

Elaboración cubo de información Sistema Gestión de matrícula

por Jordan-angie Erazo-martinez

Fecha de entrega: 06-nov-2019 03:05p.m. (UTC-0500)

Identificador de la entrega: 1208492505

Nombre del archivo: ERAZO-MARTINEZ.docx (131.68K)

Total de palabras: 6592

Total de caracteres: 37840

CAPÍTULO 1

1. INTRODUCCIÓN

Actualmente la educación ha pasado por momentos de cambios acelerados por la sociedad a nivel internacional en todos los aspectos, los planes operativos de una institución educativa suelen ser obsoletos frente a una demanda de grandes transacciones, por lo que necesariamente se debe de buscar nuevas alternativas operacionales que se acoplen a las necesidades de una institución.

El instituto tecnológico “Juan Bautista Aguirre” al igual que una gran “empresa” tiende hacer uso de grandes cantidades de datos para su respectivo análisis, todos estos datos al reunirse son considerados como información relevante para el instituto, esta información a su vez es de gran ayuda para descubrir posibles falencias dentro del plan académico y el plan operacional.

Toda organización necesita tomar buenas decisiones y para aquello se necesitan de datos y cifras concretas, que ayuden a mejorar el sistema educativo del plantel. Para esto es imprescindible tener toda la información adecuada que debe ser analizada con mucho cuidado, el imaginarse que un solo dato pudo ser alterado y este podrá generar pérdidas abrumadoras para la institución es un indicador de la minuciosidad con el que un sistema de business intelligence debe de contar.

La toma de decisiones es la elección entre dos o más acciones a seguir para obtener respuestas óptimas a los problemas o situaciones que se hagan presente. Estas a nivel gerencial son de suma importancia para la empresa por ende se deben saber implementar estrategias, técnicas y herramientas para obtener resultados eficaces.

Todas las empresas independientemente de su naturaleza, tamaño o sector van a plantear sus necesidades de información analítica las cuales deben resolver mediante la implementación de alguna herramienta, BI es la herramienta mayormente implementada en las organizaciones para la mejora en la toma de decisiones en las organizaciones.

Una herramienta de business intelligence ayuda a optimizar el tiempo de respuesta, brindando competitividad ante otras instituciones del mercado, en este caso ofreciendo una educación y servicio de calidad.

Este proyecto contribuye a la solución de las necesidades del instituto tecnológico “Juan Bautista Aguirre”, mediante el diseño y elaboración de un data mart (CUBO OLAP) para el análisis del proceso de matriculación, con la finalidad de brindar a la alta gerencia un método o sistema sencillo que permita tomar buenas decisiones referente a la parte académica de la institución.

El proceso de desarrollo del proyecto, consta con bases teóricas a la cual nos referiremos al implantar cada una de las etapas para su posterior finalización.

1.1. Planteamiento del problema⁴

La Inteligencia de Negocios es el conjunto de metodologías, prácticas y capacidades enfocadas al manejo de información que permite tomar mejores decisiones a las empresas. La práctica del BI se logra desarrollar a través de sistemas de tecnologías de la información y de un conocimiento profundo del core business de la empresa. (Silva, 2017). Permite generar conocimiento sobre la información que se manipula dependiendo de los datos que genere la organización, de esta manera al tener información que es cambiante a través del tiempo permite la manipulación de los mismos para llegar a tomar decisiones estratégicas que permitan a la empresa mayor desarrollo.

En la actualidad las empresas implementan BI (Inteligencia de Negocios) para mejorar la toma de decisiones que enfrentan a lo largo de su desarrollo, sin embargo, aquellas que la utilizan en su mayoría le dan un enfoque financiero, pero no solo puede ser en dicha área.

Una de las variadas herramientas que utiliza Business Intelligence es el procesamiento analítico en línea (OLAP), denominado cubo de información. Las herramientas OLAP proporcionan a las compañías un sistema confiable para procesar datos que luego serán utilizados para llevar a cabo análisis e informes que permiten mejorar las operaciones productivas, tomar decisiones inteligentes y optimizar la competitividad. (Morales & Cuevas, 2016). Su implementación es necesaria para el manejo masivo de grandes volúmenes de datos.

En el presente proyecto se enfocará la Inteligencia de Negocios en una institución educativa, a través del desarrollo de una solución OLAP para el análisis de la gestión de matrículas en el instituto Tecnológico “Juan Bautista Aguirre” del cantón Daule, provincia del Guayas.

El instituto tecnológico “Juan Bautista Aguirre” ubicado en el cantón Daule, Km 48 vía a Santa Lucia diagonal a Gasolinera Primax beneficia a las zonas aledañas entre cantones Lomas de Sargentillo, Isidro Ayora, Balzar, Pedro Carbo, Palestina y Daule sumando un alrededor de 254.000 habitantes, este Instituto ofrece carreras tecnológicas como: Tecnología Superior de Contabilidad, Tecnología Superior en Desarrollo de Software, Tecnología Superior en Administración. Las personas que radican en estos cantones se han beneficiado directamente gracias a las ofertas académicas de dicha institución, por este motivo los directivos de dicho instituto requieren una solución Business Intelligence que les permitirá

deducir información a nivel específico referente al área de gestión de Matrículas. de los estudiantes Para la creación de esta colección de datos se procederá a la recolección de los datos del instituto Tecnológico “Juan Bautista Aguirre”, mediante indicadores apegados al instituto se podrá presentar reportes estadísticos a la gerencia de dicha institución para la toma de decisiones

1.2. Objetivos

1.2.1. Objetivo General

El instituto “Juan Bautista Aguirre” cuenta con gran volumen de entrada de datos por lo que es necesario brindar soporte para el análisis complejo de los datos en diferentes perspectivas mediante indicadores que constan en la base de datos transaccional del instituto, Por este motivo el objetivo general es:

Diseñar y elaborar un CUBO OLAP basado en el análisis de la base de datos transaccional del instituto tecnológico “Juan Bautista Aguirre” que permita el análisis gerencial académico para la toma de decisiones.

1.2.2. Objetivos Específicos

Para la correcta toma de decisiones enfocada en el área de matriculación es necesario obtener información de la base de datos transaccional del instituto, de esta manera analizar los datos de diferentes perspectivas mediante la elaboración de data mart que permita medir indicadores de la base de datos de la institución, por este motivo los objetivos específicos son:

- Diseñar un data mart (CUBO OLAP) mediante cubos de información y el uso de la base de datos transaccional.
- Verificar la información actualizada del instituto tecnológico “Juan Bautista Aguirre” basándose en los datos almacenados en el repositorio de datos.
- Diseñar un modelo multidimensional del data mart basado en el modelo transaccional del instituto.

1.3. Justificación

Las herramientas de inteligencia de negocios o Business Intelligence son necesarias para cualquier tipo de organización, ya que son consideradas como un conjunto de tecnologías y aplicaciones que ayudan a la recopilación de datos, cuyos datos se traducen en información, dicha información es analizada por la alta gerencia para la correcta toma de decisiones.

El sistema propuesto será de mucha utilidad y de gran soporte a la gestión académica del instituto tecnológico “Juan Bautista Aguirre”, debido a que en la actualidad muchos planteles educativos no cuentan con sistemas que faciliten tener una visión en general de las matrículas de los estudiantes, estos registros pueden brindar a la institución conocimientos sólidos como por ejemplo la concurrencia de estudiantes matriculados por meses, los cursos preferidos por los alumnos tomando en cuenta la variable tiempo siendo un factor determinante para la institución.

Basándonos en la situación socio-económico de la institución “Juan Bautista Aguirre” se puede deducir que a través de una herramienta Business Intelligence se obtendrá una mejor administración en la gestión académica de los estudiantes del plantel, agilizando procesos y mejorando la toma de decisiones.

La herramienta Business Intelligence ayudará a establecer metas realistas para la institución, acelerará el tiempo de análisis y respuesta dentro de la gestión de matrículas en base a información concreta;

Dentro del proceso de elaboración de esta herramienta tecnológica en primer lugar se procederá al diseño multidimensional de un repositorio de información conocido como data mart basado en metodologías estandarizadas, entre las más utilizadas se encuentran la metodología estrella y la metodología copo de nieves, para después dar paso al diseño del Cubo OLAP que es un sistema de procesamiento analítico en línea, este último ayudará a la segmentación de los datos en función al tiempo y permitirá profundizar información que sea relevante.

1.4. Marco Teórico

1.4.1 Inteligencia de Negocios

Inteligencia de negocios o Business Intelligence es una herramienta que parte del análisis de información generada por la empresa para mostrar distintos escenarios o perspectivas y así generar reportes que permiten pronosticar resultados permitiendo tomar mejores decisiones.

H. P. Luhn (1958) define Inteligencia de Negocios como “la capacidad de aprehender las interrelaciones de los hechos presentados de tal manera que guíen la acción hacia una meta deseada”. (H. P. Luhn, 1958)

Luis Méndez del Río (2016) “Podemos entender el Business Intelligence como el conjunto de herramientas y aplicaciones que nos facilitan el acceso interactivo, análisis y manipulación corporativa de misión crítica. Estas aplicaciones aportan un conocimiento valioso que permite a la organización identificar problemas y nuevas oportunidades de negocio”.(del Río, 2006)

La inteligencia de negocios es la capacidad que tiene una organización para analizar su propio comportamiento y desarrollo histórico con el fin de conocer el antecedente de la empresa, su estado actual y lo que sucederá en el futuro, es el proceso de extracción e integración de información de hechos del negocio (empresa) para su respectivo análisis y de esta manera cumplir con los objetivos de la organización mediante la mejora de toma de decisiones.

1.4.1.1 Beneficio de la inteligencia de negocios

La inteligencia de negocios presenta varios beneficios con enfoque tanto para el cliente como para la empresa, puesto que al buscar satisfacer las necesidades del cliente se optimizan procesos y de esta manera la empresa obtiene ingresos.

Mejía (2017) identifica los siguientes beneficios:

- Acelerar y mejorar la toma de decisiones
- Optimizar los procesos de negocio internos
- Aumentar la eficiencia operativa
- Impulsar nuevos ingresos
- Generar ventajas competitivas respecto a competidores. (Mejía, 2017)

1.4.2 El proceso de toma de decisiones y los sistemas de inteligencia de negocio

Las empresas se enfrentan a tomas de decisiones constantemente pero no siempre son las correctas o las mejores, pero al implementar BI en la organización esto puede cambiar, pero ¿Cuál es el proceso de toma de decisiones? Y ¿En que falle de la toma de decisiones se implementa BI?

La toma de decisiones consta de cuatro fases: la fase de razonamiento, fase del diseño, fase de elección y fase de implementación. Primero se debe recolectar datos del entorno externo e interno para poder entrar a la primera fase, la fase de razonamiento implica integrar los datos en información y esta sea analizada de manera cuantitativa y cualitativa para detectar el problema y oportunidad, si no se logran detectar se regresa al proceso de examinación de información, pero si se logra detectar el problema y la oportunidad se continua a la segunda fase, la fase de diseño es donde se encuentran as alternativas y se analiza que curso seguir para continuar con la fase de elección, en esta fase se comparan las alternativas a seguir y se elegi una de ellas, pero como saber si la decisión que se elige es la correcta para ello se debe implementar BI (Inteligencia de Negocios) de esta manera se podrán presentar distintos escenarios y perspectivas para analizar y tomar la mejor decisión, luego se pasa a la fase de implementación para poner en practica la solución seleccionada, en este punto se pude llegar al éxito o al fracaso, si se llega el éxito entonces se culmina el proceso, pero si no es así se deberá regresar a la fase de diseño y continuar con las fases.

1.4.3 Inteligencia y análisis de negocios para el soporte de decisiones

Laudon, Kenneth C y Laudon, Jane P (2012) en su libro titulado “Sistemas de información Gerencial” establece seis elementos en el entorno de **inteligencia de negocios**.

La inteligencia de negocios es una herramienta que permite tomar **la** mejor decisión que beneficie al negocio para lo cual **se lleva a cabo el análisis de la información** que genere la empresa, pero ¿cómo se lleva a cabo dicho proceso?

Primero se debe recolectar los datos que genere la organización, estos pueden ser extraídos de fuentes internos o fuentes externas, pueden ser datos estructurados o no estructurados para luego ser almacenados en bases de datos transaccionales (global o específica) para pasar a ser analizadas, existen varias herramientas de análisis de negocios como Modelos estadísticos, Minería de Datos, OLAL e Informes de producción que pueden ser implementadas y continuar con el alineamiento de los procesos de la empresa mediante

estrategias de Negocios, Administración del desempeño, Cuadro de mando integral y pronósticos pero se debe elegir una plataforma de desarrollo ya sea MIS, DSS, ESS para así sean presentados los resultados mediante una interfaz para que sea fácil entender por parte del usuario sean estos informes, gráficos, reportes, etc.

1.4.4 Componentes de la inteligencia de negocios

BI consta de varios componentes necesario para su correcta implementación para una empresa, Cano (2007) establece los siguientes componentes:

- Fuentes de información.
- Proceso ETL.
- El propio Datamart o datawarehouse.
- El motor OLAP.
- Las herramientas de visualización. (Cano, 2007)

Fuentes de información: La fuente de información se accede de los sistemas transaccionales u operacionales de la empresa adquirida por departamentos o por información externa con datos que no son necesariamente automatizados.

Proceso ETL: es el proceso que lleva a cabo la extracción, integración y carga de los datos para alimentar el datamart o datawarehouse.

Datamart o datawarehouse: Datamart, Almacén de datos con información departamentales una base de datos especialista. Datawarehouse, almacén de datos con información integrada que se ha generado dentro de la empresa.

Cubo OLAP: Es una solución BI la cual permite que la información que reside en el Datamart o datawarehouse sea analizada que cumpla con objetivos como consultas rápidas, análisis estadísticos, actualizaciones y proveer diferentes perspectivas.

Herramientas de visualización: herramientas donde se visualice correctamente los informes o reportes. Todos estos componentes serán más detallados conforme se vaya desarrollando el tema.

1.4.5 Proceso ETL

ETL son las siglas en inglés de Extract, Transform, Load que su traducción quiere decir Extraer, Transformar, Carga, estos procesos permiten realizar migración de datos que se encuentran en una base específica otra para su respectivo análisis y explotación.

El proceso ETL permite trasladar datos desde diferentes fuentes los cuales se modifican y cargan a un repositorio (Datamart o Datawarehouse) para que por medio de aplicativos puedan ser analizados y explotados con el objetivo de convertirlos en información por consiguiente en conocimiento.

1.4.5.1 Beneficios ETL

Sin más, es momento de definir a los procesos ETL como aquellos procesos que permite a una empresa u organización manipular datos; es decir, extraerlos desde un sistema origen, transformarlos y cargarlos en un sistema destino. Esto permite eficientes tareas y obtener análisis de calidad, alcanzando objetivos empresariales y mejora en la toma de decisiones. (Trujillo, 2018)

Al implementar el proceso ETL la empresa adquiere varios beneficios como:

- Mejorar la calidad de los procesos
- Integrar sistemas (fuentes de datos)
- Facilitar la toma de decisiones
- Visualizar diversas perspectivas para encontrar falencias y mejorarlas

1.4.5.2 Fases del ETL

Las fases del proceso ETL son las siguientes:

- Extracción
- Transformación
- Carga

Extracción: Es la adquisición de la información de distintas fuentes y formato para ser analizada para lo cual se debe revisar, verificar y de ser necesario convertir al formato requerido para continuar con la siguiente fase, se realiza la identificación de datos cambiados, lectura de datos, obtención de agregados y mantenimiento de metadata (granularidad de datos).

Transformación: Se modifican los datos para que puedan ser cargados de ser necesario se lleva a cabo el proceso de formatear y limpiar los datos así también como conversión de unidades, agregar columnas, selección de columnas necesarias, División de columnas, Traducir código, generar propios valores mediante cálculos, creación de claves y mantenimiento de metadata.

Carga: Paso los datos de una base (origen) a otra (destino) de ser necesario se sobrescribe información, se almacena gran volumen de información resumida de las transacciones considerando la calidad de carga, la consolidación de los datos, minimizar la cantidad de recursos, gestionar errores y mantenimiento de metadata.

Existen dos metodologías de diseño de DW/BI (Datawarehouse/Business Intelligence) una establecida por Inmon y otra por Ralph Kimball, Inmon propone un diseño más complejo con un enfoque descendente iniciar por el DW, implementar el proceso ETL y pasar a los DM, mientras que Kimball implementa un diseño de enfoque ascendente iniciar por DM es decir datos hechos temas, implementar el proceso ETL y alimentar el DW.

Es una metodología modular centrada en áreas específicas de forma independiente, implementando un modelo dimensional el cual consta con esquemas como estrella (una tabla de hechos rodeada por tablas de dimensiones siendo 100% des normalizada) y copo de nieve (normalizada).

1.4.6 Metodología Kimball

Los autores de The Data Warehouse Lifecycle Toolkit se basó en décadas de experiencia para desarrollar el enfoque del ciclo de vida dimensional del negocio. Elegimos el nombre porque reforzó varios de nuestros principios clave para el almacenamiento exitoso de datos. En primer lugar, los proyectos de almacenamiento de datos deben centrarse en las necesidades de la empresa. En segundo lugar, los datos presentados a los usuarios comerciales deben ser dimensionales. Con suerte, esto viene como no ¡sorpresa a cualquier lector en este momento! Finalmente, mientras que el almacenamiento de datos es un proceso continuo, cada proyecto de implementación debe tener un ciclo finito con un principio y un final específicos. (Kimball, 1996)

El ciclo de vida del negocio se basa en cuatro principios que son: Centrarse en el Negocio, Construir una infraestructura, Realizar entregas en incrementos significativos y Ofrecer solución completa.

Kimball propone la metodología que ayuda a simplificar el proceso de construcción de un Datawarehouse/BI denominado ciclo de vida. Con tres diferentes áreas de enfoque: Tecnología (Camino Superior), Datos (Camino del medio) y Aplicaciones de Inteligencia de Negocios (Camino Inferior).

El ciclo de vida de Kimball consta de etapas que son las siguientes:

Planificación: Donde se define el alcance, identifica y programa las tareas, planifica asigna los recursos y se lleva documentación del **plan del proyecto, monitoreo de los procesos, identificación de problemas.**

Análisis de requerimientos: definición de requerimientos los cuales son establecidos o identificadores a partir de diferentes metodologías aplicadas al personal técnico y de negocio.

Modelo Dimensional: es un modelado alternativo del modelo entidad-relación con enfoque simétrico permitiendo eficiencia en consultas, fácil entendimiento y resistencia al cambio. Está compuesto por dos tipos de tablas (hechos y dimensiones) con esquema estrella.

Diseño físico: se configura el sistema, se establecen tipos de almacenamiento y procesadores, instalación de software y convertir el modelo lógico a modelado físico (base de datos relacional).

Diseño del sistema ETL: Al diseñar correctamente el sistema ETL lo que permitirá aumentar la calidad e integración de los datos.

Especificación y desarrollo de aplicaciones BI: establecer el estándar de los informes a generar y establecer las aplicaciones de visualización y análisis de la información.

1.4.7 Modelo dimensional

El modelo dimensional tiene un impacto en la mayoría de los aspectos de la implementación de un almacén, comenzando con la traducción de los requisitos comerciales, a través de la puesta en escena de datos y, finalmente, a la presentación de un almacén de datos a través de aplicaciones analíticas. (Kimball, 1996)

El modelo dimensional nos permite modelar los datos para su respectivo análisis, es un esquema donde se muestra el hecho del Negocio y las actividades que este realiza para de esta manera relacionarlas integrando solo datos necesarios y estos sean trasladados a un DM O DW para generar información y sea explotada.

El modelo dimensional implementa dos tipos de esquemas; estrella y copo de nieve (normalizada), el esquema más utilizado es el modelo estrella ya es 100% des normalizada su estructura es una tabla de hechos en el centro y varias dimensiones a su alrededor.

2

Hecho: es el objeto a analizar, posee atributos llamados hechos o de síntesis, y son de tipo cuantitativo. Sus valores (medidas) se obtienen generalmente por la aplicación de una función estadística que resume un conjunto de valores en único valor. (Tamayo & Javier, 2006)

La tabla de hechos está compuesta por información medible y numérica y se alimenta de las tablas transaccionales de la base de datos.

2

Dimensiones: representan cada uno de los ejes en un espacio multidimensional. Suministran el contexto en el que se obtienen las medidas de un hecho. (Tamayo & Javier, 2006)

La tabla de dimensiones está compuesta solo por información descriptiva, no numérica, ni métricas. Es la información que describe el hecho del negocio y se alimenta las tablas maestras y paramétricas.

CAPÍTULO 2

2 METODOLOGÍA

2.4 Investigación tecnológica aplicada

La metodología Kimball, es una metodología aplicada y direccionada a la construcción de un almacén de datos (Data warehouse, Data Mart), dentro del proyecto se procederá al diseño de un data mart, lo que significa que para alcanzar dicho diseño, es necesario basarse en la metodología Kimball para después dar paso a la elaboración de un Cubo OLAP.

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 nos guiarán a lo largo de todo el proyecto (Gutiérrez, 2012)

El ciclo de vida dimensional del negocio está basado en 4 conceptos:

- Centrarse en el negocio
- Construir una infraestructura de información adecuada
- Realizar entregas en incrementos
- Ofrecer la solución

En este capítulo se detalla los pasos de la investigación que se cumplieron para llevar a cabo el desarrollo del proyecto, La metodología a implementar en el proyecto es la metodología Kimball, la cual consta de cinco fases: Planificación del proyecto, Definición de requerimientos, Diseño y Construcción. La fase que corresponde al despliegue o implementación del proyecto no va a ser considerada en el presente trabajo.

2.4.1 Planificación del proyecto:

El propósito del proyecto será:

- Mejorar la distribución de información dentro del módulo de matrícula del Instituto “Juan Bautista Aguirre”
- Generar información verídica y oportuna, dirigida al usuario gerencial en forma consolidada y representativa.
- Integrar información histórica existente de la base de datos transaccional
- Facilitar el proceso de toma de decisiones a la gerencia del instituto.

El equipo de trabajo realizo reuniones descritas a continuación, además de establecer un cronograma de actividades a realizar del proyecto.

Primera reunión: Se dio con la persona encargada de la administración de la base de datos del instituto tecnológico “Juan Bautista Aguirre”, quien facilito al acceso de la información quien indico la deficiencia en reportes.

Segunda Reunión: Con la presencia del experto en Business Intelligence el Ing. Jorge Rodas quien explico cuáles podrían ser los beneficios al diseñar una data mart e implementar BI en las empresas.

Tercera Reunión: Se estableció las relaciones de actividades y cronograma como se muestra a continuación.

2.4.1.1 Diagrama de Ganntt

En el siguiente diagrama de presenta la planificación del proyecto que se llevara a cabo, el diagrama está dado por etapas las cuales contienen ciertos procesos que se llevan a cabo por determinada cantidad de días.

2.4.2 Requerimientos del proyecto

- a) **Obtención de los requerimientos:** La información se obtuvo atreves de entrevistas directas realizadas al interesado en generar reportes académicos y mediante la consulta de la base de datos de la empresa.

Fuente de información: Entrevista

¿Qué realiza la empresa?

¿Quién o quienes demandan el hecho de la empresa?

¿Dónde se encuentra ubicada la empresa?

¿Cuál es el rol o función que desempeña?

¿Qué transacción realiza la empresa?

¿Cómo se maneja el proceso actualmente?

¿Considera que la información que genera la empresa es de apoyo ¹ para la toma de decisiones?

¿Qué tipo de análisis implementa la empresa?

Dentro del proceso de matriculación de estudiantes. ¿Cuáles son los datos que se solicitan con frecuencia para la elaboración de reporte?

¿Qué tipo de herramientas utilizan para la elaboración de reportes?

- b) **Descripción del problema:** El instituto tecnológico “Juan Bautista Aguirre” se encuentra en constante crecimiento tanto educativo como económico, para la toma de decisiones a nivel de matrículas existen problemas al momento de generar reportes, no se lleva un control de los estudiantes ingresados y egresados del instituto; ya que el instituto no cuenta con un sistema que genere reportes automáticos que utilice indicadores reales.
- c) **Análisis del proyecto:** Analizando la gran demanda en reportes del instituto tecnológico “Juan Bautista Aguirre”, requiere la integración de herramientas que generan información oportuna y verídica; que sirva como base para la toma de decisiones. (Nurul Ibtisam Yaacob, Juzlinda Ghazali, Ahmad Sofian bin Shminan, 2008), para dar solución a esta problemática se propone desarrollar un Sistema Gestión de Matricula basado en una base de datos transaccional, con el objetivo de ayudar al instituto tecnológico “Juan Bautista Aguirre”, a ser más competitiva, tener un mejor control sobre el instituto generando reportes eficientes y relevantes.
- d) **Datos transaccionales:** se debe incluir modelo actual de datos para de esta manera obtener descripción de los datos cuando se desarrolle el proceso ETL.
- e) **Dimensiones y jerarquías:** Se organizan las dimensiones por niveles de manera que se presenten datos más detallados.
- f) **Dimensiones y medidas:** Se establecen las dimensiones genéricas y las medidas que describen el hecho del negocio.

2.4.3 Diseño:

Dentro de esta fase los requerimientos que han sido planteados se verán cristalizados en modelos físicos para conectar la fuente de datos para el diseño del data mart.

- a) **Selección de herramientas:** La herramienta a utilizar el sql analysis services , por ser un software libre y es capaz de integrarse con otras herramientas

- b) **Diseño Físico del Data mart:** Se crea un data mart SGM (Sistema Gestión Matricula), utilizando indicadores provenientes de la base de datos transaccional utilizada por el instituto “Juan Bautista Aguirre”, se establecen las dimensiones y la tabla de hechos.
- c) **Documentación del modelo:** Presentar la unión de las dimensiones con la tabla de hechos de acuerdo al respectivo análisis.

2.4.4 Construcción:

- **Data Mart:** Es el nuevo repositorio de datos llamado “Dm_Matricula”, donde la información será extraída de la base de datos transaccional “Bd_[itsjbaed_jba]” que servirá como base para las respectivas consultas y procedimientos
- **Creación de ETL:** Para el proceso de Extracción , Transformación y Carga(Load) de los datos, se trabajara con la herramienta sql analysis services , donde se podrá ser uso y manipulación de la base “Bd_[itsjbaed_jba]”
- **Creación de Cubo OLAP:** Una vez culminado el proceso de extracion de datos , se procederá con la construcción del cubo OLAP , con sus respectivas relaciones e indicadores
- **Preparación de interfaz:** Culminado el diseño del cubo se procederá a utilizar sentencias de navegación como son: DRILL DOWN, DRILL UP, los cuales permitirán la exploración de datos.

CAPÍTULO 3

PROPUESTA DE SOLUCIÓN

3.1 Tema

Desarrollo de un cubo OLAP para el análisis del sistema de gestión de matrícula del instituto tecnológico Juan Bautista Aguirre del cantón Daule para la correcta toma de decisiones

3.2 Descripción de la propuesta de solución

El instituto tecnológico utilizaba la herramienta ofimática de Excel como medio para análisis de los datos que genera la empresa, pero no constaba con una herramienta para generar informes o reportes, además de no constar con un adecuado procesamiento de datos.

El instituto como cualquier otra empresa que maneja grandes volúmenes de datos presenta problemas como la poca implementación de tecnología para la recolección de datos y el procesamiento de ellos para convertirlos en información valiosa para la empresa en el proceso de toma de decisiones obtenida de las diferentes áreas para cumplir con los objetivos de la empresa.

La empresa mediante la implementación de Inteligencia de Negocios puede obtener varios beneficios como:

- Disminuir brechas de interacción entre el procesamiento de reportes y usuarios
- Minimizar la posibilidad de errores en el procesamiento de datos
- Eliminación de la duplicidad de datos en los sistemas de reportes, validando el ingreso de información desde los paquetes de migración.
- Mediante el uso de la información integrada y en tiempo real, el sistema BI está preparado para brindar respuestas ágiles y eficientes.
- Proveer de escalabilidad a la solución BI mediante la integración de datos reales provenientes del sistema transaccional
- Reducir tiempos de presentación de informes relacionados al sistema de matriculación.

- Los usuarios de alta gerencia podrán realizar proyecciones semanales, mensuales y anuales de la información almacenada en la base de datos dimensional de la solución BI.

Para llevar a cabo el desarrollo del presente proyecto se implementa varios ítem, procesos o módulos estableciendo una estructura a seguir para la elaboración del Data Mart para la correspondiente integración de los datos y así transformarlos en información, analizarlos y que tenga como consecuencia la elección de la mejor decisión que beneficie a la empresa.

Enmarcando la inteligencia de negocios para este caso, el Sistema de Gestión de Matriculas presenta la siguiente estructura basada en el alcance del proyecto:

- Fuentes de Datos
- Proceso ETL (Extracción, Transformación y Carga)
- Cubo OLAP
- Presentación (Informes / Reportes)

El módulo de Fuentes de Datos se centra en la obtención de los datos que genera la empresa enfocado en los procesos operacionales de esta.

El modulo del Proceso ETL es donde se ¹ lleva a cabo la extracción de los datos los cuales deben ser transformados al formato necesario, los datos seleccionados y la actualización de las modificaciones para proceder a ser cargados.

El modulo del Cubo OLAP es donde se presentan as diferentes perspectivas de las situaciones que pueden ocurrió para que estas sean analizadas correspondientemente. Cada movimiento es información (conocimiento).

El módulo de presentación es donde se lleva a cabo mediante interfaz la presentación de la información mediante informes y reportes para que se pueda analizar la información y así tomar la mejor decisión con el fin de ³ cumplir con los objetivos de la empresa.

3.3 Especificaciones técnicas

3.3.1 Base de Datos

⁶ Una base de datos es un repositorio donde se almacena información de todo tipo sin importar el tiempo de manera que esta esté ordenada sistemáticamente para así poder recuperar información para su respectiva explotación sea el caso.

La base de datos implementada es de suma importancia ya que se puede acceder a información histórica de la empresa la se hubiera perdido si se llevaba en un registro tradicional antiguo, una Base de Datos permite obtención rápida de información, reducir recursos físicos y el recurso más importante como el tiempo al momento de realizar consultas.

En el presente proyecto se obtuvo una base de Datos extensa con varias tablas, pero no relacionadas ya que llevan un método de cascada, como se puede observar en la siguiente imagen.

Por lo cual se procedió a elaborar la siguiente base transaccional para continuar con el proceso, el diagrama relacional de Base de Datos Transaccional se muestra de la siguiente forma:

3.3.2 Involucrados

Cada rol fue diseñado. Asignado y descrito de acuerdo a las necesidades que va a tener el proyecto

3.3.3 Requerimientos

Los requerimientos se obtuvieron a partir de la conclusión de la siguiente pregunta que consta en la entrevista realizada al DBA, la cual se especifica a continuación:

Dentro del proceso de matriculación de estudiantes. ¿Cuáles son los datos que se solicitan con frecuencia para la elaboración de reporte?

Se suele solicitar datos de los estudiantes con consultas como la cantidad de estudiantes por género, ocupaciones, etnia, discapacidad, entre otras, datos de carrera con consultas de los tipos de carreras, datos del periodo lectivo, académicos y matriculas incluyendo tipos de matrículas con consultas como la cantidad de matrículas por tiempo.

3.3.4 Información

A partir de los requerimientos se estableció el modelo multidimensional de la base transaccional creada partiendo de la base de datos obtenida al inicio del proceso.

3.3.5 Dimensiones

Dimensión Estudiante (Dim_Estudiante) consta con información primordial sobre el estudiante como sus nombres, sexo, ocupación, lugar de residencia y carrera.

Dimensión Carrera (Dim_Carrera) consta con datos como tipo de carrera, modalidad, nombre y periodo académico.

Dimensión Periodo Académico (Dim_PeriodoAcademico) consta de la descripción del periodo y su respectivo estado para consultas como las matriculas realizadas en cada periodo académico.

Dimensión Periodo Lectivo (Dim_PeriodoLectivo) consta de la descripción exacta del periodo y su estado y consultar las matriculas por el periodo lectivo.

Dimensión Tiempo (Dim_Tiempo) consta con el año, mes y día de la realización de la matriculas permitiendo consultas históricas de las matriculas registradas.

3.3.6 Hecho

Hecho Matricula (Hecho_Matricula) es la tabla del hecho del Negocio que es ente caso es matricula con medidas de Estudiante, Carrera, Tipo de Matrícula, Jornada, Periodo Académico, Periodo Lectivo y Tiempo a más de información propia como observación de la matriculación.

3.3.7 SQL Lenguaje

Antes de realizar el proceso de migración de los datos de la base transaccional a nuestra base multidimensional creamos las consultas de los datos que son necesarios para su respectivo análisis y explotación de los mismo generando información concreta y específica.

A continuación, se muestran las sentencias a implementar a la migración de datos:

La presente sentencia permite consultar los datos necesarios de la dimensión Estudiante

```
--Migración Estudiante-----  
Select      E.IdEstudiante      ,      E.ApellidoPaternoEstudiante  
'ApellidoPaterno'      ,      E.ApellidoMaternoEstudiante  
'ApellidoMaterno',  
E.NombresEstudiante'Nombres', CONVERT (varchar (50) , (YEAR (GETDAT  
E ()) - YEAR (FechaNacEstudiante))) 'Edad',
```

```

E.SexoEstudiante      'Sexo',      P.NombrePais      'Pais',
PR.NombreProvincia    'Provincia',    C.NombreCanton    'Canton',
DIS.NombreDiscapacidad 'Discapacidad',
E.idTipoFinanciamiento      'Financiamiento',
E.idOcupacionEstudiante    'Ocupacion' ,    E.EstadoEstudiante
'Estado'
from SGM_M_Estudiante E, SGM_P_Pais P , SGM_P_Provincia PR ,
SGM_P_Canton C ,
SGM_M_Discapacidad DIS  where E.IdPais = P.IdPais and
E.idProvincia= PR.IdProvincia and E.IdCanton=C.IdCanton
and e.idDiscapacidad = DIS.IdDiscapacidad

```

La presente sentencia permite consultar los datos necesarios de la dimensión Carrera

```

--Migración Carrera-----
SELECT      CONVERT(int,      C.IdCarrera)      'IdCarrera'      ,
TC.NombreTipoCarrera 'Tipo' ,
MC.NombreModalidadCarrera 'Modalidad',
C.NombreCarrera      'Nombre',C.TotalPeriodosCarrera
'PeriodoAcademio'
FROM      SGM_M_Carrera      C,      SGM_P_TipoCarrera      TC,
SGM_M_ModalidadCarrera MC
WHERE      C.IdTipoCarrera      =      TC.IdTipoCarrera      AND
C.IdModalidadCarrera=MC.IdModalidadCarrera
order by convert(int,C.IdCarrera)

```

La presente sentencia permite consultar los datos necesarios de la dimensión del Periodo Lectivo.

```

-- Migración Periodo Lectivo -----
select * from SGM_M_PeriodoLectivo
select      PL.IdPeriodoLectivo      'IdPeriodoLectivo',
PL.DescripPeriodoLect 'Descripcion',

```

```
PL.EstadoPeriodoLectivo 'Estado' from SGM_M_PeriodoLectivo PL
ORDER BY PL.IdPeriodoLectivo
```

La presente sentencia permite consultar los datos necesarios de la dimensión del Periodo Lectivo.

```
-- Migración Periodo Académico-----
select * from SGM_M_PeriodoAcademico
Select IdPeriodoAcademico'IdPeriodo', NombrePeriodoAcademico
'Nombre',
EstadoPeriodoAcademico 'Estado' from SGM_M_PeriodoAcademico
```

La presente sentencia permite consultar los datos necesarios de la dimensión Tiempo

```
-- Migración Tiempo-----  
create view HechoT  
as  
select RANK() over (ORDER BY FechaMatricula) AS 'CODIGO',  
FechaMatricula,  
    DATENAME (YEAR ,FechaMatricula) 'anio', DATENAME (MONTH  
,FechaMatricula) 'mes',  
DATENAME (DAY ,FechaMatricula) 'dia' from  
SGM_T_Maricula group by FechaMatricula  
La presente sentencia permite consultar los datos necesarios  
del Hecho Matricula (hecho del Negocio)
```

```
-- Migración Hecho Matricula-----  
SELECT E.IdEstudiante , C.IdCarrera ,PA.IdPeriodoAcademico,  
PL.IdPeriodoLectivo , convert(int,T.CODIGO) 'CodigoEdad' ,  
TM.IdTipoMatricula 'DescripcionMatricula' ,  
J.IdJornada  
FROM SGM_T_Maricula M , SGM_M_Estudiante E,  
SGM_M_Carrera C, SGM_M_PeriodoAcademico PA,  
SGM_M_PeriodoLectivo PL , HechoT T , SGM_M_Jornada J ,  
SGM_M_TipoMatricula TM  
WHERE E.IdEstudiante = M.IdEstudiante AND M.IdCarrera =  
C.IdCarrera AND  
M.IdPeriodoAcademico = PA.IdPeriodoAcademico AND  
M.IdPeriodoLectivo = PL.IdPeriodoLectivo  
AND T.FechaMatricula = M.FechaMatricula AND M.IdJornada =  
J.IdJornada  
AND M.IdTipoMatricula = TM.IdTipoMatricula ORDER BY  
IdEstudiante
```

3.3.8 Migración de Datos

En este punto se ejecuta la migración de datos de una base a otra con información específica y transformación de los mismos, de ser necesario validando cargas de datos redundantes. A continuación se presenta la estructura del proceso de migración de cada dimensión y hecho del Negocio.

Migración Dim_Carrera

Migración Dim_Estudiante

Migración Dim_Periodo Lectivo

Migración Dim_PeriodoAcadémico

Migración Dim_Tiempo

Migración Hecho_Matricula, a diferencia de las demás migraciones, al migrar la tabla de hecho se lo hace directamente; es decir, la migración se la hace tomando todos los datos correspondientes de la ³base de datos transaccional, hacia el Data Mart debido a que estos datos son el resultado de las otras dimensiones con sus datos almacenados, es por esto que se recomienda empezar con el proceso ETL de una tabla de hecho al final

Flujo de control de las dimension y la tabla de hecho, se lo conoce como flujo de control por el orden secuencial en que ejecuta los procesos, el flujo de control esta compuesto por 7 tareas incluyendo la Tarea Ejecutar SQL, las demas tareas guaran el preceso de extraccion, transformacion y carga de cada una de las dimensiones que componen al Data Mart.

3.3.9 Funcionamiento

Al desarrollar el cubo OLAP se presentan las dimensiones a ser explotadas según los requerimientos establecidos. A partir del Cubo OLAP de matrícula se obtiene información que se convierte en conocimiento mediante la presentación de reportes de los movimientos de la información y datos de las dimensiones.

KPI:

Conocido también como **indicador clave de rendimiento**, son aquellas medidas que aportaron al proceso de elaboración del CUBO OLAP

Los KPI son utilizados porque:

- Permiten obtener información valiosa y en tiempo real.
- Ayuda a medir variables y cuantificar resultados.
- Permiten analizar la información de manera desglosado, tomando las mejores decisiones que generan óptimos resultados para la empresa o institución.

KpiTotal:

La función de este KPI es mostrar si el total de matrículas por año de cada periodo lectivo ha alcanzado el objetivo planteado por el equipo de gerencia del instituto tecnológico.

KpiAño:

Este Kpi ayudara al proceso de visualización y comparación entre el valor económico de las matriculas de cada año, teniendo así una visión en general sobre aquellos años en que el instituto tuvo su mayor acogida y por ende mayores ganancias.

Reportes:

El objetivo del siguiente reporte es comparar el total de carreras con más aceptación por año, utilizando el modo en que los estudiantes ven las carreras, esta puede ser dual, presencial, a distancia o dual-focalizada, las cuales pueden ser seleccionadas por el usuario.

En el siguiente reporte muestra el total de matrículas por año y si cumplen con el objetivo de la institución, para representar estos objetivos, el grafico de anillos muestra el total de matrículas por año, mientras que el grafico embudo muestra el porcentaje total de matrículas por año.

La prioridad para la administración del instituto son sus estudiantes, para aquello se diseñó una plantilla en donde se vea reflejado los estudiantes activos del instituto, en donde se pueda hacer búsquedas específicas, a la gerencia del instituto le interesa conocer cuáles son los estudiantes que poseen alguna discapacidad y la ocupación de los mismos ya sea que trabajen o estudien.

Con el siguiente reporte se logra calcular el total monetario de los periodos lectivos del instituto tecnologico desde el año 2014

3.4 Evaluación**Prueba de performance**

En esta prueba se miden y evalúan los tiempos de respuesta, los tiempos de transacción y otros requerimientos en función al tiempo. El objetivo de la prueba es verificar que se logren los requerimientos de performance. Esta es implementada y ejecutada para poner a punto los destinos de pruebas de performance como función de condiciones de trabajo o configuraciones de hardware. Para este sistema, las pruebas de performance son respecto a los tiempos de extracción de datos en la base de datos

Aspecto técnico

- El proceso de toma de datos (extracción)

La extracción de datos fue un proceso medianamente complejo ya que se debían manipular datos desde el origen llevando un análisis de los mismos, interpretando y verificando la validez de los datos al igual que su formato.

CONCLUSIONES

Implementar Inteligencia de Negocios brinda gran beneficio como la tomar la mejor decisión que permita a la empresa cumplir con cada uno de sus objetivos planteados, ejecutando el proceso ETL una vez obtenidos los datos necesarios para pasar al Datamart y generar reportes e informes una herramienta de análisis y explotador de información convirtiéndola en conocimiento relevante para la organización.

Del trabajo presentado se puede concluir que:

- Implementar herramientas de recolección de información es de suma importancia para tener claros los requerimientos.
- Existen distintos tipos de fuentes de información que sirven para alimentar un Datamart en distintos formatos.
- El proceso ETL va más allá de pasar datos de un lugar a otro, debido a que esto conlleva un proceso analítico dependiendo de lo que se busca realizar.
- Existen datos basura, no se lleva un control o no están correctamente validados campos que son necesarios para la empresa.
- Implementar una herramienta de análisis como la Inteligencia de Negocios permite mejorar la toma de decisiones de la empresa, generando utilidad y calidad.

Elaboración cubo de información Sistema Gestión de matrícula

INFORME DE ORIGINALIDAD

8%

INDICE DE SIMILITUD

4%

FUENTES DE
INTERNET

0%

PUBLICACIONES

6%

TRABAJOS DEL
ESTUDIANTE

FUENTES PRIMARIAS

1	Submitted to Universidad Cesar Vallejo Trabajo del estudiante	2%
2	docplayer.es Fuente de Internet	1%
3	Submitted to Universidad de San Buenaventura Trabajo del estudiante	1%
4	www.revistas.pucp.edu.pe Fuente de Internet	1%
5	studylib.es Fuente de Internet	1%
6	Submitted to Universidad Tecnologica del Peru Trabajo del estudiante	1%
7	reci.org.mx Fuente de Internet	1%
8	bdigital.unal.edu.co Fuente de Internet	<1%
9	Submitted to Systems Link Trabajo del estudiante	<1%

10

Submitted to Universidad Cooperativa de Colombia

Trabajo del estudiante

<1%

Excluir citas

Activo

Excluir coincidencias

< 20 words

Excluir bibliografía

Activo