

Universidad de las Ciencias Informáticas
Facultad Mártires de Artemisa



Título: Mercado de datos para el apoyo a la toma de decisiones en el comercio exterior de servicios del Ministerio del Comercio Exterior y la Inversión Extranjera (MINCEX).

**Trabajo de Diploma para optar por el título de
Ingeniero en Ciencias Informáticas**

Autores: Mirna Martínez Labrador

Yailín Avilés Avilés

Tutor: Ing. Mario Redonavich Garbey

Co-tutora: Ing. Ygraine Borges Ramírez

Artemisa, junio del 2012

“Año 54 de la Revolución”

Declaración de Autoría.

Declaramos que somos los únicos autores de este trabajo y autorizamos a la Universidad de las Ciencias Informáticas a hacer uso del mismo en su beneficio.

Para que así conste firmamos la presente a los ____ días del mes de _____ del año _____.

Mirna Martínez Labrador

Autor

Yailín Avilés Avilés

Autor

Ing. Mario Redonavich Garbey

Tutor

Ing. Ygraine Borges Ramírez

Co-Tutor



"Solo es útil el conocimiento que nos hace mejores."

Sócrates, filósofo griego.

Agradecimientos

A mis padres Mirta y Wilfredo por aconsejarme y apoyarme en todas las decisiones.

A mis abuelos, tíos, y primos porque de una forma u otra contribuyeron en mi formación como persona.

A Manso por quererme y ser paciente durante estos 5 años.

A mi compañera de tesis Yailin por animarme y darme fuerza en los momentos difíciles, sin ella no hubiese sido posible.

A mi tutor Mario Redonavich por el apoyo, dedicación y recomendaciones oportunas para la realización de esta tesis.

A todas mis amistades que compartieron junto a mí las tristezas, alegrías y horas de estudio en especial Lisi, Daily, Armando, Evelio, Yelina, Mairelis, Yadilka en fin a todos los quiero y nunca los olvidare.

Al equipo de trabajo: Yamila y Danelis por todos sus consejos y ayuda.

Al colectivo de profesores de la FRA que estuve presente en mi formación como ingeniera.

A todos gracias, Mirna

Agradecimientos

**A mi mamá por depositar toda su confianza en mí y creer que yo si podía
llegar hasta aquí.**

**A mis abuelos que siempre estuvieron presentes para mí y me enseñaron a
ser una mejor persona.**

A mi tito por enseñarme que cuando se quiere se puede.

A mi suegra que sufrió cada derrota mía como si fuese suya.

**A mis tíos y primos que supieron escuchar mis problemas y me dieron
ánimo para seguir adelante.**

**A mis sobrinas preciosas por hacerme recordar cada día que todos tenemos
un niño escondido en el alma.**

**A todas mis amistades que compartieron junto a mí las tristezas y alegrías
que trae consigo estar separados de la familia, en especial Vivi, Yiri y Lali.**

**A mi compañera de tesis Mirna por su paciencia y por enseñarme que las
cosas hay que hacerlas rápido y bien.**

**A mi tutor Mario por ser paciente y dedicarse con tanta entrega a sus
alumnas. Gracias profe.**

**A todos los profesores que hicieron posible mi preparación como futura
ingeniera.**

A la Revolución por brindarnos esta oportunidad, única en el mundo entero.

A todos, gracias.

Yailín

Dedicatoria

A mis padres, por ser parte inseparable de mi vida.

Mirna

A mi mamita por ser mi motor impulsor, por apoyarme en cada momento difícil de mi carrera, por dejar de vivir su vida para vivir la mía. Recuerda que el lucero que más brilla en el cielo tiene tu nombre... yo se lo puse.

Sencillamente te quiero.

A mi abuelo por ser mi ángel custodio y enseñarme desde pequeña que el estudio es lo primero, por velar mis noches de sueños y protegerme de todo mal. Me he quedado sin corazón porque tú te lo llevaste. Te quiero mi viejito lindo, más allá de la eternidad.

A mi tito por tener tanta paciencia durante estos cinco años, por saber esperar, por todo ese amor que me has entregado incondicionalmente, eres mi bebé único e irrepetible. Te quiero infinito.

A todos mis compañeros que iniciaron la carrera junto conmigo y que por cosas del destino no están hoy aquí. Nunca los voy a olvidar.

Yailín

Resumen

Con el desarrollo de la informática y el aumento de su impacto social cada vez son más las instituciones que deciden incorporar sistemas de apoyo a la toma de decisiones, este es el punto de partida de los almacenes de datos (AD) los cuales proporcionan un soporte de almacenamiento global único para una empresa.

Este trabajo se basa en la implementación de un mercado de datos (MD) de estadísticas de servicios para el comercio exterior de servicios del Ministerio del Comercio Exterior y la Inversión extranjera (MINCEX), en el cual se analizan las cifras de servicios con el objetivo de conocer el estado del comercio exterior de servicios del país con el resto del mundo.

La aplicación garantizará centralizar toda la información referente a las estadísticas de servicios, para facilitar la accesibilidad, el análisis estadístico, la realización de indicadores calculables y apoyar a la toma de decisiones a nivel nacional.

Se definió como metodología la propuesta por Kimball, como sistema gestor de base de datos Postgres SQL v9.0, Visual Paradigm v6.4 como herramienta de modelado y la Suite de Pentaho como herramienta de inteligencia de negocio. Durante el desarrollo de la aplicación se realizó un estudio previo de las tecnologías, herramientas y aplicaciones existentes que ayudaron en la realización de las funcionalidades antes mencionadas.

Palabras claves: almacén de datos, mercado de datos, estadísticas

Índice

Introducción	1
Capítulo 1: Fundamentos teóricos de los mercados de datos.	6
1.1 Conceptos relacionados con el tema	6
1.2 Sistemas de apoyo a la toma de decisiones (DSS)	7
1.3 Toma de decisiones en el comercio exterior de servicios.	8
1.4 Estado actual de los almacenes de datos como apoyo a la toma de decisiones	10
1.5 Almacenes de datos	10
1.6 Minería de datos.....	12
1.7 Mercado de datos.....	12
1.8 Modelos de almacenamiento de datos para los AD.....	13
1.7.1 Esquemas que se encuentran dentro del modelo dimensional	15
1.8 Tipos de almacenamiento de datos y consulta para AD	16
1.9 Herramientas de modelado.....	18
1.10 Sistemas gestores de bases de datos.....	20
1.11 Herramientas de inteligencia de negocio.....	21
1.12 Metodología	23
Capítulo 2: Análisis y diseño del mercado de datos estadísticas de servicios para el MINCEX	25
2.1 Procesos del negocio	25
2.2 Reglas del negocio.....	25
2.4 Áreas de análisis	26
2.5 Requisitos	26
2.5.1 Requisitos Informativos.....	26
2.5.2 Requisitos funcionales.....	27
2.5.3 Requisitos no funcionales	28
2.6 Casos de uso del sistema.....	29
2.7 Diseño de la solución	30
2.7.1 Modelo de datos dimensional.....	30
2.7.2 Especificación de las dimensiones.....	31
2.7.3 Especificación de los hechos	31
2.7.4 Seguridad del mercado de datos.....	33
2.7.5 Estrategia de recuperación y respaldo	33
2.8 Arquitectura del sistema	33
2.9 Arquitectura de información.....	35
2.10 Vista de despliegue	36
Capítulo 3: Implementación del mercado de datos estadísticas de servicios para el MINCEX	38

3.1 Estándar de código.....	38
3.2 Proceso de extracción, transformación y carga.....	40
3.2.1 Estrategia.....	41
3.3 Vistas materializadas.....	43
3.4 Implementación del subsistema de visualización de datos.....	44
3.4.1 Administración del sistema.....	48
Capítulo 4: Validación del mercado de datos.....	50
4.1 Pruebas de Software.....	50
4.1.1 Diseño de los Casos de Prueba.....	50
4.2 Resultados de las Pruebas.....	64
4.3 Resultados y Funcionalidades Obtenidas.....	65
Conclusiones generales.....	66
Recomendaciones.....	67
Bibliografía referenciada.....	68
Bibliografía consultada.....	70
Anexos.....	¡Error! Marcador no definido.
Glosario de términos.....	¡Error! Marcador no definido.

Índice de figuras

Figura 1. Esquema en estrella.....	15
Figura 2. Esquema en copo de nieve.....	16
Figura 3. Diagrama de Caso de Uso del Sistema (DCUS).....	30
Figura 4. Tabla del hecho importación.....	32
Figura 5. Tabla del hecho intercambio.....	32
Figura 6. Tabla del hecho exportación.....	32
Figura 7. Arquitectura del sistema.....	34
Figura 8. Mapa de navegación.....	35
Figura 9. Diagrama de despliegue.....	36
Figura 10. Proceso de ETL.....	41
Figura 11. Cubos involucrados en el proceso.....	45
Figura 12. Vista de análisis.....	46
Figura 13. Reporte definido en la solución propuesta.....	47
Figura 14. Reporte ad_hoc.....	47
Figura 15. Gráfico de barra.....	48

Índice de tablas

Tabla 1. Roles y permisos.....	33
Tabla 2. Caso de prueba “Visualizar los indicadores del boletín para las cifras de servicios”.....	54
Tabla 3. Caso de prueba “Visualizar los indicadores del proceso de importación de servicios”.....	55
Tabla 4. Caso de prueba “Visualizar los indicadores del proceso de exportación de servicios”.....	56
Tabla 5. Caso de prueba “Visualizar los indicadores del intercambio comercial de servicios”.....	56
Tabla 6. Caso de prueba “Visualizar los indicadores del proceso de cuadro”.....	58
Tabla 7. Caso de prueba “Visualizar los indicadores de donaciones exportadas”.....	58
Tabla 8. Caso de prueba “Visualizar los indicadores de donaciones importadas”.....	59

Tabla 9. Caso de prueba “Visualizar errores en los indicadores de las cifras de servicios”	60
Tabla 10. Caso de prueba “Visualizar los indicadores de serie de tiempo”	64

Introducción

Desde los inicios de la humanidad el hombre sintió la necesidad de almacenar y seguir el rastro de la información. Con la evolución de la sociedad surgieron métodos que favorecían y facilitaban la consulta de pequeñas cantidades de datos. Estos métodos fueron perfeccionados con el paso de los siglos, aunque existían insuficiencias a la hora de consultar grandes volúmenes de información. En la actualidad uno de los aspectos más importantes y con una fuerte tendencia a desarrollarse es el soporte que pueden brindar los sistemas de información para ayudar a las personas a analizar determinados problemas y a encontrar la solución para estos.

El desarrollo de las tecnologías de información y las comunicaciones (TIC) ha provocado la dependencia de aplicaciones donde se utilice el computador. Con el avance de la informática y el aumento de su impacto social, cada vez son más las instituciones u organizaciones que eligen incorporar sistemas de apoyo a la toma de decisiones, logrando así una mayor dinámica en sus procesos de negocio.

La competencia en el mundo empresarial actualmente es cada vez más estrecha. La creciente evolución y desarrollo de las ciencias de la información ha conducido a una situación en la que las organizaciones disponen de grandes volúmenes de datos con información histórica, almacenados en soportes informáticos. Así, una vez satisfecha la necesidad de disponer de un sistema de apoyo a la toma de decisiones para la gestión de datos, las organizaciones exigen más prestaciones a sus sistemas, y contemplan la posibilidad de poder extraer la información histórica almacenada que les permita analizar la organización, prever su evolución y tomar decisiones estratégicas para el futuro.

Este es el punto de partida de los almacenes de datos (AD) los cuales proporcionan un soporte de almacenamiento de datos global y único para una empresa. Facilitan el análisis de los datos mediante herramientas que permiten obtener información del mismo para apoyar el proceso de toma de decisiones. Brindan la posibilidad de proveer modelos predictivos basados en la información histórica de los procesos del negocio. Poseen procesos de optimización que garantizan la rapidez de consulta.

Cuba desde el triunfo de la revolución ha reconocido que el uso de las tecnologías y su explotación son de vital importancia para el desarrollo de una nación. Con este fin se han creado diversos centros tecnológicos para satisfacer las necesidades de información de todas las personas y esferas de la sociedad.

La Universidad de las Ciencias Informáticas (UCI) creada en el año 2002 y compuesta por varias facultades está a la vanguardia en la producción de software. Es un motor impulsor para fomentar la industria del software en Cuba. Son múltiples los proyectos productivos que se realizan en ella que contribuyen a la adquisición de nuevos conocimientos y tecnologías de avanzada como los AD.

Actualmente la facultad regional Mártires de Artemisa se encuentra inmersa en la realización del almacén de datos para el Ministerio de Comercio Exterior y la Inversión Extranjera (MINCEX), el cual constituye un elemento esencial en la política económica del país dando respuesta a la necesidad del estado cubano de asumir el manejo y control de las importaciones y exportaciones.

El MINCEX tiene como objetivo esencial el de preparar y proponer la política integral del estado en cuanto a la actividad de comercio exterior, la creación de empresas mixtas, la colaboración económica con otros países, organizaciones, asociaciones extranjeras y las inversiones que se negocien.

En el ministerio se analiza el modelo 0793 que contiene los indicadores específicos del comercio exterior de servicios con el objetivo de conocer el estado de las estadísticas de servicios del país con el resto del mundo, para así tomar decisiones estratégicas que contribuyan al desarrollo de la economía.

Los especialistas cuentan con un almacén de datos para analizar las estadísticas de servicios, el que a su vez hace uso de herramientas de escritorio y del esfuerzo humano.

Actualmente en dicha dirección, se aprecia la siguiente **situación problemática**: La toma de decisiones en el comercio exterior de servicios es insuficiente, lo cual se hace evidente cuando los especialistas no pueden acceder a la información simultáneamente, lo que provoca atrasos en la entrega de informes. Los datos estadísticos de servicios a menudo presentan inconsistencia debido a que los cálculos complejos son realizados de forma manual y no se realiza una adecuada validación de los mismos. El almacenamiento de los datos no cuenta con la

seguridad requerida lo que propicia que estos sean vulnerables ante posibles cambios que no van en correspondencia con la información que se maneja dentro del departamento. La manera de visualizar los datos se ve afectada ya que una vez presentados ante el usuario no pueden ser modificados restringiéndole la manera de analizar la información ya sea resumida o detallada.

Lo anteriormente expuesto deriva como **problema científico**: ¿Cómo apoyar la toma de decisiones en el comercio exterior de servicios del MINCEX?

Se reconoce como **objeto de estudio** de la presente investigación: Sistemas de apoyo a la toma de decisiones, delimitando como **campo de acción** la toma de decisiones en el comercio exterior de servicios.

El **objetivo general** de la investigación es: Desarrollar un mercado de datos que apoye la toma de decisiones en el comercio exterior de servicios del MINCEX.

De este objetivo se derivan los siguientes **objetivos específicos**:

1. Elaborar la fundamentación teórica del trabajo a realizar.
2. Construir el mercado de datos estadísticas de servicios para el MINCEX.
3. Validar el mercado de datos estadísticas de servicios para el MINCEX.

La **hipótesis** de la investigación es: si se desarrolla un MD “estadísticas de servicios”, se elevará la calidad de la toma de decisiones en el comercio exterior de servicios del MINCEX.

Para lograr el objetivo de esta investigación y dar solución a la hipótesis antes expuesta, es necesaria la realización de las siguientes **tareas de la investigación**:

1. Establecimiento de los fundamentos teórico-metodológicos para el desarrollo de los sistemas de apoyo a la toma de decisiones.
2. Caracterización del proceso del comercio exterior de servicios en cuanto a la toma de decisiones en el MINCEX.
3. Caracterización de las metodologías, tecnologías y herramientas a utilizar en el desarrollo del MD estadísticas de servicios para el MINCEX.
4. Desarrollo del MD estadísticas de servicios para el MINCEX.

5. Validación de la contribución lograda en cuanto al apoyo de la toma de decisiones mediante la introducción del MD en un ambiente similar al MINCEX.

Los métodos científicos utilizados fueron:

Métodos teóricos

➤ **Analítico-Sintético:** se empleó como método para el desarrollo de un mercado de estadísticas de servicios para mejorar la toma de decisiones en el comercio exterior de servicios del MINCEX.

➤ **Hipotético-Deductivo:** este método se aplica a partir de la hipótesis elaborada en la investigación, teniendo en cuenta las características y los resultados esperados de los mercados de datos.

➤ **Histórico-Lógico:** se utilizó al estudiar el objeto siendo útil en la búsqueda, y conceptualización de la historia de los sistemas de apoyo a la toma de decisiones.

Métodos empíricos

➤ **Entrevista:** se utilizó para conocer la forma en que se maneja la información en el comercio exterior de servicios mediante una entrevista a los especialistas.(ver anexo 2)

➤ **Análisis documental:** permitió el estudio y análisis de todos los documentos e informes del comercio exterior de servicios, para obtener un mejor entendimiento de los procesos.

Población y Muestra

La **población** está constituida por el MINCEX, siendo la **muestra** tres especialistas del comercio exterior de servicios.

A partir de la descripción del problema de la investigación se definen las siguientes variables:

Variable independiente

➤ Mercado de datos

Variable dependiente

➤ Toma de decisiones

Para la operacionalización de las variables ver anexo1.

El **aporte práctico** consiste en un mercado de datos estadísticas de servicios para el MINCEX.

Estructura del documento: la presente investigación está compuesta por introducción que se dedica a presentar el problema científico, así como a exponer los elementos esenciales del diseño teórico y metodológico.

Capítulo 1: Fundamentos teóricos de los mercados de datos.

En este capítulo se realizará un estudio del estado del arte de la investigación, se evidenciarán los principales conceptos de los temas tratados en el desarrollo del almacén de datos, se presentará la metodología seleccionada y las herramientas utilizadas para el tratamiento necesario de la información.

Capítulo 2: Análisis y diseño del mercado de datos estadísticas de servicios para el MINCEX.

Se definirán los requerimientos del sistema, se diseñará el diagrama de casos de uso del sistema, el modelo de datos, los subsistemas de integración y visualización.

Capítulo 3: Implementación del mercado de datos estadísticas de servicios para el MINCEX.

Se implementarán los subsistemas de integración y visualización.

Capítulo 4: Validación del mercado de datos estadísticas de servicios para el MINCEX.

Se validará el mercado de datos a través de casos de prueba basados en casos de uso.

Finalmente se presentan las conclusiones, recomendaciones, bibliografía, anexos y glosario de términos.

Capítulo 1: Fundamentos teóricos de los mercados de datos.

Introducción

En este capítulo se aborda la situación actual en la que se encuentran los almacenes de datos como sistemas de apoyo para la toma de decisiones tanto en Cuba como en el mundo y se definen los principales conceptos relacionados con el tema. Además se detallan las metodologías y herramientas utilizadas en el desarrollo del mercado de datos estadísticas de servicios para el MINCEX.

1.1 Conceptos relacionados con el tema

Sistemas de información: son un conjunto de elementos interrelacionados con el propósito de prestar atención a las demandas de información de una organización, para elevar el nivel de conocimientos que permitan un mejor apoyo a la toma de decisiones y desarrollo de acciones. (Hernández, 2008)

Toma de decisiones: consiste en encontrar una conducta adecuada para resolver una situación problemática, en la que, además, hay una serie de sucesos inciertos. Una vez que se ha detectado una amenaza real, imaginaria, probable o no, y se ha decidido hacer un plan para enfrentarse a ella, hay que analizar la situación: hay que determinar los elementos que son relevantes y obviar los que no lo son y analizar las relaciones entre ellos y la forma que tenemos de influir en ellos. Este paso puede dar lugar a problemas, cuando se tienen en cuenta aspectos irrelevantes y se ignoran elementos fundamentales del problema. Una vez determinada cual es la situación problemática y analizada en profundidad, para tomar decisiones, es necesario elaborar modelos de acciones alternativas, extrapolarlas para imaginar el resultado final y evaluar este teniendo en cuenta la incertidumbre de cada suceso que lo compone y el valor que subjetivamente se le asigna ya sea consciente o automáticamente. (García, 2005).

Servicios: el término servicios se define acorde al Sistema de Cuentas Nacionales, de Balanza de Pagos y más recientemente por el Manual de Estadísticas del Comercio Internacional de Servicios, en los términos siguientes.

“Los servicios no son entidades independientes sobre las que se puedan establecer derechos de propiedad; asimismo, no pueden intercambiarse por

separado de su producción. Los servicios son productos heterogéneos producidos sobre pedido que generalmente consisten en cambios en las condiciones de las unidades que los consumen y que son el resultado de las actividades realizadas por sus productores a demanda de los consumidores. En el momento de concluir su producción los servicios han sido suministrados a sus consumidores. Los servicios son muchas veces difíciles de separar de otros productos con los que pueden estar vinculados en diverso grado". (ONE, 2010)

1.2 Sistemas de apoyo a la toma de decisiones (DSS)

La toma de decisiones es una de las actividades que se realizan con mayor frecuencia en el mundo de los negocios, para resolver una situación problemática, en la que, además, hay una serie de sucesos inciertos. El apoyo a una decisión significa ayudar a las personas que trabajan solas o en grupo a reunir inteligencia, generar alternativas y tomar decisiones. Desde 1960 se ha venido manejando el concepto de sistemas de apoyo a la toma de decisiones (DSS por sus siglas en inglés Decision Support System). A mediados de la década de 1970 este concepto se convirtió en un espacio de investigación definiendo a los sistemas de apoyo a las decisiones como un conjunto de programas y herramientas que permiten obtener oportunamente la información requerida durante el proceso de la toma de decisiones, en un ambiente de incertidumbre. (Soto, 1999)

A partir de 1990 aproximadamente, los almacenes de datos y el procesamiento analítico en línea (OLAP) comenzaron a ampliar el ámbito de los DSS, siendo hasta la actualidad una herramienta confiable para el apoyo a la toma de decisiones.

Características de los DSS

- Rapidez en el tiempo de respuesta, ya que la base de datos subyacente suele ser un AD corporativo o un MD, con modelos de datos en estrella o copo de nieve. Este tipo de bases de datos están optimizadas para el análisis de grandes volúmenes de información.
- Disponibilidad de información histórica. En estos sistemas está a la orden del día comparar los datos actuales con información de otros períodos históricos de la compañía, con el fin de analizar tendencias, fijar la evolución de parámetros de negocio.

- Integración entre todos los sistemas/departamentos de la compañía. El proceso de extracción, transformación y carga (ETL) previo a la implantación de un sistema de soporte a la decisión garantiza la calidad y la integración de los datos entre las diferentes unidades de la empresa. Existe lo que se llama: integridad referencial absoluta.
- Suelen ser sistemas de información interactivos y amigables, con altos estándares de diseño gráfico y visual, ya que están dirigidos al usuario final.

Tipos de DSS

Sistemas de información gerencial (MIS). Los sistemas de información gerencial (MIS, Management Information Systems), también llamados Sistemas de Información Administrativa (AIS) dan soporte a un espectro más amplio de tareas organizacionales, encontrándose a medio camino entre un DSS tradicional y una aplicación CRM/ERP implantada en la misma compañía.

Sistemas de información ejecutiva (EIS). Los sistemas de información ejecutiva (EIS, Executive Information System) son el tipo de DSS que más se suele emplear en inteligencia de negocios, ya que proveen a los gerentes de un acceso sencillo a información interna y externa de su compañía, y que es relevante para sus factores clave del éxito.

Sistemas expertos basados en inteligencia artificial (SSEE). Los sistemas expertos, también llamados sistemas basados en conocimiento, utilizan redes neuronales para simular el conocimiento de un experto y utilizarlo de forma efectiva para resolver un problema concreto. Este concepto está muy relacionado con la minería de datos.

Sistemas de apoyo a decisiones de grupo (GDSS). Un sistema de apoyo a decisiones en grupos (GDSS, Group Decision Support Systems) es "un sistema basado en computadoras que apoya a grupos de personas que tienen una tarea u objetivo común, y que sirve como interfaz con un entorno compartido". El supuesto en que se basa el GDSS es que si se mejoran las comunicaciones se pueden mejorar las decisiones. [Sistemas de Apoyo a la Toma de Decisiones]

1.3 Toma de decisiones en el comercio exterior de servicios.

La toma de decisiones es un proceso que está en correspondencia con la presentación de la información de servicios, que permitirá a los directivos del MINCEX tomar medidas de menor a mayor trascendencia.

Los indicadores del comercio exterior de servicio se presentan a los especialistas mediante el archivo M793 con formato dbf, que es suministrado por la Oficina Nacional de Estadísticas e Información (ONEI) con periodicidad trimestral. Este archivo contiene los ingresos y egresos efectuados por país, empresa y servicio.

La información suministrada es sometida a un proceso de cuadro que consiste en validar los datos para garantizar la consistencia de la información y a su vez propiciar que la toma de decisiones se enfoque sobre datos verídicos. Posteriormente se realizan una serie de reportes que muestran el comportamiento de las estadísticas de servicios en cuanto a:

- Gráficos que muestren la tendencia del total de las exportaciones e importaciones por los períodos del año en valor.
- Gráficos que muestren el crecimiento con relación a igual período del año anterior.
- Tablas que muestren la variación absoluta a nivel de entidad de un trimestre con relación al anterior.
- Tablas que muestren a nivel de entidad la variación de un trimestre con el anterior en %.
- Estructura de las exportaciones e importaciones a nivel de entidades.
- Estructura de las exportaciones e importaciones a nivel de servicios a 5 dígitos de CPCU.
- Estructura de las exportaciones e importaciones a nivel de países

Todos estos reportes se plasman en un boletín que consiste en un documento formal el cual es enviado a los directivos sirviéndoles para tomar decisiones y establecer estrategias que permitan el crecimiento de la economía.

1.4 Estado actual de los almacenes de datos como apoyo a la toma de decisiones

Desde hace varias décadas las organizaciones empresariales han buscado soluciones que les ayuden a atender sus necesidades a la hora de tomar decisiones sobre sus procesos de negocio. Con el fin de lograr este objetivo en la década de los noventa se comenzaron a dar los primeros pasos en tecnología de AD.

Basándose en el estudio de esta tecnología y el desarrollo de la misma, gran cantidad de empresas han comenzado a migrar hacia su modernización para posicionarse a la delantera en la toma de decisiones ejemplo de ello están:

- El AD para la cadena multinacional de distribución Carrefour (España), que proporciona los datos de manera entendible, los costos laborales, la productividad y los ingresos por ventas. (Ortiz, 2009)
- El AD INE (Instituto Nacional de Estadísticas), está estructurado para facilitar el análisis de datos estadísticos, creación de informes para datos demográficos y realización de estudios complejos. (Rodríguez, 2010)

Cuba no ha quedado ajena a la tecnológica de los AD, ejemplo de ello se tiene:

- El AD comercial de la corporación CIMEX, el cual facilita el análisis de la información mediante la creación de reportes, visualización de los gráficos y la información, propiciando una adecuada toma de decisiones.
- El AD estadísticos de la Oficina Nacional de Estadísticas e Información (ONEI), el cual se dedica fundamentalmente a integrar la información estadística del país y visualizar la información requerida para el control del plan de la economía.
- En el XIII Concurso Nacional de Computación se presentó un MD de facturación perteneciente al AD Cubacel. (Burgaleta, 2007)

1.5 Almacenes de datos

La tecnología de AD ha venido madurándose y posicionándose como la variante más acertada para la realización de análisis de información histórica. Los AD surgieron para darle solución a los problemas que existían en las empresas o instituciones administrativas pues tenían grandes volúmenes de información, dificultando la toma de decisiones por parte de los directivos. En la actualidad las

definiciones más reconocidas son las planteadas por **William H. Inmon** y **Ralph Kimball**, las cuales se mencionan a continuación.

Según **William H. Inmon** (considerado por muchos el padre del AD), un AD es un conjunto de datos orientados por temas, integrados, variable en el tiempo y no volátiles, que tienen por objetivo dar soporte a la toma de decisiones. (Inmon, 2005)

Según **Ralph Kimball** (considerado el principal promotor del enfoque dimensional para el diseño de almacenes de datos), un AD es una copia de los datos transaccionales específicamente estructurada para la consulta y el análisis. (Kimball, 2007)

Partiendo de los conceptos mencionados anteriormente se puede apreciar que los autores hacen referencia a un mismo tema aunque lo expresan de manera diferente. Se puede afirmar que un AD no es más que un conjunto de datos históricos relacionados entre sí, orientados a temáticas o procesos de negocios almacenados de forma integrada, cuyas consultas apoyan la toma de decisiones.

Entre los principales aportes de un almacén de datos se encuentran:

- Proporciona una herramienta para la toma de decisiones en cualquier área funcional, basándose en información integrada y global del negocio.
- Proporciona la capacidad de aprender de los datos del pasado y de predecir situaciones futuras en diversos escenarios de cualquier esfera.
- Supone una optimización tecnológica y económica, en entornos de centros de información estadística o de generación de informes.

En la actualidad se puede afirmar que los avances alcanzados en el desarrollo de los AD, confirman que ya es una tecnología madura, estable y que soluciona las problemáticas presentadas en las empresas.

Características de los AD

Orientada a un tema: la información se clasifica en base a los aspectos que son de interés para la empresa. Siendo así, que los datos tomados están en correspondencia con los procesos de negocio.

Integrado: integra datos recolectados de diferentes sistemas operacionales de la organización y/o fuentes externas. Los datos deben ser consistentes siempre dentro del AD.

Variación de tiempo: los cambios producidos con el paso del tiempo en los datos almacenados también se van registrando en el almacén, lo que permite brindar una perspectiva histórica de la información, que abarca un espacio de tiempo indefinido.

No volátil: los datos que son almacenados no sufren ninguna actualización solo son incrementados.

1.6 Minería de datos

Debido a que la información en las bases de datos está más desestructurada que en otras fuentes de información, para su utilización es necesario un proceso de tratamiento y análisis exhaustivo. Estas tareas de análisis de la información de las bases de datos se denominan minería de datos (en inglés data mining).

La definición de minería más citada en la literatura sobre el tema expuesta por Fayyad en el año 1996 plantea que: "minería de datos es un proceso no trivial de identificación de patrones de datos válidos, nuevos, potencialmente usables y comprensibles". (Fayyad, 1996).

Es un mecanismo de explotación, consistente en la búsqueda de información valiosa en grandes volúmenes de datos. La minería de datos se centra en llenar la necesidad de descubrir el por qué, para luego predecir y pronosticar las posibles acciones con cierto factor de confianza para cada predicción. Es el análisis de archivos y bitácoras de transacciones, trabaja a nivel del conocimiento con el fin de descubrir patrones, relaciones, reglas, asociaciones o incluso excepciones útiles para la toma de decisiones.

1.7 Mercado de datos

Al hacer un análisis de la bibliografía existente sobre los AD se puede constatar que son muchos los autores que han expresado su opinión en cuanto a la definición del término mercado de datos (MD), pero una de las más abarcadoras es la propuesta por Ralph Kimball quien considera que constituye un conjunto

flexible de datos, idealmente basado en el dato más atómico posible (granular) para ser extraído de las fuentes operacionales y presentado en un modelo simétrico (dimensional), que es más resistente cuando se enfrentan con las más inesperadas consultas de los usuarios. (Kimball, 2007)

Un MD es una alternativa de solución al igual que AD a los problemas antes planteados porque el diseño y construcción son similares, además de poseer una secuencia común. La diferencia entre estas dos estructuras se basa principalmente en que los MD están enfocados en un área de negocio específica, mientras que un AD entrega información a nivel corporativo. (Inmon y col, 2007)

Se puede resumir, basado en los conceptos antes mencionados, que un AD está compuesto por varios MD y no son más que un conjunto de datos específicos referidos a un departamento que son creados para satisfacer las necesidades específicas de un grupo de usuarios.

Características de los MD

Entre las características de los MD se destacan:

- Posee usuarios limitados.
- Tiene un área específica.
- Mejora el tiempo de respuesta al usuario final por causa de la reducción en el volumen de información a ser accedido.
- Provee datos estructurados apropiadamente para satisfacer los requerimientos de las herramientas de acceso de usuario final.

1.8 Modelos de almacenamiento de datos para los AD

Para desarrollar un AD el modelo de almacenamiento que se utiliza es el modelo dimensional que tiene como finalidad lograr que la manipulación de los datos sea la correcta. A continuación se presentan características de este modelo.

El modelo dimensional: posee la misma información que el modelo relacional pero la organiza de forma diferente para garantizar la velocidad y eficiencia en la recuperación de la misma. Una de sus características principales es que no necesita una predefinición de los reportes debido a que se diseña de forma tal que cubra todas las variantes que los usuarios necesiten para consultar la

información almacenada. Este modelo se compone por hechos, dimensiones y medidas. (Ponniiah, 2009)

Tablas de hechos: la tabla de hecho es la tabla primaria en el modelo dimensional donde el rendimiento de las mediciones numéricas del negocio es almacenado. (Kimball y col, 2007) La principal condición que deben cumplir las tablas de hechos es que el hecho debe almacenarse de tal forma que su valor sea numérico y a su vez sea aditivo para así poder realizar cálculos sobre él, ya sea por ciento, sumas e igualdades.

Medidas: las medidas más útiles para incluir en una tabla de hechos son aquellas medidas numéricas que pueden calcularse con la suma de varias cantidades de la tabla. Es decir, las medidas candidatas son los datos numéricos, pero no cada atributo numérico es una medida candidata.

Tablas de dimensiones: las tablas de dimensiones definen cómo están los datos organizados lógicamente y proveen el medio para analizar el contexto del negocio. Representan los ejes del cubo, y los aspectos de interés, mediante los cuales el usuario podrá filtrar y manipular la información almacenada en la tabla de hechos. Las tablas de dimensiones están compuestas por un identificador único y al menos un atributo que describe los criterios de análisis relevantes de la organización, dentro de estas tablas puede existir:

Jerarquía: es un conjunto de atributos descriptivos que se rigen por un orden preestablecido. Una jerarquía implica una organización de niveles dentro de una dimensión, con cada nivel representando el total agregado de los datos del nivel inferior. Además, definen cómo los datos son sumados desde los niveles más bajos hasta los más altos. Una jerarquía puede contener todos los valores que se encuentran en una dimensión, pero no es necesario.

Granularidad: Es el nivel más bajo de información que será almacenado en la tabla de hechos e indica el grado de detalle asociado a un hecho particular. La granularidad depende directamente del número de dimensiones que se asocian a la tabla de hechos. El primer paso para diseñar una tabla de hechos es determinar la granularidad. El gran factor y el más decisivo de la granularidad es el tiempo, ya que mientras menor sea el intervalo de tiempo, mayor será el grado de detalle obtenido. Se deben considerar otros factores como la carga del

procesador, espacio de almacenamiento y satisfacer a cabalidad los requerimientos del cliente.

1.7.1 Esquemas que se encuentran dentro del modelo dimensional

Esquema en estrella

Su estructura base es similar a una tabla central y un conjunto de tablas que la atienden radialmente. Este modelo, resulta ser asimétrico, pues hay una tabla dominante en el centro con varias conexiones a las otras tablas. Las tablas dimensiones tienen sólo la conexión a la tabla hecho y ninguna más. (Herrera, 2007)

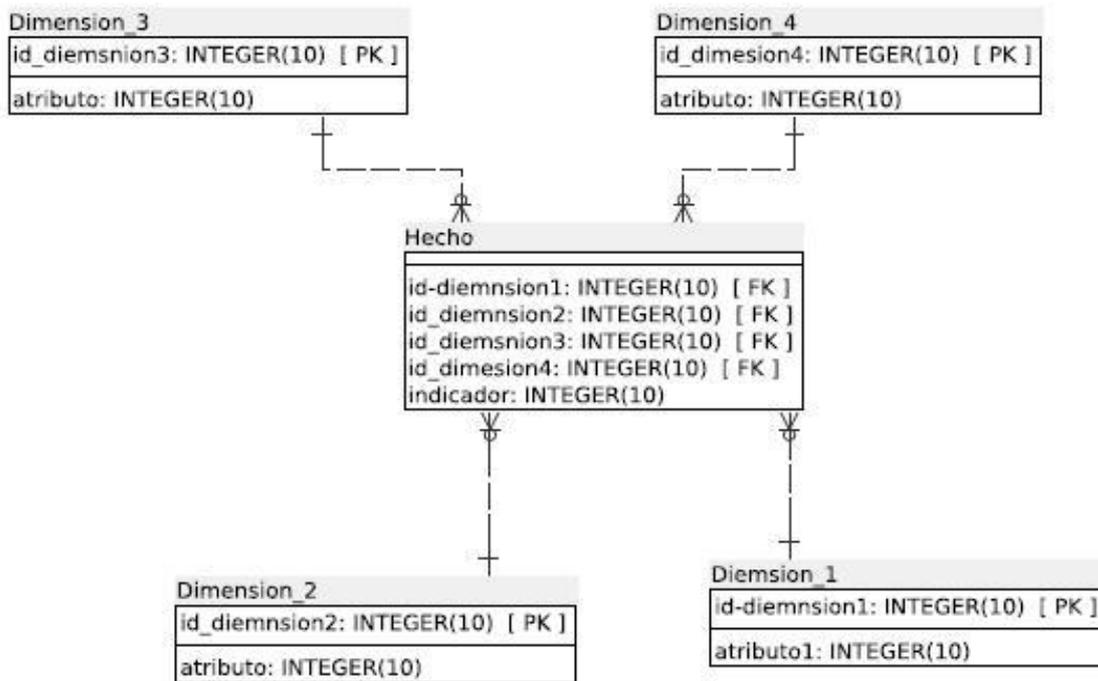


Figura 1. Esquema en estrella

Esquema en copo de nieve

Es una estructura más compleja que el esquema de estrella. Se da cuando las dimensiones se implementan con más de una tabla de datos. Aunque puede reducir espacio por la mínima redundancia de datos, tiene la contrapartida de peores rendimientos al tener que crear más tablas de dimensiones y más relaciones entre las tablas, lo que tiene un impacto directo sobre el rendimiento. (Technologies, 2010)

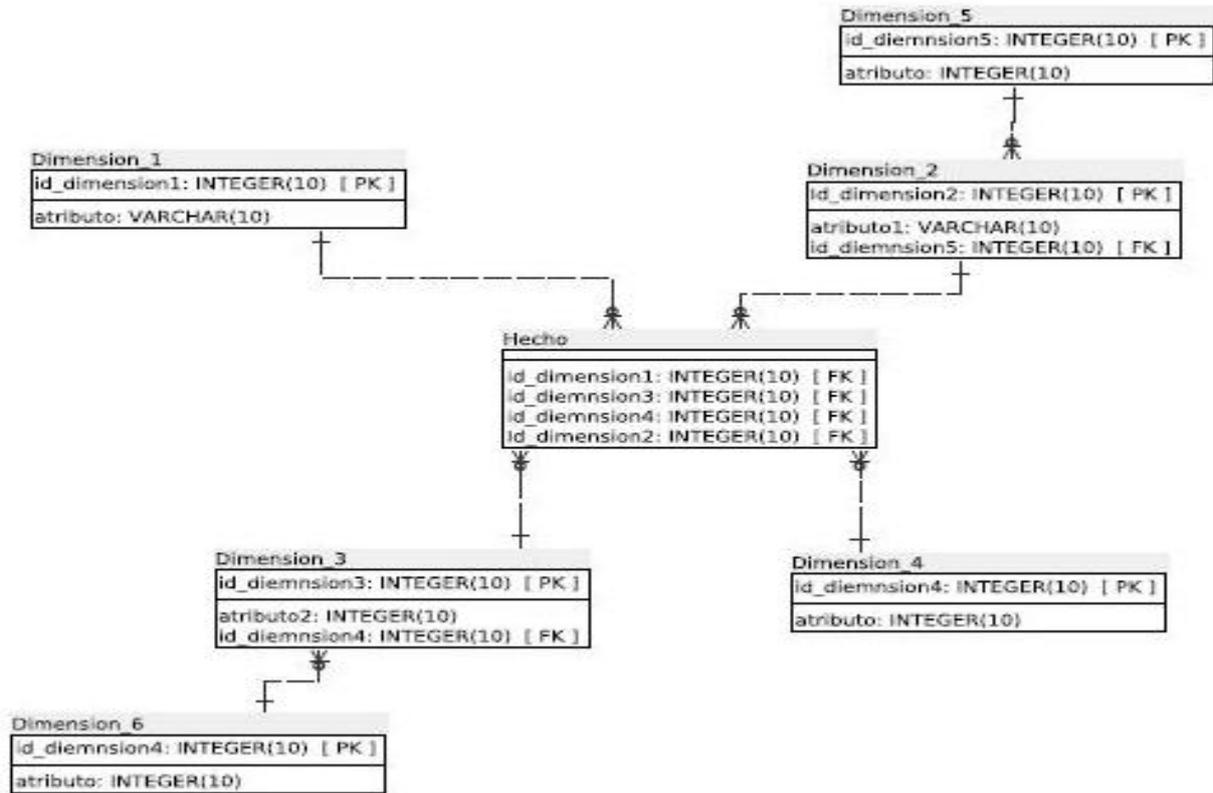


Figura 2. Esquema en copo de Nieve

Los datos dentro de la tabla dimensión, que proveen información del negocio o que describen alguna de sus características, son llamados datos de referencia. Entonces, se puede afirmar que una tabla de dimensión posee una clave primaria y uno o más datos de referencia.

Justificación de modelo de almacenamiento de datos a utilizar

Para la realización del AD del MINCEX se utiliza el esquema en estrella porque es el más simple, tiene mejor rendimiento y velocidad que el de copo de nieve además de que se adapta perfectamente a las necesidades del mercado de datos.

1.8 Tipos de almacenamiento de datos y consulta para AD

El Procesamiento Analítico en Línea (OLAP por sus siglas en inglés), es una solución que se utiliza en la inteligencia de negocio, cuyo principal objetivo se basa en lograr que el proceso de consultas de grandes cantidades de datos sea más rápido. Para ello, utiliza estructuras multidimensionales o cubos OLAP, que contienen datos resumidos de grandes bases de datos. De manera general, el

procesamiento analítico en línea puede verse como la síntesis, análisis y consolidación de grandes volúmenes de datos con un enfoque multidimensional.

Existen tres modelos para realizar este proceso: ROLAP, MOLAP y HOLAP. Seguidamente, se explican brevemente las características principales de cada uno de ellos.

ROLAP

En el Procesamiento Analítico Relacional en Línea, en inglés Relational Online Analytical Process (ROLAP), los datos son guardados en filas y columnas de forma racional. Este modelo presenta los datos a los usuarios en forma de dimensiones del negocio, con el objetivo de ocultar las estructuras de almacenamiento a los usuarios y presentar los datos multidimensionalmente. (Rodríguez, 2010)

El modelo ROLAP, es usado fundamentalmente sobre la información que no se consulta frecuentemente, debido a que resulta muy útil cuando se desea consultar información que se almacena durante muchos años.

MOLAP

El Procesamiento Analítico Multidimensional en Línea, en inglés Multidimensional Online Analytical Process (MOLAP), almacena los datos multidimensionalmente a diferencia del modelo ROLAP. (Rodríguez, 2010)

La estructura de los datos en este modelo, es estática para que la lógica al procesar el análisis multidimensional pueda ser basada en métodos bien definidos, con el propósito de establecer las coordenadas del almacenamiento de los datos.

Para realizar el acceso a la información almacenada de forma más rápida y efectiva, las estructuras de almacenamiento se organizan en grandes arreglos dimensionales, que son una copia de la fuente de datos y persisten físicamente en la misma estación de trabajo donde está instalado el AD.

HOLAP

El modo de almacenamiento Procesamiento Analítico Híbrido en Línea, en inglés Hybrid Online Analytical Process (HOLAP), como su nombre lo indica, es un híbrido entre los métodos ROLAP y MOLAP, que permite almacenar una parte de

los datos como en un sistema MOLAP y el resto como en uno ROLAP. (Leyva, 2010) Además posee dos tipos de particionamiento:

Particionamiento vertical: almacena las agregaciones como un MOLAP para mejorar la velocidad de las consultas, y los datos se detallan en ROLAP para optimizar el tiempo en que se procesa el cubo.

Particionamiento horizontal: en este modo, HOLAP almacena una sección de los datos. Normalmente, los datos más recientes se almacenan en modo MOLAP para mejorar la velocidad de las consultas y los datos más antiguos se guardan en ROLAP.

Justificación del modo de almacenamiento de datos a utilizar

Se decide seleccionar para el desarrollo de la investigación al modelo ROLAP para las consultas, debido a que el sistema gestor de base de datos seleccionado es Postgres y el mismo no soporta el modo de almacenamiento MOLAP, otra de las razones de usar ROLAP es que es capaz de usar datos pre-calculados si éstos están disponibles, o de generar dinámicamente los resultados desde los datos elementales si es preciso.

1.9 Herramientas de modelado

Las herramientas CASE son un conjunto de programas que brindan a los desarrolladores asistencia para el desarrollo del ciclo de vida de un proyecto. Con la introducción de estas herramientas se han podido confeccionar y modificar los diagramas que mejoran la calidad y productividad del software. Tienen como objetivo fundamental automatizar los aspectos clave en el proceso de desarrollo.

Dentro de las características más importantes de las herramientas CASE están:

- Alto rendimiento.
- Calidad de software desarrollado.
- Archivos de datos de conocimiento.
- Facilidades de uso.

Entre las herramientas CASE libres más utilizadas se pueden mencionar:

Umbrello: rápida, ligera, sencilla de usar, no se pone lenta cuando los proyectos son enormes. Requiere de más opciones de generación de documentación de los diseños modelados dentro de ella. Permite generar código en diversos lenguajes.

ArgoUML: es una herramienta profesional de lenguaje de modelado unificado (UML), de uso libre. Incluye todos los diagramas UML 1.3, permite ingeniería inversa de código Java y generación de código. Disponible para las plataformas más importantes e integración con Eclipse. Se ha desarrollado con los criterios de código abierto y licencia libre.

Características

- Provee toda la funcionalidad deseada en una herramienta para modelar bajo UML.
- Genera código en varios lenguajes.
- Sus diseños son exportables a XML y pueden ser importados por algunos Frameworks.

Visual Paradigm for UML: es una herramienta profesional para modelar utilizando UML, que implementa el ciclo de vida completo del desarrollo de software: análisis y diseño orientados a objetos, construcción, pruebas y despliegue. Ayuda a la construcción de aplicaciones más rápidas, con la calidad requerida y a un menor costo. Permite representar todos los tipos de diagramas de clases, código inverso, generar código desde diagramas y generar documentación. Esta herramienta proporciona además, abundantes tutoriales, demostraciones interactivas y proyectos UML.

Power Architect: herramienta de modelado de datos que fue creada por los diseñadores de almacenamiento de datos y tiene muchas características únicas dirigidas específicamente para el arquitecto de almacenamiento de datos.

Justificación de la herramienta de modelado a utilizar

Para el desarrollo del sistema se deciden tomar como herramientas de modelado el Visual Paradigm v6.4 y el Power Architect v3.0 debido a las ventajas que ambas ofrecen y que son herramientas distribuidas bajo licencia de software libre permisiva y con su código fuente disponible libremente. Visual Paradigm genera de forma automática códigos desde diagramas, además brinda la posibilidad de

generar documentación. Power Architect v3.0 permite a los usuarios de la herramienta ingeniería inversa de bases de datos existentes, realizar perfiles de datos en bases de datos de origen y generar automáticamente los metadatos de ETL.

1.10 Sistemas gestores de bases de datos

Un sistema gestor de base de datos (SGBD) es un programa, que suministra a los usuarios los medios necesarios para describir, recuperar y manipular los datos almacenados en una base de datos, manteniendo su integridad, confidencialidad y seguridad. Las funciones esenciales de un SGDB son la descripción, manipulación y utilización de los datos. Algunos de los sistemas gestores de bases de datos más conocidos son:

MySql: es un sistema de gestión de bases de datos relacional, licenciado bajo la GPL de la GNU. Su diseño multihilo le permite soportar una gran carga de forma muy eficiente. Este gestor de bases de datos es, probablemente, el gestor más usado en el mundo del software libre, debido a su gran rapidez y facilidad de uso. Esta gran aceptación es debida, en parte, a que existen infinidad de librerías y otras herramientas que permiten su uso a través de gran cantidad de lenguajes de programación, además de su fácil instalación y configuración.

Apache Derby: es un SGDB escrito en Java que puede ser empotrado en aplicaciones Java y utilizado para procesos de transacciones online. Tiene un tamaño de 2 MB de espacio en disco. Inicialmente distribuido como IBM Cloudscape, Actualmente se distribuye como Sun Java DB.

PostgreSQL: es un potente motor de bases de datos, que tiene prestaciones y funcionalidades equivalentes a muchos gestores de bases de datos comerciales. Es más completo que MySQL ya que permite métodos almacenados, restricciones de integridad, vistas, entre otras funcionalidades. Una de las características que comparte con otros motores de bases de datos es el hecho de ser multiplataforma y de tener varias herramientas para administrar.

Justificación de sistema gestor de base de datos a utilizar

Como sistema gestor de bases de datos se eligió el PostgreSQL v9.0 ya que puede ser ejecutado sobre la plataforma más usada en todo el mundo: Windows,

además de brindar avanzadas características en materia de seguridad, soporte de aplicaciones, seguimiento y control, rendimiento y almacenamiento de datos especializados. Entre las principales características se encuentran:

- La cantidad máxima de BD que permite es ilimitada.
- El máximo de campos por tabla es 250 a 1600 en dependencia de los tipos de datos usados.
- Alta capacidad de almacenar información, con una alta velocidad de respuesta ante consultas complejas y/o extensas.

1.11 Herramientas de inteligencia de negocio

Las herramientas de inteligencia de negocios es un tipo de software de aplicaciones diseñado para colaborar con la inteligencia de negocios (BI) en los procesos de las organizaciones. Específicamente se trata de herramientas que asisten el análisis y la presentación de los datos. Seguido se muestran varias herramientas libres de BI.

JasperReports: es una herramienta de creación de informes que tiene la habilidad de entregar contenido enriquecido a los usuarios en ficheros con extensión pdf, html, xls, csv y xml. Su propósito principal es ayudar a crear documentos de tipo páginas, preparados para imprimir en una forma simple y flexible.

Suite de Pentaho: ofrece una plataforma la cual incluye un servidor, herramientas clientes y tecnologías que posibilitan el espectro entero de la BI. Se considera una fácil solución de software que integra el poder de análisis e integración de datos para proporcionar puntos de vista que mejoran la toma de decisiones. Incluye, basado en roles, interfaces de auto-servicio para todo tipo de usuarios dentro de un marco bien definido y un punto centralizado de administración. Esto ayuda a las organizaciones a simplificar y acelerar el despliegue de inteligencia empresarial. También incluye la comunidad de Pentaho que es gratis y posee herramientas como ETL, análisis OLAP, los metadatos, la minería de datos, informes, cuadros de mando y una plataforma que permite crear complejas soluciones a los problemas empresariales.

LogiReport: aplicación de BI gratuita basada en Web de LogiXML. Entre las características más importantes de esta solución está su bajo coste, su rápida implementación y la sencillez de uso.

Justificación de la herramienta de inteligencia de negocio a utilizar

Como herramienta de Inteligencia de negocio se seleccionó la versión 3.6 de la suite de la comunidad de Pentaho la cual está compuesta por:

➤ **BI-Server v3.10:** funciona como un sistema basado en administración web de informes, el servidor de integración de aplicaciones y un motor de flujo de trabajo ligero. Está diseñado para integrarse fácilmente en cualquier proceso de negocio. Pentaho BI Server se considera una fácil solución de software que integra el poder de análisis e integración de datos para proporcionar puntos de vista que mejoran la toma de decisiones. Incluye, basado en roles, interfaces de auto-servicio para todo tipo de usuarios dentro de un marco bien definido y un punto centralizado de administración. Esto ayuda a las organizaciones a simplificar y acelerar el despliegue de inteligencia empresarial. Este sistema contiene además los servidores Mondrian y Apache Tomcat.

➤ **Kettle - Spoon v4.2 (Data Integration):** provee una consistencia, una sola versión de todos los recursos de información, que es uno de los más grandes desafíos para las organizaciones de hoy en día. Pentaho Data Integration permite una poderosa ETL (Extracción, Transformación y Carga). El uso de Kettle permite evitar grandes cargas de trabajo manual frecuentemente difícil de mantener y de desplegar.

➤ **Pentaho Report Designer v3.8.3:** editor basado en eclipse con prestaciones profesionales de calidad y con capacidad de personalización de informes a las necesidades de negocio destinado a desarrolladores. Incluye asistentes para facilitar la configuración de propiedades. Está estructurado de forma que los desarrolladores pueden acceder a sus prestaciones de forma rápida. Incluye un editor de consultas para facilitar la confección de los datos que serán utilizados en un informe.

➤ **Pentaho Metadata Editor v4.1:** es una herramienta que permite crear dominios de metadatos y modelos. El objetivo es mapear la estructura física de la base de datos a un modelo lógico de negocio.

➤ **Schema Workbench v3.3.0:** es un entorno visual para el desarrollo y prueba de cubos OLAP, los cuales proveen un mecanismo para buscar datos con rapidez y tiempo de respuesta uniforme, independientemente de la cantidad de datos en el cubo o la complejidad del procedimiento de búsqueda. Con esta aplicación, se puede configurar una conexión JDBC como el modelo físico, para luego elaborar el esquema lógico de manera simple y efectiva. Para ello, el entorno ofrece un editor de esquemas con la fuente de datos subyacente para su validación, además de permitir la ejecución de consultas MDX contra el esquema y la base de datos.

1.12 Metodología

Los AD con el paso del tiempo han alcanzado un lugar sobresaliente entre los sistemas destinados para el análisis de información histórica y apoyo a la toma de decisiones.

A su vez se han creado y perfeccionado metodologías que ayuden a la realización de los AD. En este sentido se destacan varias metodologías pero existen dos que son las más reconocidas en el mundo, Metodología de Kimball y Metodología de Inmon, en honor a sus creadores Ralph Kimball y William H. Inmon, personalidades más influyentes en el área de los AD.

A continuación se presentan varias metodologías:

Metodología Kimball: se centra esencialmente en el diseño de la base de datos y este a su vez se enfoca en la creación de las tablas de hechos. Esta metodología tiene gran aceptación en el mundo para el desarrollo de una solución de Inteligencia de Negocios, permite que su solución se realice en un período corto de tiempo. (Kimball y col, 2007)

Metodología de Inmon: propone la creación de un repositorio de datos corporativo como fuente de información consolidada, persistente, histórica y de calidad, donde la información se encuentre en 3FN. De esta forma, se crea primero un DWH del cual se van a nutrir los diferentes MD que van a satisfacer las necesidades específicas de cada proceso del negocio.

Metodología para el Diseño Conceptual de AD: tiene como aspecto novedoso la incorporación de los casos de uso para guiar el proceso de desarrollo,

definiendo además varias transformaciones para llevar del modelo relacional al dimensional, obteniendo las estructuras que posteriormente conforman el repositorio de datos. (Zepeda, 2008)

Justificación de la metodología a utilizar

La metodología a utilizar en la confección del AD para el MINCEX debe ser una metodología robusta que sea capaz de validar durante el transcurso de vida del proyecto que se cumplan los objetivos impuestos por el cliente, para ello se ha seleccionado la metodología de Kimball adaptándola a las características particulares del UCI y el centro de desarrollo Facultad Regional Mártires de Artemisa, los elementos que influyeron en esta decisión se exponen a continuación:

- Esta metodología divide el mundo de BI entre el hecho y las dimensiones es muy eficaz y conduce a una solución completa en un tiempo razonable.
- Es iterativo, donde se construye una pieza a la vez (mercado de datos) garantizando mayor velocidad de respuesta a los clientes.
- La forma de almacenar la información es de fácil entendimiento por parte del usuario lo que permite mayor comprensión para el análisis de los datos que se encuentran integrados.
- Es una metodología resistente y adaptable ante los cambios.

Conclusiones del capítulo

Después de desarrollar este capítulo se concluye que la tecnología a utilizar son los MD, las herramientas a utilizar son Visual Paradigm v6.4, Power Architect v3.0, PostgreSQL v9.0, la Suite de Pentaho v3.6 y la metodología seleccionada es la metodología de Kimball.

Capítulo 2: Análisis y diseño del mercado de datos estadísticas de servicios para el MINCEX.

Introducción

El siguiente capítulo está enmarcado en exponer las características del sistema propuesto para que cumpla con las expectativas del cliente y usuarios finales.

2.1 Procesos del negocio

La propuesta de solución está orientada a los procesos identificados en el análisis de las estadísticas de servicios, enfocándolo principalmente en el diseño dimensional del mercado de datos.

Exportación: se dice que tiene lugar una exportación de servicios cuando los residentes venden, obsequian o donan servicios a no residentes.

Importación: se dice que tiene lugar una importación de servicios cuando los residentes compran o aceptan obsequios o donaciones de servicios de no residentes.

Intercambio comercial: es el término utilizado por los especialistas del MINCEX para analizar el intercambio de servicios y países con respecto a Cuba.

2.2 Reglas del negocio

Las reglas del negocio describen las políticas, normas y restricciones presentes en una organización además de ser la entrada fundamental para los procesos de diseño del almacén, ETL y BI a continuación se describen las reglas del negocio para la solución propuesta.

- País no puede ser Cuba en las tablas de hecho exportación, importación e intercambio comercial.
- Las importaciones y exportaciones por empresa/servicio/país deben crecer o mantenerse de un trimestre al siguiente.
- El país no debe repetirse por empresa/servicio aun en el caso del 999.
- Los valores serán en miles de pesos con un decimal.
- CPCU: La clasificación consiste en secciones (identificadas por el primer dígito), divisiones (identificadas por el primero y el segundo dígitos), campo (identificados por los tres primeros dígitos), clases (identificadas por los cuatro

primeros dígitos) y subclases (identificadas por los cinco dígitos en conjunto).

➤ REUP: los primeros tres dígitos son el código del organismo, después un dígito el tipo de empresa, luego 5 dígitos el código de la empresa.

2.4 Áreas de análisis

Las áreas de análisis responden a la agrupación de la información según su propósito, cuyo criterio depende de las necesidades del departamento de análisis estadístico, para facilitar el análisis de los indicadores del comercio exterior de servicios.

1. Análisis de los indicadores específicos de exportación en el comercio exterior de servicios.
2. Análisis de los indicadores específicos de importación en el comercio exterior de servicios.
3. Análisis de los indicadores específicos de intercambio comercial en el comercio exterior de servicios
4. Análisis de los indicadores específicos de donaciones exportadas en el comercio exterior de servicios.
5. Análisis de los indicadores específicos de donaciones importadas en el comercio exterior de servicios.
6. Análisis de los indicadores del comercio exterior de servicio por serie de tiempo.
7. Análisis de los indicadores específicos del comercio exterior de servicios en el boletín.
8. Análisis de los indicadores específicos del proceso de cuadro del comercio exterior de servicios.
9. Análisis de errores en los indicadores del comercio exterior de servicio.

2.5 Requisitos

2.5.1 Requisitos informativos

Los requisitos informativos representan los pedidos de información más usadas por el cliente, para la solución, fueron detectados 157 los cuales se encuentran distribuidos por áreas de análisis.

Análisis de los indicadores específicos del proceso de cuadro del comercio exterior de servicios: 5 requisitos informativos

Análisis de los indicadores específicos de exportación en el comercio exterior de servicios: 26 requisitos informativos

Análisis de los indicadores específicos de importación en el comercio exterior de servicios: 26 requisitos informativos

Análisis de los indicadores específicos de intercambio comercial en el comercio exterior de servicios: 5 requisitos informativos

Análisis de los indicadores específicos de donaciones exportadas en el comercio exterior de servicios: 26 requisitos informativos

Análisis de los indicadores específicos de donaciones importadas en el comercio exterior de servicios: 26 requisitos informativos

Análisis de los indicadores del comercio exterior de servicio por serie de tiempo (mismos trimestres de años diferentes, período de años en los 4 trimestres, trimestre de un mismo año): 33 requisitos informativos

Análisis de los indicadores específicos del comercio exterior de servicios en el boletín: 10 requisitos informativos

2.5.2 Requisitos funcionales

Los requisitos funcionales son capacidades o condiciones que el sistema debe cumplir para satisfacer las expectativas del cliente. A continuación se enumeran los requisitos funcionales identificados para el mercado de datos.

Autenticar usuario

Crear usuario

Modificar usuario

Visualizar usuario

Eliminar usuario

Crear rol

Modificar rol

Visualizar rol

Eliminar rol

Crear vista de análisis

Modificar vista de análisis

Visualizar vista de análisis

Eliminar vista de análisis

Crear reporte ad-hoc

Modificar reporte ad-hoc

Visualizar reporte ad-hoc

Cargar dimensiones de estadísticas

Cargar dimensiones de servicios

Cargar hechos de servicios

2.5.3 Requisitos no funcionales

Los requerimientos no funcionales son propiedades o cualidades que el producto debe tener, a continuación se muestran los requisitos no funcionales para la solución.

Requisitos de usabilidad:

1. Dominar conocimientos básicos sobre Base de datos.
2. Tener conocimiento en análisis y tratamiento de la información.
3. Tener al menos una persona que domine el trabajo con la herramienta “Pentaho Data Integration” para realizar o modificar las transformaciones para la integración de datos.
4. Los usuarios deben recibir capacitación de las herramientas utilizadas en la confección de las tablas de salida y cómo interpretar la información de los mismos.

Requisitos de disponibilidad:

1. El sistema debe estar disponible 100% entre las 8:00 am y las 5:00 pm de lunes a viernes.
2. La carga de los datos se realizará después de las 7:00 pm.

3. El sistema de integración será accedido para su mantenimiento una vez por mes. En este plazo de mantenimiento se validarán las estructuras de auditoría y se establecerán estrategias que permitan clasificar posibles errores y darles solución en caso de ser posible.

4. El área de almacenamiento del proceso de ETL debe estar disponible las 12 horas del día lo que representa 50% horas de uso.

Requisitos de interfaz:

1. Los reportes estadísticos deben contar con una interfaz simple que facilite la interacción usuario-aplicación.

2. Las interfaces de salida no serán cargadas con información innecesaria.

3. Los gráficos serán con colores claros ajustándose a los estándares establecidos de un buen diseño.

Requisitos de hardware:

1. En el proceso de integración es necesaria la utilización de una memoria mínima de 2GB para el proceso de transformación.

2. Se debe contar con un área de almacenamiento intermedio de 500GB mínimos.

3. Para la visualización y la inteligencia de negocio se necesita una memoria de 8GB.

4. Las estaciones de trabajo (pc clientes) deben contar con impresoras (para garantizar la impresión de las tablas de salida).

Requisitos de software:

1. Se debe disponer de la instalación de la herramienta BI SERVER.

2. Debe existir un navegador asociado al sistema operativo que se escoja para lograr que las interfaces web de las tablas de salida puedan visualizarse.

3. El lenguaje para la programación dentro del repositorio será MDX para realizar consultas a la base de datos e implementar las funciones necesarias.

4. El perfilado de datos se realizará con el JavaScript, Pentaho, herramienta libre especializada en estas funcionalidades.

2.6 Casos de uso del sistema

En la fase de requerimiento se modelan los casos de uso que ejercen un papel protagónico, pues se agrupan todos los requisitos funcionales y de información. Responsables de especificar la comunicación y el comportamiento del sistema y los usuarios. En la siguiente figura se muestra el Diagrama de Casos de Uso del Sistema (DCUS) para la solución.

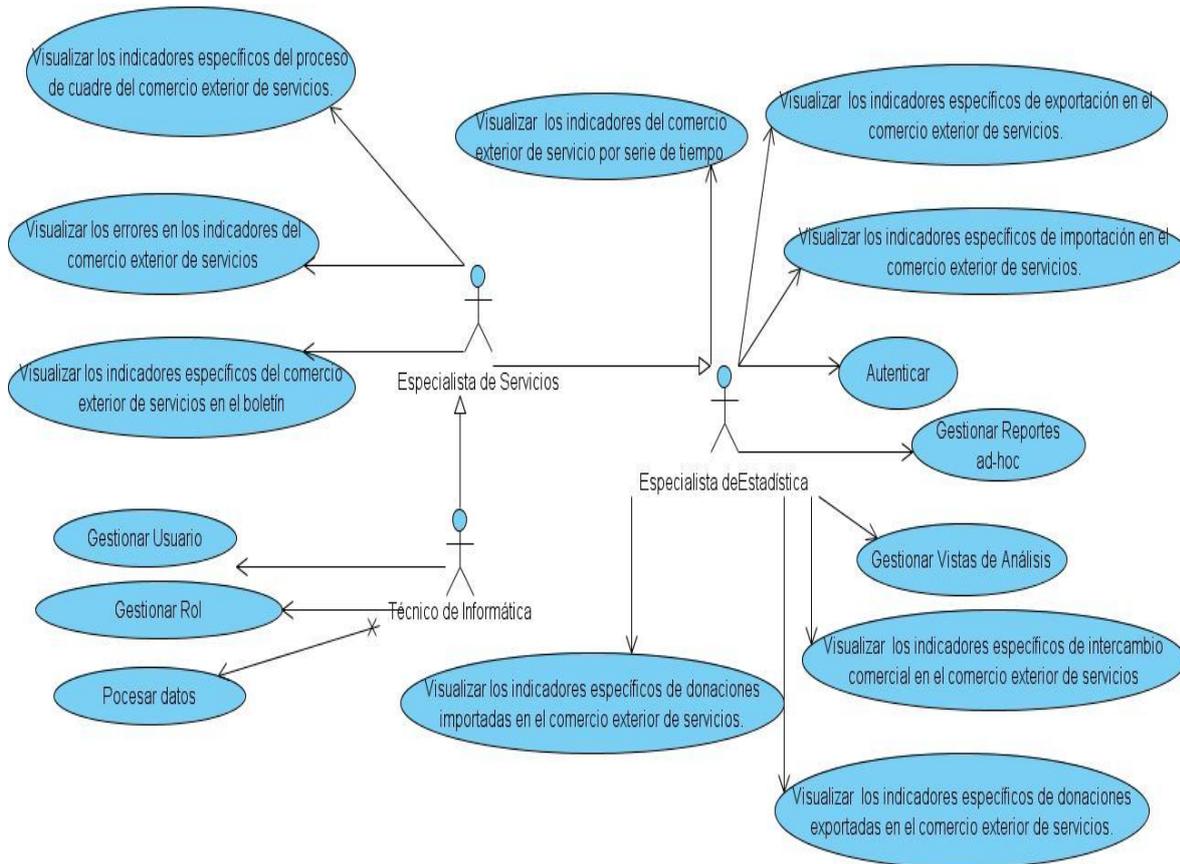


Figura 3. Diagrama de Caso de Uso del Sistema (DCUS)

2.7 Diseño de la solución

El diseño corresponde a la representación de las estructuras dimensionales identificadas en el análisis, donde posteriormente se van a almacenar las estadísticas de servicios.

2.7.1 Modelo de datos dimensional

El modelo de datos dimensional está integrado por las dimensiones, hechos y, variables cuantitativas llamadas medidas. En la solución propuesta se escogió el modelo tipo estrella para su desarrollo. Debido a esto, se hace necesario conocer

cada uno de sus componentes, los cuales se explican a continuación con el propósito de facilitar su entendimiento.

2.7.2 Especificación de las dimensiones

Se identificaron 5 dimensiones que tienen relación de uno a muchos con las tablas de hecho.

- **dim_pais:** agrupa información relacionada con los países, especificando las zonas, regiones, subregiones, la región policome y el país de interés.
- **dim_empresa:** agrupa información relacionada con las empresas exportadoras e importadoras que emiten información al MINCEX. Está compuesta por: el nombre, el tipo de la empresa, el organismo y el REUP.
- **dim_CPCU:** agrupa información relacionada con el Clasificador de Productos de Cuba. Está compuesta por sección, división, campo, clase y subclase.
- **dim_dpa:** agrupa información relacionada con la División Política Administrativa de Cuba. Está compuesta por municipio y provincia.
- **dim_tiempo:** agrupa información relacionada con la periodicidad con que se recoge la información del modelo 0793. Está compuesta por: año, trimestre y mes.

2.7.3 Especificación de los hechos

El diseño está compuesto por 3 tablas de hecho; Intercambio, Importación y Exportación. Por cada tabla de hecho se confeccionan dos vistas materializadas las cuales van a almacenar los datos de las estadísticas de servicio para el año actual y los últimos dos años.

- **hech_importacion:** agrupa los indicadores recogidos en el modelo 0793, relacionados con la importación de servicios y contiene las medidas cuanta por pagar al inicio del año, servicios facturados, pagos realizados, cuentas pagar al final del año, donaciones importadas y total de importaciones.



Figura 4. Tabla del hecho importación

➤ **hech_intercambio:** agrupa los indicadores recogidos en el modelo 0793, relacionados con el intercambio de servicios y contiene las medidas total de exportaciones, total de importaciones, balanza comercial e intercambio comercial.



Figura 5. Tabla del hecho intercambio

➤ **hech_exportacion:** agrupa los indicadores recogidos en el modelo 0793, relacionados con la exportación de servicios y contiene las medidas cuentas por cobrar al inicio del año, servicios facturados, pagos realizados, cuentas por cobrar al final del año, donaciones exportadas y total de exportaciones.



Figura 6. Tabla del hecho exportación

2.7.4 Seguridad del mercado de datos

La seguridad en cualquier sistema es de vital importancia por ello toda aplicación debe proteger y mantener la integridad de los datos que maneja. Por tal motivo el esquema de seguridad está respaldado por la asignación de roles y permisos garantizando que el usuario autenticado pueda acceder a la información que le corresponde.

Rol	Permisos
Especialista en Servicios	Analiza los indicadores del comercio exterior servicios, además de realizar el proceso de cuadro, boletín y analizar los errores que pueda tener la información.
Especialista de estadísticas	Analiza los indicadores del comercio exterior de servicios.
Técnico de Informática	Puede realizar las acciones de los especialistas del departamento de Análisis Estadístico, y además se encarga de gestionar usuarios y roles, además de realizar el procesamiento de los datos.

Tabla 1. Roles y permisos

2.7.5 Estrategia de recuperación y respaldo

Para garantizar la estabilidad y consistencia en los datos se realizarán salvases mensuales, trimestrales y anuales.

2.8 Arquitectura del sistema

La arquitectura de software define la relación entre los componentes identificados. Para la solución propuesta se utilizó una arquitectura ascendente orientada al desarrollo de mercados de datos, guiados por la metodología de Kimball.

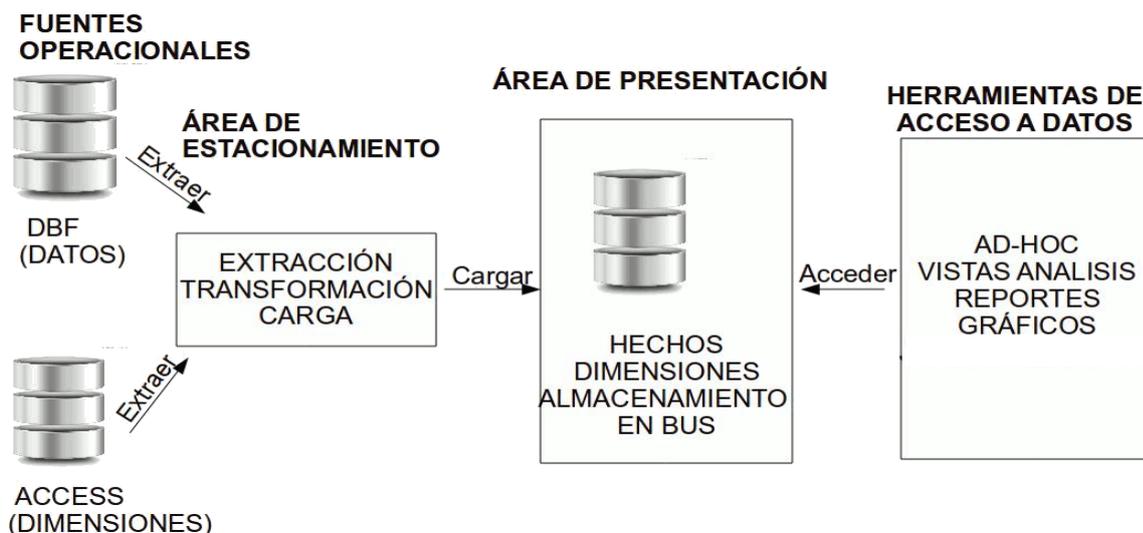


Figura 7. Arquitectura del Sistema

Fuentes operacionales

Las fuentes de datos operacionales utilizadas para el mercado de datos son las siguientes:

Archivos M793 en formato dbf que contiene los indicadores de los procesos de importación y exportación.

Base de datos access con los nomencladores necesarios para el análisis de los indicadores.

Área de estacionamiento

Representación simbólica de los procesos de ETL. Está constituida por los archivos de la herramienta de integración de datos y todas las tablas de almacenamiento temporal necesarias para ejecutar las transformaciones, para el diseño de la solución se realizaron 16 transformaciones.

Área de presentación

Son todas las tablas del mercado de datos, en las cuales se almacena de manera permanente las cifras de servicios. Las tablas se almacenan en esquemas separados en dependencia de sus funciones y posible compatibilidad con otros mercados del mismo almacén.

Herramientas de acceso a datos

Subsistemas de visualización. Contiene todos los archivos de las estructuras

predefinidas para el análisis de las cifras de servicios, ya sean reportes, gráficos o vistas de análisis.

2.9 Arquitectura de información

La arquitectura de información describe cómo estará distribuida la información y las normas de diseño del sistema. Seguidamente se muestra el mapa del sistema que define un esquema de navegación a seguir por el usuario.



Figura 8. Mapa de Navegación

Descripción de los elementos de la arquitectura

Nivel principal: se enmarca en las Estadísticas de Servicios, la cual controla el comercio exterior de servicios del país con el resto del mundo.

Área de análisis: responde a la agrupación de la información según su propósito cuyo criterio depende de las necesidades del departamento de estadísticas donde se aplica el sistema.

Libro de trabajo: estructura organizativa que agrupa los reportes generados dentro de las áreas de análisis. Se crea teniendo en cuenta criterios de organización de la información.

Reporte: unidad básica utilizada para mostrar la información almacenada con un formato establecido por el departamento de estadística y el mismo será impreso para la consulta de los directivos.

Vista de análisis: unidad básica utilizada para mostrar la información almacenada. Permite agrupar y detallar la información para los indicadores escogidos.

Gráfico: permite visualizar indicadores a través de colores y tamaños.

2.10 Vista de despliegue

La vista de despliegue muestra el hardware del sistema, el software instalado en él y los protocolos de comunicación utilizados para conectar unos nodos con otros.

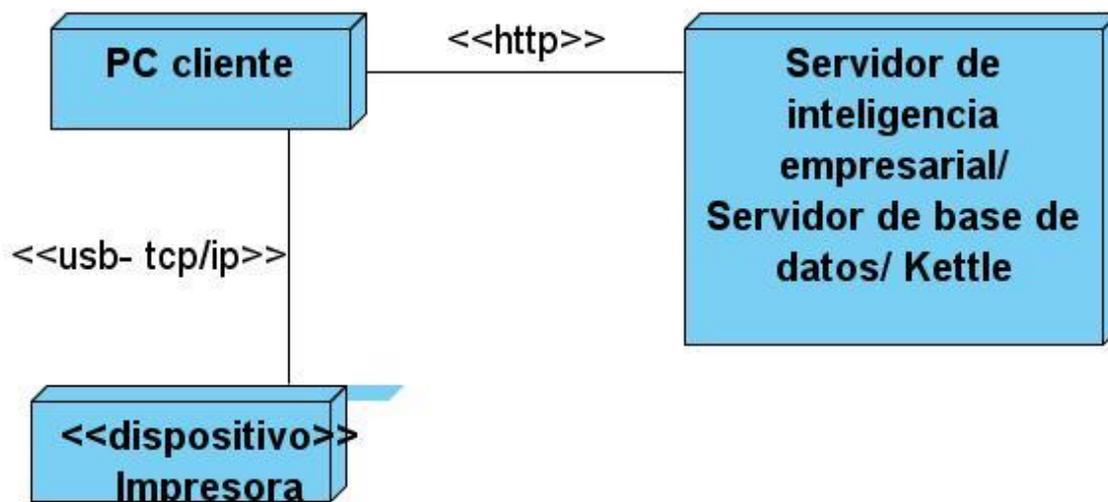


Figura 9. Diagrama de despliegue

Nodo PC cliente

Existen 2 computadoras en el departamento de estadísticas, las cuales presentan 512 MB de RAM, y tienen como sistema operativo Windows.

Nodo servidor

El servidor es un Dual Core con 2GB de RAM y 500GB de disco duro. Trabaja con el sistema operativo Ubuntu Server 10.4 y en él se almacena el AD para el departamento de estadísticas.

Nodo impresora

Existen 4 impresoras, una de ellas es HP LaserJet 1320, otra es LaserJet 2015 y las dos restantes son Epson LX-300. Las impresoras están conectadas a las computadoras mediante conexión usb y además existe una de ellas que está en red.

Conclusiones del capítulo

En este capítulo se realizó un estudio detallado del negocio y se comenzó a desarrollar la propuesta de solución del sistema. Se identificaron los requisitos funcionales, no funcionales y de información, además de las áreas de análisis, reglas del negocio que se deberán tener en cuenta, las necesidades de usuarios, los casos de uso del sistema y el modelo de datos dimensional.

Capítulo 3: Implementación del mercado de datos estadísticas de servicios para el MINCEX.

Introducción

En el presente capítulo se abordará todo lo relacionado con la implementación de los subsistemas de integración y visualización correspondientes a la solución propuesta.

3.1 Estándar de código

El estándar de código es un término que describe convenciones para escribir código fuente en ciertos lenguajes de programación. Con él se logra dar uniformidad al código para que sea entendible por otros programadores y facilitar la revisión del código por los especialistas del departamento de Calidad. A continuación se muestra el estándar de código utilizado para la implementación del mercado de datos.

➤ Base de Datos

Almacén de Datos.....[nombre]_dwh

Almacén de Datos Operacional.....[nombre]_ods

Catálogo Personalizado.....[nombre]_meta

Base de datos de Intercambio.....[nombre]_middle

➤ Esquema

Objetos de Configuración.....config_[nombre esquema]

Dimensiones Compartidas.....dimensiones_estadística

Esquemas de datos.....mart_[temática]

➤ Tablas

Configuración.....cf_[nombre tabla]

Dimensiones.....dim_[nombre]

Copo de Nieve.....cp_[nombre dimensión 1_nombre dimensión 2]

Hechos.....hech_[nombre]

Vistas Materializadas.....vm_[nombre]_#a

➤ Secuencias

Secuencias dimensiones.....seqDim_[nombre_campo]

Secuencias configuración.....seqConf_[nombre_campo]

Secuencias hechos.....seqHech_[nombre_campo]

➤ Funciones

ETL.....fc_etl_[nombre tabla]

Sincronización.....fc_scn_[nombre tabla]

Disparador.....fc_tgr_[nombre tabla]

Particionamiento.....fc_prt_[nombre tabla]

Replicación.....fc_rp_[nombre tabla]

➤ Vistas

Vistas.....v_[nombre]

➤ Índices

Índices.....Idx_[nombre tabla]-[nombre campo]

➤ Restricciones

Claves Primarias.....pk_[tabla]_id

Clave Foránea.....fk_[tablaFuente]_id

Unicidad.....unq_[nombre]

➤ Disparador

Control de Cambio.....tgr_cdc_[nombre]

Sincronización.....tgr_snc_[nombre]

Réplica.....tgr_rp_[nombre]

➤ Columnas de Tabla

Dimensión.....[nombre dimensión]_id, [nombre de nivel]_codigo, [nombre de nivel]_nombre, [nombre de nivel]_descripcion, [nombre de nivel]_[nombre atributo]

Hechos.....[nombre dimensión]_id, *

➤ Proceso de ETL

Entrada de tabla.....ed_et_[nombre de la tabla]

Entrada de excel.....ed_ee_[nombre de la hoja de cálculo]

Entrada de Access.....ed_ea_[nombre de la tabla]

Selecciona campos y renombrar campos....sr #

Insertar y Actualizar Datos.....sd_ia_[nombre de la tabla]

Salida tabla.....sd_st_[nombre de la tabla]

Salida a Access.....sd_sa_[nombre de la tabla]

Salida a Excel....sd_se_[nombre del fichero]

Buscar datos....bd_[nombre de la tabla dónde vas a buscar]

Código java....js #

Partir campo.....pc_[campo que se partirá]-> [nombre del primer campo]/ [nombre del segundo campo]

Insertar columna.....ic->[nombre de la primera columna]/ [nombre de la segunda columna]....

Calculadora....cl->[nombre de la primer columna obtenida]/[nombre de la segunda columna obtenida]

Fila única.....fu_[nombre de la primera columna]/[nombre de la segunda columna]

Ordenar fila.....of_[nombre de la primer columna] /[nombre de la segunda columna]

Filtrar fila.....ff #

Agrupar por.....ap #

Eliminar.....e_[nombre de la columna]

3.2 Proceso de extracción, transformación y carga

ETL, este término viene de inglés de las siglas Extract-Transform-Load que significan Extraer, Transformar y Cargar y se refiere a los datos en una empresa. ETL es el proceso que organiza el flujo de los datos entre diferentes sistemas en una organización y aporta los métodos y herramientas necesarias para mover datos desde múltiples fuentes a un almacén de datos, reformatearlos, limpiarlos y

cargarlos en otra base de datos, data mart ó bodega de datos.

Extracción

La extracción convierte los datos a un formato preparado para iniciar el proceso de transformación. En esta fase se leen y comprenden los datos de la fuente y se copia la parte que interesa en el área de almacenamiento.

Transformación

En la transformación se aplican una serie de reglas de negocio o funciones sobre los datos extraídos para convertirlos en datos que serán cargados. En esta fase se incluye la limpieza de datos, se eliminan los campos que no son necesarios y se puede realizar la combinación de varias fuentes.

Carga

La fase de carga es el momento en el cual los datos de la fase anterior (transformación) son cargados en el sistema de destino.

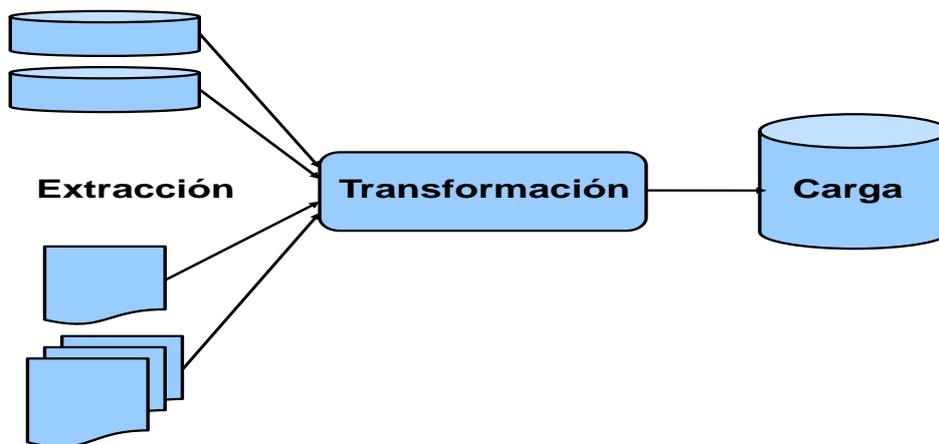


Figura 10. Proceso de ETL

3.2.1 Estrategia

El proceso de extracción de los datos se realiza a partir de una fuente conformada por dos ficheros access denominados servicios.mdb y baseAduana.mdb que contienen las tablas con los datos brutos para posteriormente transformarlos. Una vez efectuada la transformación, se realizan las validaciones necesarias teniendo en cuenta las reglas del negocio identificadas y en caso de encontrarse valores incorrectos, el flujo se desvía hacia un fichero excel donde son almacenados con la correspondiente descripción del error. Si se comprueba que los datos poseen la

calidad requerida, se les realiza un conjunto de transformaciones y se procede a su inserción en el mercado de datos.

Transformaciones para la dimensiones

Para realizar la transformación de las dimensiones se cargan los datos que se encuentran almacenados en las tablas de configuración ubicadas en la base de datos denominada mincex_dwh. A continuación se extraen los datos de las tablas anteriormente mencionadas, se transforman teniendo en cuenta las reglas del negocio y finalmente se cargan ya transformados en las tablas de dimensiones ubicadas en la misma base de datos. Para realizar toda la extracción, transformación y carga de los datos se tuvieron en cuenta algunos componentes que conforman a este proceso, los cuales se mencionan a continuación con una breve referencia de que funcionalidad cumple cada uno:



Selecciona/Renombrar: Se utiliza para obtener los campos luego de realizar alguna transformación y para renombrar los campos buscando que exista coincidencia con los nombres de los campos pertenecientes a la tabla destino donde se va a realizar la carga final.



Búsqueda en base de datos: Se utiliza para verificar coincidencia entre los datos de la tabla de entrada y una tabla especificada que tengan relación entre sí para obtener nuevos datos.



Filtrar Filas: Se utiliza para validar datos y teniendo en cuenta el resultado de estas validaciones envía los datos hacia fuentes de origen diferentes.

Transformaciones para los hechos

Para realizar la transformación de los hechos, se extrajeron los datos del fichero access denominado servicios.mdb, se realizaron búsquedas en cada una de las tablas de dimensiones relacionadas con los hechos y teniendo en cuenta las reglas del negocio se cargaron los datos en cada una de las tablas de hechos ubicadas en la base de datos denominada mincex_dwh. Para desarrollar este proceso se utilizaron los componentes empleados en las transformaciones de las dimensiones, también se utilizaron otros los cuales se mencionan a continuación:



Ordenar filas: Se utiliza para ordenar filas especificando los campos.



Filas únicas: Se utiliza para evitar que se repitan datos.



Valor java script: Se utiliza para validar las transformaciones teniendo en cuenta las reglas del negocio.



Calculadora: Se utiliza para realizar cálculos definiendo el nombre del campo que va a ser calculado.



Eliminar: Se utiliza para eliminar los datos redundantes especificando el campo que se va a tomar como referencia.

Trabajos

Un trabajo o job es similar al concepto de proceso. Un proceso es un conjunto sencillo o complejo de tareas con el objetivo de realizar una acción determinada. En los trabajos se pueden utilizar pasos específicos que son diferentes a los disponibles en las transformaciones. Además, se pueden ejecutar una o varias transformaciones ya diseñadas y orquestar una secuencia de ejecución de ellas. Los trabajos estarían en un nivel superior a las transformaciones.

Para realizar el trabajo con las dimensiones, solo se tuvo en cuenta cargar estas haciendo referencia a las transformaciones ya creadas anteriormente. De la misma manera se trabajó con los hechos. Para realizar estas acciones se utilizó el componente que se muestra a continuación:



Transformación: Se utiliza para cargar las transformaciones.

3.3 Vistas materializadas

Una vista materializada almacena físicamente los datos resultantes de ejecutar la consulta definida en la vista. Este tipo de vistas materializadas realizan una carga inicial de los datos cuando se definen y posteriormente con una frecuencia establecida se actualizan los datos de la misma. Con la utilización de vistas materializadas se logra aumentar el rendimiento de las consultas SQL además de ser un método de optimización a nivel físico en modelos de datos muy complejos

o con muchos datos.

Como parte de la solución se definieron las funciones dentro de las Funciones Disparadoras ubicadas en el esquema mart_servicios que son las que se programan y en Disparadores ubicados dentro de cada una de las tablas de hechos y vistas materializadas se guardan los disparadores o triggers que son los que automatizan a cada una de las funciones, así cuando se realice algún tipo de evento como insertar, actualizar o eliminar los disparadores serán activados y seguidamente se ejecutará la función que esté relacionada con esos disparadores.

3.4 Implementación del subsistema de visualización de datos

Cubos OLAP

Para crear el sistema se han definido cada uno de los cubos que responden a todas las tablas de hechos y vistas materializadas existentes, las cuales a su vez abarcan a todas las dimensiones ya descritas en otros capítulos. Estos cubos fueron nombrados de acuerdo a los procesos de negocio con los que están relacionados. Los mismos fueron desarrollados mediante el uso de la herramienta Schema Workbench y el tipo de almacenamiento definido para estos cubos es el almacenamiento ROLAP. Todo este proceso es guardado en un archivo .xml quedando de la siguiente manera:



Figura 11. Cubos involucrados en el proceso

Para crear los cubos inicialmente se definieron estos decidiendo cuáles eran los atributos necesarios para realizar el análisis. Posteriormente se especificó la estructura de las dimensiones definiendo las jerarquías con sus respectivos niveles y las medidas utilizadas.

Vistas de análisis

Para obtener una vista generalizada de todos los datos y poder realizar el análisis de los mismos se utilizaron las vistas de análisis permitiéndole al usuario interactuar con la información de forma detallada o resumida además de brindar la opción de exportar a excel, pdf y mostrar gráficos.

		Medidas	
Organismo	Region	● Total de Exportaciones	● CCIA
[-] Empresa	[+] Pais	9950,00	2611,60
[+] MICONS	[+] Pais	2,00	0,20
[+] MINCEX	[+] Pais	0,00	0,00
[+] MINIL	[+] Pais	7,00	0,00
[+] SIME	[-] Pais	9941,00	2611,40
	[+] Africa	73,00	16,80
	[+] América	4995,00	1721,90
	[+] Asia	1031,00	164,80
	[+] Europa	3840,00	705,40
	[+] Medio Oriente	2,00	2,50

Figura 12. Vista de análisis

Reportes

En la implementación de la solución se generaron reportes útiles para la Dirección de Economía y Planificación del MINCEX a partir de los datos registrados que van a estar orientados a darle respuesta a los pedidos de información identificados.

Los reportes ad_hoc facilitan la creación de reportes específicos que no se encuentran definidos en la solución propuesta, esto es posible realizando los metadatos, brindándole al usuario la información de manera entendible. A continuación se muestra la figura de un reporte definido en la solución propuesta y un reporte ad_hoc:



Exportación y Porcentaje de Servicios por Servicio

Total Exportado: 9.950

Servicio	MPS	Por ciento
Otros servicios auxiliares	6	6%
Otros servicios de ..	9	9%
Otros servicios ..	113	113,6%
Servicios auxiliares y otro ..	46	46,2%
Servicios científicos y otro..	553	555,8%
Servicios comerciales al ..	6	6%
Servicios comerciales al ..	11	11,1%

Fecha actual: 13 abril 2012

Cifras preliminares emitidas para el MINCEX

Fuente: Oficina Nacional de Estadísticas e Información

CONFECCIONADOS _____ HOJAS _____

DISTRIBUCION DE ENVIO: _____

Figura 13. Reporte definido en la solución propuesta

Empresa	País	Total exportaciones
ELEKA	REINO DE LOS PAISES BAJOS	0
ELEKA	REPUBLICA DE POLONIA	364
ELEKA	REPUBLICA DEL LIBANO	1
ELEKA	RUMANIA	1
ELEKA	REPUBLICA DE NICARAGUA	18
ELEKA	REUBLICA DEL PERU	67
ELEKA	LIBIA	1
ELEKA	REPUBLICA FEDERACION RUSA *	0
ELEKA	REINO DE NORUEGA	2
ELEKA	REUBLICA DEL PERU	3
ELEKA	REPUBLICA DE HAITI	3
ELEKA	REPUBLICA ISLAMICA DE IRAN	0

Figura 14. Reporte ad_hoc

Gráficos

En la solución propuesta fue utilizado el gráfico de barras teniendo en cuenta que este tipo de gráfico permite de forma muy precisa el reconocimiento visual de patrones y tendencias, sustentados en la capacidad del ser humano de detectar cambios en la longitud de una forma, por lo que también resulta una excelente herramienta para visualizar comparaciones entre las magnitudes de los datos que se representan, facilitando al observador el proceso de búsqueda de similitudes o contrastes.

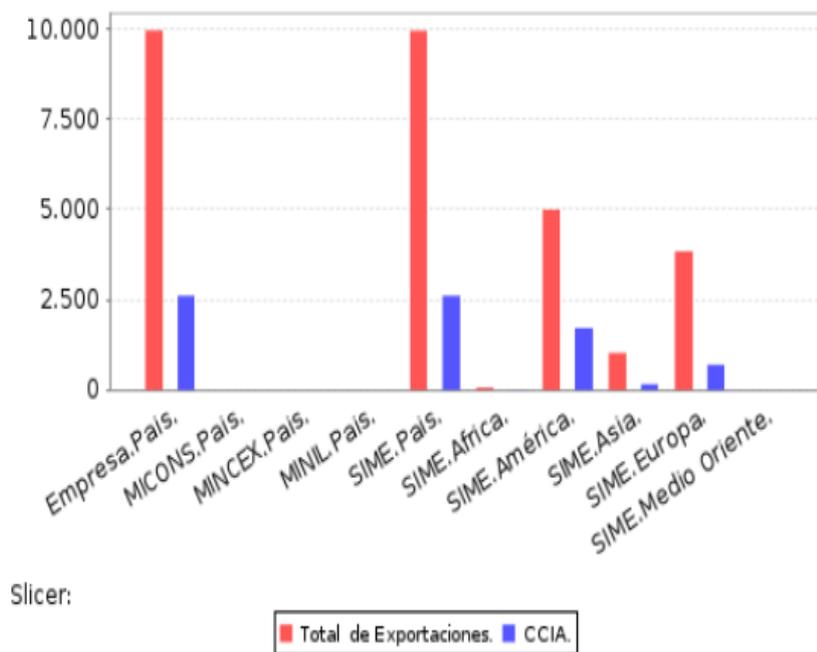


Figura 15. Gráfico de barra

3.4.1 Administración del sistema

Para administrar el sistema se utiliza la consola de administración que permite realizar las configuraciones necesarias para garantizar la seguridad y la accesibilidad de la información, permitiendo desplegar la aplicación satisfactoriamente, de esta forma si se desea emigrar la información hacia otra base de datos se evitan conflictos dentro de la aplicación.

Conexiones JNDI

Toda la información mostrada proviene de la base de datos y esto es posible mediante las estructuras de visualización que obtienen su funcionamiento a través

de la conexión JNDI definida en la consola de administración. Esta conexión caracterizada por permitir a los clientes descubrir y buscar objetos y datos a través de un nombre, ofrece la posibilidad de mantener funcionando eficientemente las estructuras de visualización si la base de datos es cambiada de lugar.

Usuarios y roles

Para garantizar la seguridad de la aplicación se definen dentro de la consola de administración los usuarios y roles a los cuales se les asignan permisos de acuerdo a las funciones que cumpla cada uno, creando de esta manera un sistema totalmente seguro y preparado para responder ante cualquier infracción o violación del sistema haciéndolo más eficiente y menos vulnerable.

Conclusiones del capítulo

Durante el desarrollo de este capítulo se abordó de manera general todo lo relacionado con la implementación del MD Estadísticas de Servicios. Siguiendo el estándar de código definido por el equipo de trabajo fueron definidos los procesos de carga de los datos y a su vez la automatización de los mismos mediante ETL y las vistas materializadas. Se creó el cubo de datos que posteriormente fue utilizado en las vistas de análisis y se generaron los reportes necesarios para la dirección de Economía y Planificación del MINCEX.

Capítulo 4: Validación del mercado de datos estadísticas de servicios para el MINCEX.

Introducción

En el capítulo se exponen las pruebas realizadas al mercado de datos, así como los resultados obtenidos en cada una de las pruebas, verificando que se cumplan las necesidades del cliente y el producto cuente con la calidad requerida.

4.1 Pruebas de software

Las pruebas son un conjunto de actividades en las cuales un sistema o componente es ejecutado bajo unas condiciones o requerimientos específicos, donde los resultados son observados y registrados para dar una evaluación del sistema. Por esta razón, se hace necesario aplicar diferentes pruebas luego de culminada la etapa de implementación, con el propósito de verificar no solo que el producto cumpla con los requisitos del cliente, sino también para eliminar los posibles defectos que pudiera tener.

4.1.1 Diseño de los casos de prueba

Para la validación del mercado de datos, se diseñaron 15 casos de prueba, con el propósito de verificar los requisitos funcionales, agrupados en 15 casos de uso del sistema que fueron definidos durante la etapa de análisis. A continuación se muestran las pruebas realizadas a los casos de uso.

Visualizar los indicadores del boletín para las cifras de servicios	
Escenario: Visualizar reportes.	
Subescenario	Visualizar reportes de exportación.
Descripción	Muestra los reportes del boletín para el proceso de exportación.
Perfiles de análisis	Tiempo, cpcu, país, empresa, dpa
Indicadores	a Valor exportado

medir	
Flujo central	Se autentifica y selecciona la carpeta de exportación en el lateral izquierdo superior de la aplicación, siguiendo el siguiente esquema de navegación: servicios/reportes/exportación. Se selecciona el reporte deseado en el lateral izquierdo inferior.
Respuesta del sistema	Sistema muestra el reporte seleccionado en el lateral derecho de la aplicación. Cada reporte refleja el comportamiento de los indicadores a medir en los últimos dos años, en dependencia de los perfiles de análisis definidos.
Resultados obtenidos	Satisfactorios.
Subescenario	Visualizar reportes de importación.
Descripción	Muestra los reportes del boletín para el proceso de importación.
Perfiles de análisis	Tiempo, cpcu, país, empresa, dpa
Indicadores a medir	Valor importado
Flujo central	Se autentifica y selecciona la carpeta de importación en el lateral izquierdo superior de la aplicación, siguiendo el siguiente esquema de navegación: servicios/reportes/importación. Se selecciona el reporte deseado en el lateral izquierdo inferior.
Respuesta del sistema	El sistema muestra el reporte seleccionado en el lateral derecho de la aplicación. Cada reporte refleja el comportamiento de los indicadores a medir en los últimos dos años, en dependencia de los perfiles de análisis definidos.

Resultados obtenidos	Satisfactorios.
Subescenario	Visualizar reportes de intercambio comercial.
Descripción	Muestra los reportes del boletín para el proceso de intercambio comercial.
Perfiles de análisis	Tiempo, cpcu, país, empresa, dpa
Indicadores a medir	Valor intercambio
Flujo central	Se autentifica y selecciona la carpeta de importación en el lateral izquierdo superior de la aplicación, siguiendo el siguiente esquema de navegación: servicios/reportes/importación. Se selecciona el reporte deseado en el lateral izquierdo inferior.
Respuesta del sistema	El sistema muestra el reporte seleccionado en el lateral derecho de la aplicación. Cada reporte refleja el comportamiento de los indicadores a medir en los últimos dos años, en dependencia de los perfiles de análisis definidos.
Resultados obtenidos	Satisfactorios
Escenario: Visualizar vistas de análisis.	
Subescenario	Visualizar vistas de análisis de exportación.
Descripción	Muestra las vistas del boletín para el proceso de exportación.
Perfiles de análisis	Tiempo, cpcu, país, empresa, dpa
Indicadores a medir	Valor exportado

Flujo central	Se autentifica y selecciona la carpeta de exportación en el lateral izquierdo superior de la aplicación, siguiendo el siguiente esquema de navegación: servicios/vistas de análisis/exportación. Se selecciona la vista deseada en el lateral izquierdo inferior.
Respuesta del sistema	El sistema muestra la vista seleccionada en el lateral derecho de la aplicación. Cada vista refleja el comportamiento de los indicadores a medir en los últimos dos años, en dependencia de los perfiles de análisis definidos.
Resultados obtenidos	Satisfactorios.
Subescenario	Visualizar vistas de análisis de importación.
Descripción	Muestra las vistas del boletín para el proceso de importación.
Perfiles de análisis	Tiempo, cpcu, país, empresa, dpa
Indicadores a medir	Valor importado
Flujo central	Se autentifica y selecciona la carpeta de importación en el lateral izquierdo superior de la aplicación, siguiendo el siguiente esquema de navegación: servicios/vistas de análisis/importación. Se selecciona la vista deseada en el lateral izquierdo inferior.
Respuesta del sistema	El sistema muestra la vista seleccionada en el lateral derecho de la aplicación. Cada vista refleja el comportamiento de los indicadores a medir en los últimos dos años, en dependencia de los perfiles de análisis definidos.
Resultados	Satisfactorios.

obtenidos	
Subescenario	Visualizar vistas de análisis de intercambio comercial.
Descripción	Muestra las vistas del boletín para el proceso de intercambio comercial.
Perfiles de análisis	Tiempo, cpcu, país, empresa, dpa
Indicadores a medir	Valor intercambio
Flujo central	Se autentifica y selecciona la carpeta de intercambio en el lateral izquierdo superior de la aplicación, siguiendo el siguiente esquema de navegación: servicios/vistas de análisis/intercambio. Se selecciona la vista deseada en el lateral izquierdo inferior.
Respuesta del sistema	El sistema muestra la vista seleccionada en el lateral derecho de la aplicación. Cada vista refleja el comportamiento de los indicadores a medir en los últimos dos años, en dependencia de los perfiles de análisis definidos.
Resultados obtenidos	Satisfactorios.

Tabla 2. Caso de prueba “Visualizar los indicadores del boletín para las cifras de servicios”

Visualizar los indicadores del proceso de importación de servicios.	
Escenario: Visualizar reportes.	
Descripción	Muestra los reportes del proceso de importación de servicios.
Perfiles de análisis	Tiempo, cpcu, país, empresa, dpa
Indicadores a	Valor importado

medir	
Flujo central	Se autentifica y selecciona la carpeta de reportes en el lateral izquierdo superior de la aplicación, siguiendo el siguiente esquema de navegación: importación/reportes. Se selecciona el reporte deseado en el lateral izquierdo inferior.
Respuesta del sistema	El sistema muestra el reporte seleccionado en el lateral derecho de la aplicación. Cada reporte refleja el comportamiento de los indicadores a medir en el último año, en dependencia de los perfiles de análisis definidos.
Resultados obtenidos	Satisfactorios.

Tabla 3. Caso de prueba “Visualizar los indicadores del proceso de importación de servicios”

Visualizar los indicadores del proceso de exportación de servicios.	
Escenario: Visualizar reportes.	
Descripción	Muestra los reportes del proceso de exportación de servicios.
Perfiles de análisis	Tiempo, cpcu, país, empresa, dpa
Indicadores a medir	Valor exportado
Flujo central	Se autentifica y selecciona la carpeta de reportes en el lateral izquierdo superior de la aplicación, siguiendo el siguiente esquema de navegación: exportación/reportes. Se selecciona el reporte deseado en el lateral izquierdo inferior.
Respuesta del sistema	El sistema muestra el reporte seleccionado en el lateral derecho de la aplicación. Cada reporte refleja el comportamiento de los indicadores a medir en el último año, en dependencia de los perfiles de análisis definidos.

Resultados obtenidos	Satisfactorios.
-----------------------------	-----------------

Tabla 4. Caso de prueba “Visualizar los indicadores del proceso de exportación de servicios”

Visualizar los indicadores del proceso de intercambio comercial de servicios.	
Escenario: Visualizar reportes.	
Descripción	Muestra los reportes del proceso de intercambio comercial de servicios
Perfiles de análisis	Tiempo, cpcu, país, empresa, dpa
Indicadores a medir	Valor del intercambio
Flujo central	Se autentifica y selecciona la carpeta de reportes en el lateral izquierdo superior de la aplicación, siguiendo el siguiente esquema de navegación: intercambio comercial/reportes. Se selecciona el reporte deseado en el lateral izquierdo inferior.
Respuesta del sistema	El sistema muestra el reporte seleccionado en el lateral derecho de la aplicación. Cada reporte refleja el comportamiento de los indicadores a medir en el último año, en dependencia de los perfiles de análisis definidos.
Resultados obtenidos	Satisfactorios.

Tabla 5. Caso de prueba “Visualizar los indicadores del intercambio comercial de servicios”

Visualizar los indicadores del proceso de cuadro.	
Escenario: Visualizar reportes.	
Subescenario	Visualizar vistas de análisis de importación
Descripción	Muestra las vistas de cuadro de cifras del proceso importación.

Perfiles de análisis	de	Tiempo, cpcu, país, empresa, dpa
Indicadores a medir	a	Cantidad de empresas, valor importado
Flujo central		Se autentifica y selecciona la carpeta de importación en el lateral izquierdo superior de la aplicación, siguiendo el siguiente esquema de navegación: proceso de cuadro/vistas de análisis/importación. Se selecciona la vista deseada en el lateral izquierdo inferior.
Respuesta del sistema	del	El sistema muestra la vista seleccionada en el lateral derecho de la aplicación. Cada vista refleja el comportamiento de los indicadores a medir en el último año, en dependencia de los perfiles de análisis definidos.
Resultados obtenidos		Satisfactorios.
Subescenario		Visualizar vistas de análisis de exportación
Descripción		Muestra las vistas de cuadro de cifras del proceso exportación.
Perfiles de análisis	de	Tiempo, cpcu, país, empresa, dpa
Indicadores a medir	a	Cantidad de empresas, valor exportado
Flujo central		Se autentifica y selecciona la carpeta de exportación en el lateral izquierdo superior de la aplicación, siguiendo el siguiente esquema de navegación: proceso de cuadro/vistas de análisis/exportación. Se selecciona la vista deseada en el lateral izquierdo inferior.
Respuesta del sistema	del	El sistema muestra la vista seleccionada en el lateral derecho de la aplicación. Cada vista refleja el comportamiento de los indicadores a medir en el último

	año, en dependencia de los perfiles de análisis definidos.
Resultados obtenidos	Satisfactorios.

Tabla 6. Caso de prueba “Visualizar los indicadores del proceso de cuadro”

Visualizar los indicadores de donaciones exportadas.	
Escenario: Visualizar reportes.	
Descripción	Muestra los reportes del proceso de exportación de servicios.
Perfiles de análisis	Tiempo, cpcu, país, empresa, dpa
Indicadores a medir	Donaciones exportadas
Flujo central	Se autentifica y selecciona la carpeta de reportes en el lateral izquierdo superior de la aplicación, siguiendo el siguiente esquema de navegación: donaciones exportadas/reportes. Se selecciona el reporte deseado en el lateral izquierdo inferior.
Respuesta del sistema	El sistema muestra el reporte seleccionado en el lateral derecho de la aplicación. Cada reporte refleja el comportamiento de los indicadores a medir en el último año, en dependencia de los perfiles de análisis definidos.
Resultados obtenidos	Satisfactorios.

Tabla 7. Caso de prueba “Visualizar los indicadores de donaciones exportadas”

Visualizar los indicadores de donaciones importadas.	
Escenario: Visualizar reportes.	
Descripción	Muestra los reportes del proceso de importación de servicios.
Perfiles de	Tiempo, cpcu, país, empresa, dpa

análisis	
Indicadores a medir	Donaciones importadas
Flujo central	Se autentifica y selecciona la carpeta de reportes en el lateral izquierdo superior de la aplicación, siguiendo el siguiente esquema de navegación: donaciones exportadas/reportes. Se selecciona el reporte deseado en el lateral izquierdo inferior.
Respuesta del sistema	El sistema muestra el reporte seleccionado en el lateral derecho de la aplicación. Cada reporte refleja el comportamiento de los indicadores a medir en el último año, en dependencia de los perfiles de análisis definidos.
Resultados obtenidos	Satisfactorios.

Tabla 8. Caso de prueba “Visualizar los indicadores de donaciones importadas”

Visualizar errores en los indicadores de las cifras de bienes.	
Escenario: Visualizar vistas de análisis.	
Subescenario	Visualizar vistas de análisis de importación
Descripción	Muestra las vistas de análisis que pudieran reflejar posibles errores en las cifras de servicios.
Perfiles de análisis	Tiempo, cpcu, país, empresa, dpa
Indicadores a medir	Valor importado
Flujo central	Se autentifica y selecciona la carpeta de importación en el lateral izquierdo superior de la aplicación, siguiendo el siguiente esquema de navegación: proceso de cuadro/vistas de análisis/importación. Se selecciona la vista deseada en el lateral izquierdo inferior.

Respuesta del sistema	El sistema muestra la vista seleccionada en el lateral derecho de la aplicación. Cada vista refleja el comportamiento de los indicadores a medir en el último año, en dependencia de los perfiles de análisis definidos.
Resultados obtenidos	Satisfactorios.
Subescenario	Visualizar vistas de análisis de exportación
Descripción	Muestra las vistas de análisis que pudieran reflejar posibles errores en las cifras de servicios.
Perfiles de análisis	Tiempo, cpcu, país, empresa, dpa
Indicadores a medir	Valor exportado
Flujo central	Se autentifica y selecciona la carpeta de exportación en el lateral izquierdo superior de la aplicación, siguiendo el siguiente esquema de navegación: proceso de análisis de errores/vistas de análisis/exportación. Se selecciona la vista deseada en el lateral izquierdo inferior.
Respuesta del sistema	El sistema muestra la vista seleccionada en el lateral derecho de la aplicación. Cada vista refleja el comportamiento de los indicadores a medir en el último año, en dependencia de los perfiles de análisis definidos.
Resultados obtenidos	Satisfactorios.

Tabla 9. Caso de prueba “Visualizar errores en los indicadores de las cifras de servicios”

Visualizar los indicadores de serie de tiempo	
Escenario: Visualizar reportes.	
Subescenario	Visualizar reportes de exportación
Descripción	Muestra los reportes del proceso de exportación en series

		de tiempo.
Perfiles de análisis	de	Tiempo, cpcu, país, empresa, dpa
Indicadores a medir	a	Valor exportado
Flujo central		Se autentifica y selecciona la carpeta de exportación en el lateral izquierdo superior de la aplicación, siguiendo el siguiente esquema de navegación: series de tiempo/reportes/importación. Se selecciona el reporte deseado en el lateral izquierdo inferior.
Respuesta del sistema	del	El sistema muestra el reporte seleccionado en el lateral derecho de la aplicación. Cada reporte refleja el comportamiento de los indicadores a medir en el último año, en dependencia de los perfiles de análisis definidos.
Resultados obtenidos		Satisfactorios.
Subescenario		Visualizar reportes de importación
Descripción		Muestra los reportes del proceso de exportación en series de tiempo.
Perfiles de análisis	de	Tiempo, cpcu, país, empresa, dpa
Indicadores a medir	a	Valor importado
Flujo central		Se autentifica y selecciona la carpeta de importación en el lateral izquierdo superior de la aplicación, siguiendo el siguiente esquema de navegación: series de tiempo/reportes/importación. Se selecciona el reporte deseado en el lateral izquierdo inferior.
Respuesta del sistema	del	El sistema muestra el reporte seleccionado en el lateral derecho de la aplicación. Cada reporte refleja el

	comportamiento de los indicadores a medir en el último año, en dependencia de los perfiles de análisis definidos.
Resultados obtenidos	Satisfactorios.
Subescenario	Visualizar reportes de importación
Descripción	Muestra los reportes del proceso de exportación en series de tiempo.
Perfiles de análisis	Tiempo, cpcu, país, empresa, dpa
Indicadores a medir	Valor importado, valor exportado, valor intercambio, balanza comercial
Flujo central	Se autentifica y selecciona la carpeta de intercambio en el lateral izquierdo superior de la aplicación, siguiendo el siguiente esquema de navegación: series de tiempo/reportes/intercambio. Se selecciona el reporte deseado en el lateral izquierdo inferior.
Respuesta del sistema	El sistema muestra el reporte seleccionado en el lateral derecho de la aplicación. Cada reporte refleja el comportamiento de los indicadores a medir en el último año, en dependencia de los perfiles de análisis definidos.
Resultados obtenidos	Satisfactorios.
Escenario: Visualizar gráficos.	
Subescenario	Visualizar gráficos de exportación
Descripción	Muestra los gráficos del proceso de exportación en series de tiempo.
Perfiles de análisis	Tiempo, cpcu, país, empresa, dpa
Indicadores a	Valor exportado

medir	
Flujo central	Se autentifica y selecciona la carpeta de exportación en el lateral izquierdo superior de la aplicación, siguiendo el siguiente esquema de navegación: series de tiempo/gráficos/importación. Se selecciona el gráfico deseado en el lateral izquierdo inferior.
Respuesta del sistema	El sistema muestra el gráfico seleccionado en el lateral derecho de la aplicación. Cada gráfico refleja el comportamiento de los indicadores a medir en el último año, en dependencia de los perfiles de análisis definidos.
Resultados obtenidos	Satisfactorios.
Subescenario	Visualizar gráficos de importación
Descripción	Muestra los gráficos del proceso de exportación en series de tiempo.
Perfiles de análisis	Tiempo, cpcu, país, empresa, dpa
Indicadores a medir	Valor importado
Flujo central	Se autentifica y selecciona la carpeta de importación en el lateral izquierdo superior de la aplicación, siguiendo el siguiente esquema de navegación: series de tiempo/gráficos/importación. Se selecciona el gráfico deseado en el lateral izquierdo inferior.
Respuesta del sistema	El sistema muestra el gráfico seleccionado en el lateral derecho de la aplicación. Cada gráfico refleja el comportamiento de los indicadores a medir en el último año, en dependencia de los perfiles de análisis definidos.
Resultados obtenidos	Satisfactorios.

Subescenario	Visualizar gráficos de importación
Descripción	Muestra los gráficos del proceso de exportación en series de tiempo.
Perfiles de análisis	Tiempo, cpcu, país, empresa, dpa
Indicadores a medir	Valor importado, valor exportado, valor intercambio, balanza comercial
Flujo central	Se autentifica y selecciona la carpeta de intercambio en el lateral izquierdo superior de la aplicación, siguiendo el siguiente esquema de navegación: series de tiempo/gráficos/intercambio. Se selecciona el gráfico deseado en el lateral izquierdo inferior.
Respuesta del sistema	El sistema muestra el gráfico seleccionado en el lateral derecho de la aplicación. Cada gráfico refleja el comportamiento de los indicadores a medir en el último año, en dependencia de los perfiles de análisis definidos.
Resultados obtenidos	Satisfactorios.

Tabla 10. Caso de prueba “Visualizar los indicadores de serie de tiempo”

4.2 Resultados de las pruebas

Seguidamente se detallarán de forma breve cada una de las pruebas aplicadas al mercado de datos, así como los resultados obtenidos en cada una de ellas:

Prueba de las estructuras de pre-visualización o estructuras multidimensionales

En esta prueba se verifica que las estructuras construidas (cubos importación, exportación, intercambio, importación 1 año, exportación 1 año, intercambio 1 año, importación 2 año, exportación 2 año, intercambio 2 año) respondan a todo el alcance de información que fue definido en el negocio. Teniendo como precondition estructuras sólidas que respondan a las necesidades de información que se definieron en el negocio. Se mide con los siguientes aspectos:

- Eficiente: Las estructuras cumplen con el alcance de información previsto.

- Parcial: Las estructuras cumplen en un 80% con el alcance de información previsto.
- Deficiente: Las estructuras cumplen menos del 80% con el alcance de información previsto.

El resultado obtenido fue: Eficiente.

4.3 Resultados y funcionalidades obtenidas

El desarrollo del mercado de datos se realizó en cuatro fases. Durante el transcurso de cada fase se realizaron un conjunto de actividades; tales como capacitación e investigación y estudio de temas relacionados con el desarrollo de almacenes de datos y elaboración de documentos. Al finalizar cada fase arrojó un resultado práctico o una nueva funcionalidad al sistema de la manera que se explica a continuación:

Estudio preliminar: Fin del proceso de levantamiento de la información. Se obtuvo un cuestionario resuelto con las características del sistema a implementar.

Fase de requerimientos: Se establecieron cada uno de los requerimientos funcionales, no funcionales e informativos. Se diseñaron y describieron los casos de uso del sistema. Se realizó un estudio profundo de las fuentes de datos, factible para el proceso de ETL aún no implementado.

Fase de arquitectura y diseño: Se diseñó la base de datos que corresponde al modelo dimensional del mercado de datos. Se estableció la arquitectura de software y la arquitectura de información. Se obtuvo un sistema capaz de almacenar las estadísticas de servicios.

Fase de implementación y prueba: Se obtuvo un sistema funcional capaz de extraer y transformar las estadísticas de servicios y almacenarlas en la base de datos. En dicho sistema se pueden analizar las estadísticas de servicios con el uso de estructuras predefinidas como reportes, gráficos y vistas de análisis. Además, fue sometido a un proceso de validación a través de caso de pruebas.

Conclusiones del capítulo

En el capítulo, se detallaron los diferentes casos de prueba utilizadas para la validación del mercado de datos y se abordaron las diferentes pruebas realizadas para verificar el cumplimiento de los objetivos definidos, conjuntamente con los resultados obtenidos luego de ser aplicadas.

Conclusiones generales

Finalmente se concluye que:

- El análisis del modelo 0793 correspondiente a las estadísticas de servicios proporcionó que se obtuvieran los indicadores que luego fueron utilizados para el diseño del MD.
- El estudio de las principales tendencias de implementación de los MD permitió escoger las herramientas factibles para el diseño e implementación del sistema.
- El proceso de ETL se implementó satisfactoriamente al insertar en el MD estadísticas de servicios los datos requeridos para los posteriores análisis de información.
- La implementación de los subsistemas de visualización definidos responde a las necesidades de información y permite una eficiente presentación de la información requerida para la toma de decisiones.
- La validación del MD mediante las pruebas resultó satisfactoria haciendo posible que la propuesta cuente con la calidad requerida.

Recomendaciones

Los objetivos generales de este trabajo fueron alcanzados, pero durante su desarrollo han surgido ideas que serían recomendables para su futuro perfeccionamiento:

- Perfeccionar la construcción del MD de servicios e incorporarlo al AD del MINCEX.
- Que se logre una unificación de las estadísticas de bienes y servicios para que se pueda trabajar con más información y de forma más organizada.
- Que se tome el contenido de esta investigación como referencia para próximos sistemas de mercados y almacenes de datos que se desarrollen.

Bibliografía referenciada

- **Burgaleta Altuna, Boris Luis.** Data Mart de facturación. Data warehouse Cubacel, 2007. [Disponible en: <http://www.uh.cu/infogral/areasuh/vri/archivos/Otros/Compu/compu02.htm>]
- **Hernández Peñaloza, Lucía Victoria.** Diseño y Construcción de un Data Mart para la mantención de Indicadores de Sostenibilidad de la Industria del Salmón. Chile, 2008.
- **Herrera, C.** Adictos al trabajo, 2007. [Disponible en: <http://www.adictosaltrabajo.com/tutoriales/tutoriales.php?pagina=datawarehouse4>]
- **Imhoff, Claudia, Galemme, Nicholas y Geiger, Jonathan G.** Mastering Data Warehouse Desing, Relational and Dimentional Techniques. EUA, Wiley Publishing Inc, 2009.
- **Inmon, William H.** Building the Data Warehouse. EUA, Wiley Publishing Inc, 2005.
- **Inmon, William H., Strauss y Neushloss.** DW 2.0 The Architecture for de Next Generation. EUA, Wiley Publishing Inc, 2007.
- **Kimball, Ralph, y otros.** The Data Warehouse Lifecycle Toolkit. EUA, Wiley Publishing Inc, 2007.
- **Leyva Osorio, Reinaldo.** Mercado de Datos Estadístico de Inmigración y Extranjería para el Departamento de Turismo y Comercio de la Oficina Nacional de Estadísticas. UCI. Ciudad Habana, 2010. 78.
- **Ortiz Sierra, Julio Ernesto.** Diseño e Implementación de un Mercado de Datos para la Oficina Nacional de Estadísticas. UCI. Ciudad Habana, 2009. 132.
- **Ponniah, Paulraj.** Data Warehousing Fundamentals. EUA, Wiley Publishing Inc, 2009.
- **Rodríguez Monges, Marianela.** Mercado de datos Indicadores relacionados con el Comercio Agropecuario para la Oficina Nacional de Estadísticas. UCI. Ciudad Habana, 2010. 81.
- **Rodríguez Sans, Miguel.** ANÁLISIS Y DISEÑO DE UN DATAMART PARA EL SEGUIMIENTO ACADÉMICO DE ALUMNOS EN UN ENTORNO

UNIVERSITARIO, 2010. [Disponible en: <http://es.scribd.com/doc/52203545/6/Panorama-actual-de-los-Data-Warehouse-y-los-Data-Mart>]

➤ **Soto, Lauro.** Sistemas De Apoyo Toma De Decisiones, 2010. [Disponible en: <http://www.mitecnologico.com/Main/SistemasDeApoyoTomaDeDecisiones>

➤ **Technologies, B. P.** Visual Paradigm, 2010. [Disponible en: <http://www.visual-paradigm.com>]

➤ **Zepeda Sánchez, Leopoldo Zenaido.** Metodología para el Diseño Conceptual de Almacenes de Datos. UNIVERSIDAD POLITÉCNICA DE VALENCIA. Valencia, 2008. 211.

Bibliografía consultada

- **Almacenando Datos.** Modelo Dimensional, 2010. [Disponible en: <http://proyectopentahodw.wordpress.com/2010/05/04/modelo-dimencional>]
- **Álvarez, Jaime Velázquez.** Implementación del modelo de datos ArcHydro en la región hidrológica Balsas. México, 2010.
- **Antunez, Ivette Marrero.** La inteligencia de negocios desde la perspectiva cubana: retos y tendencias. La Habana, Cuba , 2008.
- **Armstrong, Smith.** Oracle Discoverer 10g Handbook. San Francisco, California, 2006.
- **Bernabeu, Ricardo Darío.** HEFESTO: Metodología propia para la Construcción de un Data Warehouse, 2008. [Disponible en:http://www.google.com/cu/url?sa=t&source=web&ct=res&cd=1&ved=0CAYQFjAA&url=http%3A%2F%2Fwww.dataprix.com%2Fprintpdf%2F288&rct=j&q=Metodo log%C3%ADa+SQLBI&ei=Nb_PS5L3NYGdlgfcroSiCw&usg=AFQjCNGD0prG8PrOfwBRU4r5Pwq4CtYiMg. 20]
- **Burgaleta Altuna, Boris Luis.** Data Mart de facturación. Data warehouse Cubacel, 2007. [Disponible en: <http://www.uh.cu/infogral/areasuh/vri/archivos/Otros/Compu/compu02.htm>]
- **Burgaleta Altuna, Boris Luis.** Data Mart de facturación. Data warehouse Cubacel, 2010. [Disponible en: <http://www.uh.cu/infogral/areasuh/vri/archivos/Otros/Compu/compu02.htm>]
- **Business Intelligence+Informática estratégica.** Sinnexus, 2007. [Disponible en: <http://www.sinnexus.com>]
- **Candás Romero, Jorge.** Bid, 2006. [Disponible en: <http://www.ub.edu/bid/17canda2.htm>]
- **Cañete, Patricio.** Business Weblog, 2006. [Disponible en: <http://www.pcanete.com.ar/leer.asp?idx=414>]
- **Company Headquarters.** Visual Paradigm, 2009. [Disponible en: <http://www.visual-paradigm.com>]