores impartir sus clases de manera online, ayudando a muchas personas que no pueden acceder a la educación por diferentes circunstancias como discapacidad, enfermedad, etc.
- 2.- Entrenadores físicos personalizados: Esta manera de empleo ayuda a las personas a mantener una condición física saludable.
- 3.- Personas creativas: El empleo que permite a las personas desarrollar contenidos para los dispositivos, como juegos o videos; los famosos Youtubers que ganan millones por enseñar a la gente sus videos peculiares y de monotonía.
- 4.- Consultores de imagen: Aquel empleo que permite a las personas asesorar a las personas u empresas sobre un producto o servicio para mejor comercialización.
- 5.- Genetistas digitales: El empleo dedicado a la atención de la salud, la fisioterapia, la estética y la odontología por medio de canales online y redes sociales.
- 6.- Ingenieros hospitalarios: El empleo que ayuda a los hospitales con el mantenimiento de sus equipos sofisticados por medio de la tecnificación online.
- 7.- Tiendas virtuales: El comercio electrónico es sin duda el mayor utilitario de los proyectos tecnológicos y de desarrollo, en la actualidad se puede comprar en todo el mundo de manera digital, a través de las tiendas digitales que muestran sus catálogos de productos, así como sus maneras diversificadas de pago como dinero electrónico, transferencias bancarias o pagos digitales como PayPal, Bitcoin, etc. La facilidad es que el producto llega al lugar de preferencia del comprador, el pago no es excesivo y todo desde la comodidad de una tablet, smartphone o computador personal.

Los profesionales y emprendedores deben estar preparados para las nuevas oportunidades laborales que proporcionan los proyectos tecnológicos, las empresas requieren profesionales

capaces de trabajar en el mundo digital e innovar para generar más fluctuación de la información, la economía y el comercio. Un ejemplo claro es la formación en Big Data, que congrega la información de múltiples Data Warehouse y miles de DataMart, toda esta información para tener reportes más explícitos y prospectivos del comportamiento del mercado online, las oportunidades y amenazas para la toma de decisiones.

4.11.2.- Medidas orientados a la igualdad de género

Entre las medidas orientadas a la igualdad de género que propicia el DataMart en consenso con la inteligencia de negocios y los proyectos tecnológicos, se destacan:

1.- Alfabetización digital: El acceso a un proyecto tecnológico, de inteligencia de negocios conciliado con los recursos de un DataMart permite el registro de hombres y mujeres a un conocimiento global en todos los campos de la actividad humana. Las mujeres y los hombres pueden capacitarse y disuadir su desconocimiento en temas que sean de su interés, almacenados en Data Warehouse de información de cultura general, un ejemplo claro son las enciclopedias digitales que potencia el conocimiento de los hombres y mujeres para un mejor desenvolvimiento en la sociedad y el trabajo.

2.- Creación y fortalecimiento empresarial: Las mujeres y los hombres en igual de género, en la actualidad han creado emprendimientos tecnológicos gracias a la inteligencia de negocios, las base de datos, DataMart y Data Warehouse, porque las redes sociales se invaden con publicidad de empresas de cosmetología, de venta de artículos importados, joyería, juguetería, etc. El comercio electrónico permite la igual de género para el emprendimiento y el desarrollo económico de hombres y mujeres.

3.- Desarrollo profesional y empleo: En la actualidad los proyectos tecnológicos involucrados con la inteligencia de negocios y la gestión de información con herramientas de DataMart han permitido que las mujeres en muchos países, accedan a la educación y logren formarse profesionalmente en localidades culturalizadas por el machismo y los estereotipos. La competitividad de hombres y mujeres está ligada a su voluntad por navegar en almacenes de información o el acceso a las carreras a distancia que ofrecen Universidades para buscar globalizar el conocimiento, esto favorece a mujeres que ya no solo dependen del esposo para su manutención, sino que generan empleo y con muchos mejores regalías que la de los hombres, todo dependiendo de la creatividad, voluntad y eficiente administración de sus objetivos e intereses.

4.- e-comunicación de género: El empoderamiento de las mujeres es una palabra que se ha convertido en un llamamiento del género femenino para buscar el respeto de los hombres y la igualdad de género; se visualiza en las redes sociales y los medios de comunicación, convocatorias masivas para buscar la igualdad de género, y es que con la tecnología y con los almacenes de datos la información de mujeres maltratadas o de hombres maltratadores es conocida y comunicada para establecer medidas preventivas y correctivas. Las mujeres convocan fácilmente a marchas mundiales para exigir su respeto; por su lado los hombres se unen a marchas y convocan el apoyo a la vida, a la familia y el amor a las mujeres, como madres, amigas, hermanas e hijas.

5.- e-salud: Las oportunidades para el acceso a la atención a tratamientos de salud en medios electrónicos involucrados en Data Warehouse y DataMart, ha evolucionado con los proyectos tecnológicos, tanto que permiten a las mujeres al igual que a los hombres disponer de fuentes de consulta para tratamientos sobre patologías de cualquier enfermedad, así como información de profesionales altamente especializados para una atención más personalizada; sin inconvenientes, hombres y mujeres tiene igualdad para la navegación en repositorios de doctores online.

6.- Ocio y cultura: La inteligencia de negocios y los DataMart junto con los proyectos tecnológicos han conseguido diseñar juegos, aplicaciones, software y herramientas digitales para smartphone, tablets, computadoras de escritorio y personales; que permiten el entretenimiento y el aprendizaje de hombres y mujeres en iguales condiciones, aplicaciones de noticias, cocina, cultura general, belleza, fitness, etc.

7.- Mainstreaming de género: Los proyectos tecnológicos de inteligencia de negocios y DataMart deben permitir la reorganización, el desarrollo, la mejora y la evaluación de los procesos políticos para que la igualdad de género se adhiera en las políticas y en las jerarquías gubernamentales, esto se ha conseguido por medio de la globalización de la información y el conocimiento, hombres y mujeres pueden conocer en repositorios, sus derechos y obligaciones para con la sociedad.

4.11.3.- Medidas orientadas a la inclusión social

Entre las medidas orientadas a la inclusión social que patrocina el desarrollo de un proyecto tecnológico como una solución DataMart para la inteligencia de negocios o el mismo uso requerido de los Data Warehouse, son las siguientes:

1.- Comunidades virtuales: En el campo social el acceso a la información como en la redes sociales por medio de DataMart u Data Warehouse favorecen la integración de las personas a las opiniones y decisiones públicas; favorecen la comunicación entre familiares que se encuentran dispersos, amigos, conocidos, profesionales de la misma carrera e intercambiar ideas y conocimiento para la mejora personal y profesional, lo que repercute en el desempeño de las organizaciones y la inteligencia de negocios.

2.- Internet como medio de participación: En el campo político la participación de la tecnología es muy relevante, DataMart de ciudadanos y partidos políticos ha podido establecer resultados ágiles en votaciones u orientaciones de la población por la planificación y el trabajo de los gobernantes, la participación de la población está directamente proporcional al uso de comunidades virtuales y redes sociales con información que está vinculada a la funcionalidad de la inteligencia de negocios, porque a una empresa también la afectan las políticas.

3.- Desarrollo de habilidades y capacidades para el uso de la tecnología como requisito para el acceso al empleo, servicios, etc.: En el campo educativo la inteligencia de negocios y los DataMart son necesarios para crear oportunidades en los profesionales que está a la vanguardia del uso de la tecnología. Actualmente existen modalidades de educación como el e-learning, b-learning o m-learning que ayuda el acceso a la educación en cualquier parte del mundo y en Universidades de prestigio. La inteligencia de negocios se alimenta del talento e investigación de muchos profesionales que dedican sus estudios al desarrollo administrativo y la gestión de proyectos tecnológicos por el beneficio de la sociedad. Una persona preparada puede tener mayores oportunidades laborales o crear emprendimientos.

4.- Mejora el mercado laboral, accesos, desarrollo del mercado electrónico: En el campo económico la inteligencia de negocios y el diseño de DataMart tiene sus resultados más efectivos porque permiten cualificar la oferta y la demanda de los productos, valorar el criterio y el gusto de los compradores, calificar catálogo de productos y estimar los precios de conveniencia. El comercio electrónico son los que mayormente explotan el paradigma de la tecnología de la inteligencia de negocios porque permite comprar cualquier producto en cualquier parte del mundo con acceso a base de datos de catálogo de productos con tablas dimensionales; este tipo de comercio mejor la economía de todos los países, el emprendimiento y las oportunidades laborales.

4.11.4.- Medidas orientadas a la mejora de la sostenibilidad

Entre las medidas orientadas a la mejora de la sostenibilidad se encuentran las siguientes:

- 1.- Satisfacer al cliente/usuario o patrocinador del DataMart: Es importante satisfacer al usuario del DataMart entregándole valor en cortos espacios de tiempo, para no incurrir en descontentos del usuario con los resultados finales; también para recibir calificaciones favorables de la opinión del usuario y provocar la pérdida de tiempo y recursos.
- 2.- Aceptar el cambio de los requisitos del DataMart, en indeterminado momento: La metodología ágil como SCRUM, se adaptan al cambio de manera efectiva, esta es la ventaja de no funcionar con una planificación estricta e inamovible
- 3.- Entregar valor frecuentemente: Esta medida se enfoca en el principio OODA (orientar, observar, decidir y actuar) porque mientras más proporcional sea el avance de una solución mejor será el resultado final deseado y con ello la sostenibilidad de la DataMart. La entrega de valor es iterativa porque se trabaja con iteraciones, se detiene, observa, ajusta y vuelve a iterar. O también incremental por que se sustenta valor poco a poco y no directamente el resultado final que no puede ser de complacencia del patrocinador o del discente.
- 4.- Trabajar directamente con el cliente/usuario: La colaboración con el cliente de manera constante, indica el resultado obtenido del DataMart y su ajuste con los requerimientos del usuario en cada iteración.
- 5.- Trabajar con el equipo de la solución motivado: El equipo de trabajo motivado ayuda a mejorar la sostenibilidad de la solución y el negocio porque el desempeño de los colaboradores genera inteligencia organizacional como un activo o capital humano para la mejora continua.
- 6.- Buscar directamente la comunicación: La comunicación directa es más efectiva que los medios electrónicos u otras herramientas de comunicación; en ella se enfrenta un dialogo más sincero y real de los necesidades, requerimientos o satisfacción del usuario.
- 7.- Medir el proceso del DataMart: Es importante valorar las métricas para estipular la marcha de la solución, esto puede ser posible con trocear el trabajo a realizar en tareas que aporten a la inteligencia del negocio.
- 8.- Mantener un ritmo laboral sostenible: Una medida para la mejora de la sostenibilidad es mantener el trabajo constante sin declives que perjudique la calidad del proyecto, en este sentido no se puede exigir al equipo horas extensas de trabajo si no pueden mantener un ritmo constante para la consecución de la solución.

9.- Intentar el éxito en la operatividad del DataMart: El talento del equipo de trabajo depende de la potencialidad con la que se gestione el desarrollo de la solución, no se puede esperar alto rendimiento sino se capacita, se gratifica, se brinda oportunidades o facilita el trabajo del equipo.

10.- Depurar la información que no aporta valor: La metodología ágil junto con la sostenibilidad aportan como medida de mejora que no hay que perder el tiempo en tareas que no aportan valor a la solución.

11.- Permitir que el equipo sea autodidacta para lograr mejores resultados: La creatividad y procesos autodidactas del equipo de trabajo es necesario explotarlos y dejarlos fluir sin cuartarlos para mejorar al máximo el potencial; el equipo participa en las decisiones, por lo que su grado de compromiso será el máximo para con el proyecto y la inteligencia del negocio.

12.- Reflexionar sobre éxitos y fracasos de la inteligencia de negocios: La toma de decisiones establecidas en instrumentos tecnológicos y herramientas como el DataMart permiten la toma de decisiones pero hay que tener presente que la decisión siempre será la que marque la diferencia; con la memoria corporativa la inteligencia del negocio puede ser más asertiva y aprender de las decisiones fracasadas.

CAPÍTULO 5. CONCLUSIONES Y TRABAJO FUTURO

5.1.- Principales conclusiones

- El TFM analizó la importancia del Data Warehouse (DW) para organizaciones y empresas como estrategia de la inteligencia de negocios. En este sentido, la literatura permitió destacar dos enfoques de las organizaciones para el diseño de DW y por ende de DataMart que son enfoques de arriba hacia abajo (top-down) y de abajo hacia arriba (bottom-up), y se mostró las ventajas y desventajas de cada enfoque. Para el Centro de Educación Continua (CEC) de la Universidad Técnica de Machala, se adoptó el enfoque de abajo hacia arriba conocido como enfoque de Ralph Kimball.
- El TFM permitió discutir el efecto del DW en el rendimiento de la base de datos y de la plataforma del CEC de la Universidad Técnica de Machala y el tiempo de procesamiento de consultas, que en la actualidad tiene problemas de lentitud de procesamiento de la información y de reportes, así como la confiabilidad de los datos. Al comparar los resultados de las consultas que propone una base de datos construida con DataMart, está claro en cada caso que la recuperación de datos es mucho más rápida desde el esquema en estrella organizado en el DW que desde la base de datos transaccional.
- El manejo de la información en el CEC necesita ser mejorado para poder generar mejores servicios a la memoria organizacional de la Universidad Técnica de Machala, un centro que forma a los estudiantes de manera continua y permanente tiene que establecer consultas inmediatas, así como la toma de decisiones eficientes en lo que refiere a la preparación de los estudiantes. Con el uso de los DataMart, la información puede ser organizada y resumida en reportes que aventajan la toma de decisiones con respecto a la parte académica y administrativa del CEC.
- El TFM ha podido demostrar que el DataMart en conjunto al DW y la inteligencia empresarial pueden ser implementados por cualquier organización, sin importar la razón social de su funcionalidad, como es el caso de las universidades; además, que el DataMart debe tener un objetivo claro de apoyar tanto a los usuarios comerciales como a los ejecutivos. Es posible diseñar un DW para toda la entidad a partir de la DataMart eficiente. El DataMart prototipo del TFM permite la escalabilidad en el CEC, entidad de educación superior.

- Las instituciones de educación superior han almacenado datos operativos durante mucho tiempo, y han seguido acumulando enormes cantidades de datos a una velocidad masiva a medida que las bases de datos operativas se vuelven más valiosas, ayuda a aumentar la velocidad a la que las entidades estudiantiles tienen éxito. Con el TFM se ha demostrado que el DataMart junto con el DW recopila, consolida, organiza y resume estos datos estructurados y no estructurados para que estos datos inteligentes puedan usarse para informar las decisiones estudiantiles.
- Otra importante conclusión que deriva del TFM es que se observó que la clave para el DW es el diseño de DataMart comprensible y que satisfagan los requerimientos de los usuarios. Un buen diseño de datos crea un DW viable. Los usuarios de los sistemas educativos saben qué datos necesitan y cómo quieren usarlos. El DataMart del TFM se centra en los usuarios para determinar qué datos se necesitan, ubicar las fuentes de los datos y organizarlos en un modelo dimensional que represente las necesidades del área Administrativa del CEC. El análisis de datos y los informes no son suficientes para aprovechar los beneficios de la inteligencia empresarial en una entidad tan dinámica como es la educación superior. La información requerida, necesaria y correcta también debe encontrar una manera de llegar a las personas adecuadas en tiempo suficiente. Pero, lo que es más importante, se debe actuar sobre la información para activar los beneficios, en este sentido es conveniente que la inteligencia de negocios sea parte del proceso administrativo del CEC, dentro del sector de gestión Admisión y Registro; Financiación Estudiantil y Finanzas; por otra parte, la BI no solo debe ayudar a comprender el pasado, sino también trabajar para encontrar nuevas oportunidades y tendencias emergentes en el futuro.
- El sentido epistemológico del TFM ha permitido comprender que un DW es una gran base de datos utilizada para revelar y examinar los datos en la base de datos. Los datos recopilados en el DW se actualizan desde los sistemas operativos, como en un mercado. Los datos superan los datos operativos almacenados para otras operaciones por esta razón es que la organización de la información con un programa de BigData y específicamente en el diseño de DataMart ayudan de manera directa en la toma de decisiones porque reportan solo la información que el usuario necesita para ver la situación de la entidad. Los datos en el DataMart en conjunto con la DW generado a partir de sistemas de fuente de datos integrados no necesita procesos de extracción, transformación y carga (ETL) para producir bases de datos o bases de datos de DW operativas, por esta

razón se deduce que los DW se subdividen en DataMarts, y a su vez; los DataMarts almacenan los datos en subconjuntos de DW.

- Los DataMart como estrategias de inteligencia de negocios, está comprobado que pueden mejorar la toma de decisiones porque junto con la base de datos muestra consultas directas de datos que hacen que las partes interesadas puedan observar los mejores caminos para llegar a diversas soluciones de un problema, en el caso de CEC, con el diseño del DataMart se puede prever que muchas dificultades que tenían los administradores al momento del registro y admisión de estudiantes, los resultados financieros o el aprovechamiento académico, servirá para la toma de decisiones marcando un impacto positivo en los estudiantes que quieren una educación continua sin tapujos en la administración de la información que procesa la Universidad Técnica de Machala.

5.2.- Líneas de trabajo futuro

- El DataMart derivado como prototipo de contribución, en la fase actual, ha diseñado un DataMart que mejora las consultas y funcionalidad de las bases de datos de la plataforma del CEC, en el futuro se espera implementar un DW que ayude principalmente en el proceso de toma de decisiones relacionado con los estudiantes, su aprovechamiento y sus aspectos financieros y que se pueda integrar al sistema de todo el campus universitario. Se espera mejorar el prototipo DataMart para el área Administrativa con técnicas heurísticas de redes neuronales para mostrar información más clara y precisa para la toma de decisiones, además expandiendo las funciones fundamentales para los estudiantes del CEC, como Admisión y Registro, Asuntos Estudiantiles y Finanzas.
- Es factible que el internet de las cosas IoT permita trabajar en conjunto a las tecnologías de la información y la comunicación para buscar una educación de calidad, un ejemplo claro la educación a distancia como e-learning, b-learning o m-learning o las plataformas digitales que permiten a los estudiantes investigar para mejorar su formación, es posible entrelazar los DW del CEC en oferta a las nuevas tecnologías de la información. En este sentido, la actualización y mantenimiento de la base de datos es muy costosa por el tiempo que se invierte, implementando un DW para el CEC que se vincule a las plataformas de educación virtual, ayudará a generar mayor oferta y demanda con respecto a la educación continua y permanente.
- En un DW el mayor problema para la creación del front-end son las vistas y visualizaciones de los reportes o la información que las DataMart procesan, por esta razón se espera como

trabajo futuro indagar en los algoritmos de selección de vista materializada. A pesar de que la mayoría de los trabajos en este campo han recibido diferentes enfoques para la selección de vistas a materializar es poco considerado el costo de mantenimiento de la vista y el espacio de almacenamiento. Se necesita motivar investigaciones para seleccionar las vistas que se materializarán en función de los umbrales dinámicos, como la frecuencia de uso de la vista. Por lo tanto, un DW debe diseñar un modelo para la selección de vista materializada basada en la frecuencia de uso de la vista en diferentes períodos sobre el aspecto académico, financiero y de la modalidad de educación.

- En base a la evaluación del prototipo de DataMart diseñado en el TFM, así como las potencialidades de los diferentes lenguajes de programación, para originar un BigData del Centro de Educación Continua que pueda ser una muestra que el modelo propuesto puede integrarse con los algoritmos existentes de selección de vistas materializadas para dar mejores decisiones de materialización, e incluso un mejor rendimiento de la base de datos, ya que restringe el conjunto de vistas antes del proceso de materialización, ahorrando tiempo y recursos. En el futuro, se evaluará la integración de la ventana de inteligencia empresarial en las herramientas de colaboración, con aplicativos que trabajen con las plataformas móviles.

Referencias bibliográficas

- Bentley, W. (2008). Análisis de sistemas y diseño y métodos. México, Editorial McGraw-Hill, Primera Edición, pp. 62-69.
- Bernabeu, R. (2010). Data warehousing: Investigación y Sistematización de Conceptos-Hefesto: Metodología propia para la construcción de un Datawarehouse. Córdoba, Argentina, pp. 72-73.
- Davies, P. (2014). Sistemas de información. España, Editorial Reverté, Primera Edición, pp.33-51.
- Bibliomática.com. Minería de datos y diseño de Data Warehouse. Recuperado de: <http://bibliomatica.blogspot.com/2014/10/tema-7-almacenes-de-datos-y-mineria-de.html>
- Bustos, S., & Mosquera, V. (2013). Análisis, Diseño e Implementación de una solución Business Intelligence para la generación de indicadores y control de desempeño en la Empresa OTECEL S.A. utilizando la metodología Hefesto V2.0. Ecuador, Quito, pp.17-29.
- Cano, J. (2017). Inteligencia de negocios: Competir desde la información. Madrid, España, Editorial ESADE, Primera Edición, pp.32-41.
- CEAACES. (2015). Modelo de Evaluación Institucional de Universidades y Escuelas Politécnicas del Ecuador. Quito, Ecuador.
- Coronel, C., & Morris, S. (2011). Bases de datos diseño, implementación y administración. México, México D.F. Editorial Cengage Learning, pp.27-32.
- DATAPRIX.com (2016). ¿Qué son los Data Warehouse y los Data Mart? Recuperado de: <http://www.dataprix.com/que-es-un-datawarehouse>
- Daza, A. (2016). Data Mining y desarrollo de almacenamiento de datos estructurados. Lima, Perú, Editorial Macro EIRL, pp.44-46.

- EvaluandoSoftware. (2016). Evaluando bases de datos. Recuperado de: <http://www.evaluandosoftware.com/sistemas-oltp-procesamiento-administracion-mantenimiento-transacciones/>
- Ferrer, S. (2015). Los sistemas de procesamiento y manejo de la información. Recuperado de: <http://pertutatis.cat/la-piramide-de-losdiferentes-tipos-de-sistemas-de-informacion/>
- Ferrer, S. (2015). ¿Cuáles son los sistemas de toma de decisiones según la inteligencia de negocio? Recuperado de: <http://pertutatis.cat/que-son-los-sistemas-de-a-poyo-a-la-toma-de-decisiones-dds/>
- Gómez, V., & Suárez R. (2010). Sistemas de información, herramientas prácticas para la gestión. México, México D.F., Editorial Alfaomega Grupo Editor S.A., pp.86-87.
- INTERAKTIV. (2012). Esquemas de Data Mart para la toma de decisiones. Recuperado de: http://www.interaktiv.cl/blog/wp-content/uploads/2012/04/8-Eschema_FisicoMolap-Rolap-Holap.pdf
- Joyanes, A., (2013). Big Data análisis de grandes volúmenes de datos en organizaciones, México, México D.F., Editorial Alfaomega, Segunda Edición, pp. 37-45.
- Kendall, K., & Laudon, K. (2010). Sistemas de información, fundamentos y análisis. México, México D.F. Editorial Person Educación, Primera Edición, pp. 49-57.
- LatinoBI. (2013). Introducción a la inteligencia de negocios. Recuperado de: <http://www.latino-bi.com/espanol/fundamentos-bi/introduccion-al-bi.php>
- Mannino, M. (2007). Administración de bases de datos diseño y desarrollo de aplicaciones. México, México D.F. Editorial McGraw Hill, Primera Edición, pp.48-69.
- OPENRED. (2014). Open Red Soluciones en Software Libre. Recuperado de: <http://www.openred.es/index.php/pentaho>

Pérez, C., & Satin, D. (2006). Data Mining y la lógica del diseño de las base de datos concentradas en Data Warehouse. México, México D.F., Editorial Alfaomega, pp.58-59.

Piattini, M., & Marcos, E. (2007). Data Mart y Data Warehouse una visión a la tecnología y diseño de bases de datos. México, México D.F. Editorial Alfaomega, pp.52-63.

Romero, A. (2012). Gestión del conocimiento organizacional e inteligencia de negocios. Recuperado de: <https://es.slideshare.net/alfredoromero/gestin-de-conocimiento-organizacional-1-parte-resumida>

The Data Warehouse Institute. (2013). Fundamentos de la Inteligencia de Negocios. Recuperado de: <http://www.latino-bi.com/espanol/fundamentos-bi/introduccion-al-bi.php>

Toro, J. (2012). Pautas para el estudio de los Data Warehouse: Gestión de la calidad en las instituciones de Educación Superior Chile. Santiago de Chile, Chile, Editorial CINDA-RIL, pp.80-84.

Turban, E. (2012). Information Technology for Management. Estados Unidos, New York, Editorial Wiley, Primera Edición, pp. 47-69.

Bibliografía

Anahory, S., & Murray, D. (1991). Data Warehouse en el mundo real. Editorial Adisson Wesley, Primera Edición, p.p.35-37.

Inmon, W. (2002). Construyendo una Data Warehouse, Editorial Wiley Computer Publishing, Segunda Edición, pp. 59-87.

Jarkemaurizio, L., Vassiliou, Y. (2000). Fundamentos de las Data Warehouse, Editorial Springer, Primera Edición, pp.101-113.

Kalnis, P., Mamoulis, N., & Papadias, D. (2002). View selection using randomized search. Data & Knowledge Engineering, pp.89-111.

- Kimball, R., & Ross, M. (2010). The Kimball Group Reader; Relentlessly Practical Tools for Data Warehousing and Business Intelligence. Indianapolis, Editorial Wiley, pp.22-29.
- Kimball, R., & Ross, M. (2013). The Data Warehouse Toolkit: The definitive guide to dimensional modeling. Indianapolis, Editorial Wiley, pp.37-38.
- Lapa, A. (2016). Desarrollo de un datamart para información táctica del impuesto predial del servicio de administración tributaria de huamanga. Tesis de pregrado, Universidad de San Cristóbal de Huamanga, Ayacucho, Lima.
- Lee, M. (2001). Speeding Up Materialized View Selection in Data Warehouses Using a Randomized Algorithm. International Journal of Cooperative Information Systems, 10(3), pp. 327-353.
- Nima, J. (2009). Soluciones OLAP con Microsoft SQL Server AnalysisServices. Recuperado de: <http://www.eumed.net/libros-gratis/2009c/574/index.htm>.
- Tamayo, M. y Moreno, F. (2006). Análisis del modelo de almacenamiento MOLAP frente al modelo de almacenamiento ROLAP. Ingeniería e Investigación. Recuperado de: <http://www.redalyc.org/html/643/64326317/>
- Torres, L. (2007). Business Intelligence. Recuperado de: <http://www.gravitar.biz/index.php/bi/bi-terminologia-1>.
- Universidad Internacional de La Rioja (UNIR) (2019). Diseño del TFM. Escuela Superior de Ingeniería y Tecnología. Máster Universitario en Diseño y Gestión de Proyectos Tecnológicos.

Anexos

Anexo I: PRUEBAS DE ACEPTACIÓN

Las Pruebas de Aceptación son las características con las que debe cumplir un requerimiento (historia de usuario) para ser aceptado, está relacionado principalmente con la funcionalidad.

Las pruebas de aceptación son evaluadas por el facilitador que en este caso es el administrador de la aplicación. El DataMart tiene que cumplir con los requerimientos funcionales y ser desarrollado en el tiempo estipulado en el cronograma según las fases de análisis, diseño, transferencia y pruebas. Como resultado de la evaluación de las pruebas se determinó la siguiente tabla I.1.

Tabla I.1. Pruebas de Aceptación

Detalle	Fecha Inicio Revisión	Fecha Fin Revisión	Pruebas Exitosas	Pruebas Fallidas
02 Historias de Usuario	01-12-2019	15-01-2020	10	0

Fuente: Elaboración propia

La documentación de las actividades realizadas para evaluar las pruebas de aceptación se detallan en la tabla I.2.

Tabla I.2. Prueba de Aceptación 01

PRUEBA DE ACEPTACIÓN 01	
ID: PA-01	Nombre: El nombre de las tablas y columnas deben cumplir con un estándar.
Descripción: En el estándar establecido con el cliente se estableció que los nombres de las tablas y las columnas deben empezar con las iniciales correspondientes.	
Responsable: Arreaga Santacruz, Sally Astrid	Fecha: 27/11/2019
Pre-Condiciones: Debe estar creada la Base de Datos, con las tablas y columnas.	

(Continuación)

(Continuación)

Pasos de ejecución:

- Ingresar a ORACLE 11g.
- Conectarse con la base de datos.
- Seleccionar el esquema público.
- Revisar que los nombres de las tablas y las columnas cumplan con el estándar establecido.

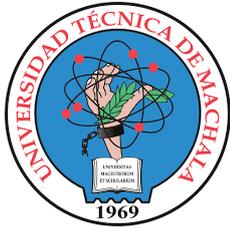
Pos-Condiciones: Los nombres de las tablas deben terminar con la letra ADC (ACADÉMICO y las columnas con las 3 primeras letras de la tabla correspondiente o con toda la palabra si el campo lo amerita.

Resultado: satisfactorio.

Anexo II: CIERRE DEL PROYECTO Y PLANTILLAS

5.- CIERRE DEL PROYECTO

5.1.- ACTA DE CIERRE DEL PROYECTO

Código de documento:	CEC-UTM-DM-24-2020	
Fecha: febrero 2020	Versión: V 1.0	
Proyecto:	Diseño del DataMart	
Unidad:	Área Administrativa del CEC	
Autora:	Sally Astrid Arreaga Santacruz	

Generalidades de cierre del proyecto	
Fecha de inicio	
Fecha entregable final	
Fecha de cierre	

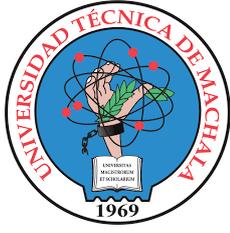
Aceptación del entregable			
Entregable del proyecto	Fecha aceptación	Firma del Patrocinador	Firma del Director del proyecto
Acta de constitución			
EDT			
Cronograma			
Presupuesto			

Aceptación cierre del proyecto

A través de la presente acta, se finaliza el proyecto cuya denominación es: Diseño de un DataMart para la toma de decisiones del Área Administrativa del CEC de la Universidad Técnica de Machala. En forma directa el Patrocinador y el Director del proyecto presentan su rúbrica de aceptación de toda la documentación inherente al entregable final.

----- Director del proyecto	----- Patrocinador
--------------------------------	-----------------------

Conocimiento adquirido:

Código de documento:	CEC-UTM-DM-25-2020	
Fecha: febrero 2020	Versión: V 1.0	
Proyecto:	Diseño del DataMart	
Unidad:	Área Administrativa del CEC	
Autora:	Sally Astrid Arreaga Santacruz	

Memoria Organizacional	
ID línea de investigación	
Área del conocimiento	
Fecha	
Oportunidad	

Título conocimiento aprendido
Descripción del suceso
Descripción de la influencia en los objetivos del proyecto
Acciones implementadas preventivas o correctivas
Recomendaciones

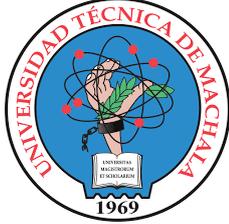
6.- PLANTILLAS DEL PROYECTO

Acta de constitución del proyecto

Código de documento:			
Fecha:	Versión:		
Proyecto:			
Unidad:			
Autora:			
Patrocinador:			
Fecha inicio		Fecha final tentativa:	

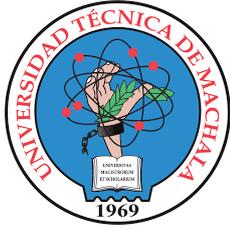
Objetivos del proyecto	
Objetivo General	
Objetivos Específicos	
Justificación del proyecto	
Descripción entregable final del proyecto	
Principales riesgos	
Requisitos de alto nivel	
Mapa de hitos	
Hito	Fecha de vencimiento
Realizado por:	Aprobado por:

Declaración del alcance:

Código de documento:			
Fecha:	Versión:		
Proyecto:			
Unidad:			
Autora:			
Patrocinador:			
Fecha inicio		Fecha final tentativa:	

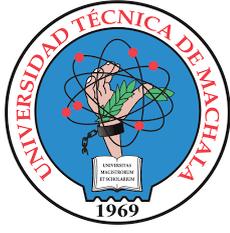
Declaración del alcance	
Criterios de aceptación de la solución	
1.- Técnicos: 2.- De calidad: 3.- Administrativos: 4.- Comerciales:	
Entregables del proyecto	
Proceso del proyecto	Entregable
Exclusiones del proyecto	
Restricciones del proyecto	
Internos	Externos

Plan de gestión de la integración:

Código de documento:			
Fecha:	Versión:		
Proyecto:			
Unidad:			
Autora:			
Patrocinador:			
Fecha inicio		Fecha final tentativa:	

Roles de la gestión de la integración						
Rol	Persona asignada	Responsabilidad	Nivel autoridad			
Plan de documentación						
Documento	Formato	Acceso rápido	Disponibilidad	Seguridad	Recuperación de información	Retención de información
Ítems de integración						
Código	Nombre	Tipo de ítem	Fuente	Formato	Observación	
Gestión del cambio						
Roles gestión de cambio						
Rol	Persona asignada	Responsabilidades	Nivel de autoridad			
Tipo de cambio						
Proceso de gestión de cambios						
Plan de contingencia frente a solicitudes de cambio - urgentes						
Verificación y auditoría de configuración						

Plan de gestión del alcance:

Código de documento:			
Fecha:	Versión:		
Proyecto:			
Unidad:			
Autora:			
Patrocinador:			
Fecha inicio		Fecha final tentativa:	

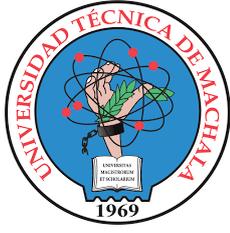
Proceso de definición del alcance
Proceso de elaboración de la EDT
Proceso para la elaboración del diccionario EDT
Proceso para verificación del alcance
Proceso para control del alcance

Acta de aceptación de entregables:

Código de documento:		
Fecha:	Versión:	
Proyecto:		
Unidad:		
Autora:		
Patrocinador:		
Fecha inicio		Fecha final tentativa:

Generalidades del entregable		
Entregable		
Versión		
Fecha cronograma		
Fecha verificación		
Fecha entregable		
Descripción del entregable		
Observaciones		
	Nombre	Firma
Entregado por:		
Aprobado por:		

Plan de gestión de requisitos:

Código de documento:		
Fecha:	Versión:	
Proyecto:		
Unidad:		
Autora:		
Patrocinador:		

Actividades de los requisitos
Actividades de gestión de configuración
Priorización de requisitos
Métricas de la solución
Estructura de trazabilidad

Plan de gestión de tiempo:

Código de documento:		
Fecha:	Versión:	
Proyecto:		
Unidad:		
Autora:		
Patrocinador:		

Proceso de definición de actividades
Proceso de secuencia de actividades
Proceso de estimación de duración
Proceso desarrollo del cronograma
Proceso control de calendario

Plan de gestión de costos:

Código de documento:		
Fecha:	Versión:	
Proyecto:		
Unidad:		
Autora:		
Patrocinador:		

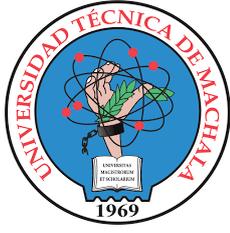
Nivel de precisión	Unidades de medida	Umbral de control
Reglas de medición desempeño		
Gestión del proyecto		
Observaciones		

Plan de gestión de la calidad:

Código de documento:		
Fecha:	Versión:	
Proyecto:		
Unidad:		
Autora:		
Patrocinador:		

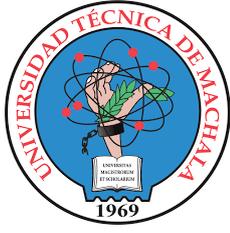
Políticas de calidad				
Línea base de calidad				
Factor de calidad	Objetivo de calidad	Métrica usable	Frecuencia y tiempo	Frecuencia y reporte
Plan de mejora de procesos				
Roles para la gestión de calidad				
Organigrama para la calidad				
Procesos de gestión de la calidad				

Plan de gestión de riesgos:

Código de documento:		
Fecha:	Versión:	
Proyecto:		
Unidad:		
Autora:		
Patrocinador:		

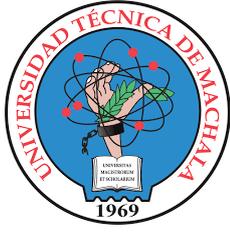
Metodología de gestión de riesgos			
Proceso	Descripción	Herramienta	Fuente información
Periodicidad de la gestión de riesgos			
Proceso	Momento de ejecución	Entregable de EDT	Periodicidad de ejecución
Formatos de gestión de riesgos			

Acta de cierre del proyecto:

Código de documento:		
Fecha:	Versión:	
Proyecto:		
Unidad:		
Autora:		
Patrocinador:		

Generalidades del cierre			
Fecha inicio			
Fecha aceptación			
Fecha de cierre			
Aceptación formal de entregables			
Entregable	Fecha aceptación	Firma Patrocinador	Firma Director
Acta de constitución			
EDT			
Presupuesto			
Cronograma			
Aceptación de cierre del proyecto			
Rúbricas			
----- Director del proyecto		----- Patrocinador	

Conocimiento aprendido:

Código de documento:		
Fecha:	Versión:	
Proyecto:		
Unidad:		
Autora:		
Patrocinador:		

Referencia	
ID Línea de investigación	
Área de conocimiento	
Fecha	
Oportunidad	
Título conocimiento aprendido	
Descripción del conocimiento aprendido	
Descripción del impacto en los objetivos	
Acciones implantadas preventivas y correctivas	
Conocimiento organizacional	
Recomendaciones	