

# Universidad de las Ciencias Informáticas

Facultad Regional de Artemisa



## Trabajo de diploma para optar por el título de Ingeniero en Ciencias Informáticas

Mercado de datos estadísticas de bienes para el  
departamento de análisis estadístico del Ministerio del  
Comercio Exterior y la Inversión Extranjera.

Autores:

Dianelys Alonso Martínez  
Yamila Monteagudo Sánchez

Tutor:

Ing. Mario Redonavich Garbey

Co-Tutor:

Ing. Ygraine Borges Ramirez

# Declaración de Autoría

---

## 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 \_\_\_\_\_.

---

Yamila Monteagudo Sánchez  
Autor

---

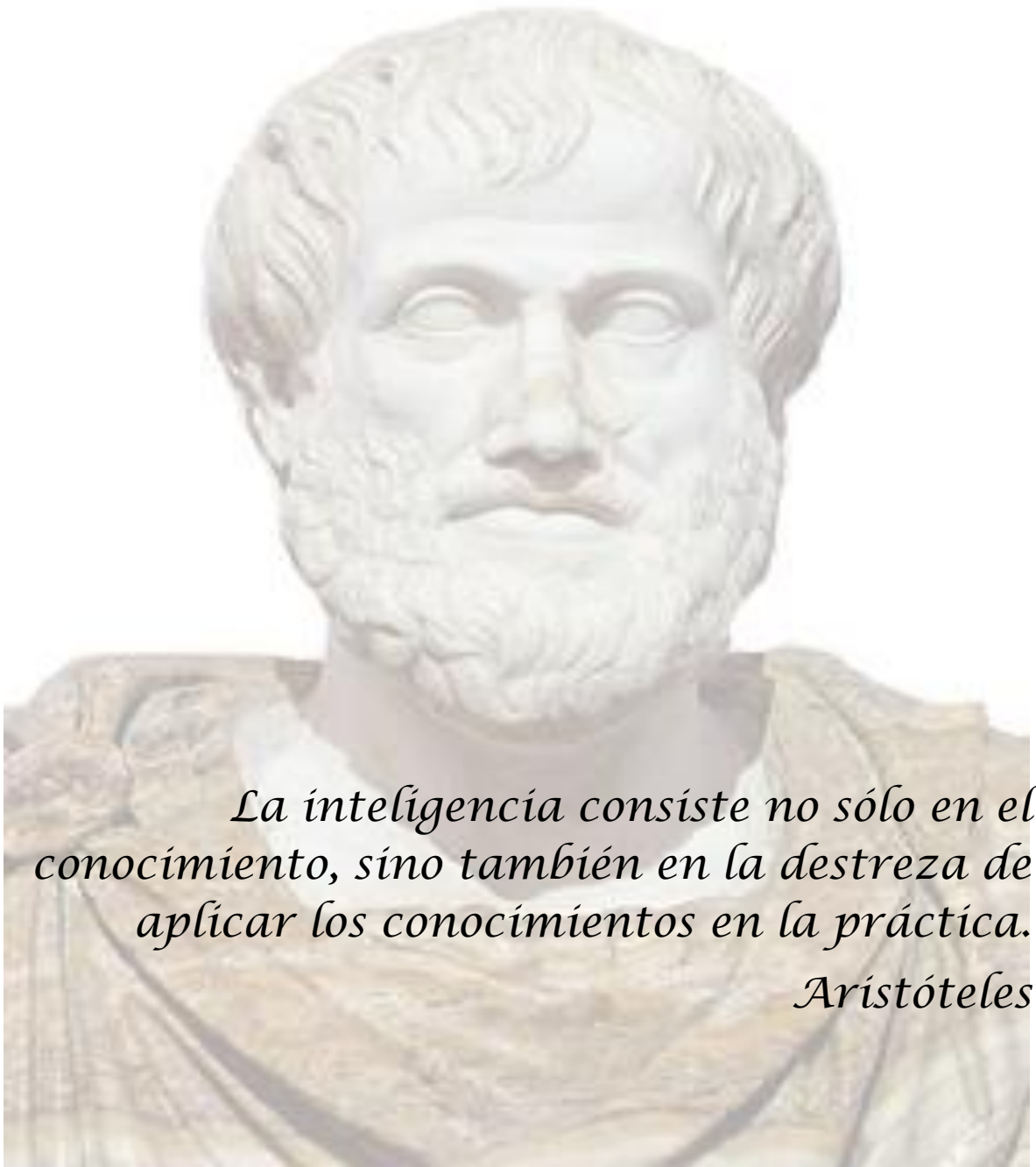
Dianelys Alonso Martínez  
Autor

---

Ing Mario Redonavich Garbey  
Tutor

---

Ing Ygraine Borges Ramirez  
Co - Tutor



*La inteligencia consiste no sólo en el conocimiento, sino también en la destreza de aplicar los conocimientos en la práctica.*

*Aristóteles*

## Agradecimientos

---

### *Agradecimientos*

*Yamila*

*Quisiera agradecerles a mis padres y a mi familia por apoyarme en este largo camino.*

*A mis amigos que han estado conmigo en los buenos y malos momentos.*

*A Estre, a Nati, Rufi, Douglas, Ivan y Yari por escucharme y ayudarme desinteresadamente.*

*A todo aquel que creyó en mí y que puso su grano de arena para ayudarme a construir mi sueño.*

*A mi tutor por enseñarme a pensar y por siempre estar ahí. Gracias a su dedicación lo logramos.*

*Dianelys*

*Le agradezco a todo aquel que de una forma u otra ayudó en el desarrollo de este trabajo.*

*A mis padres que siempre estuvieron presentes.*

*A Roberto Blanco por haberme apoyado durante toda la carrera.*

*A mi tío Jorge y a mi tía Haymelis por haberme apoyado.*

*Un agradecimiento especial a mi tutor Mario Redonavich por el apoyo incondicional que siempre nos brindó y por ser buena guía.*

*A mi compañera Amarelys Álvarez por su desempeño en el inicio de este trabajo y a mi compañera de tesis Yamila Monteagudo por su esfuerzo, apoyo y dedicación en nuestro trabajo de diploma.*

*A todos ellos mis más sinceros agradecimientos.*

## Dedicatoria

---

### *Dedicatoria*

*Yamila*

*Dedico esta tesis a mi padre que con su educación me llevó de la mano por el buen camino y una vez allí sigue a mi lado.*

*A mi madre que me ha apoyado en todo momento de mi vida y que siempre ha querido lo mejor para mí.*

*A mi hermana, mis sobrinos, mi familia.*

*A alguien que luchó mucho por esta carrera pero que no está para ver los frutos, gracias por todo.*

*A mis amigos que me han enseñado que todos somos seres humanos pero que en nuestras manos está hacer la diferencia.*

*Dianelys*

*Dedico este trabajo de diploma a mi madre Olga Lidia Martínez, para cumplir uno de sus sueños.*

*A mi padre Juan Alonso y a mi hermana Yanelys Erika para que sirva de ejemplo en el futuro.*

# Resumen

---

## Resumen

La cantidad de información relevante y la capacidad de análisis sobre la misma son factores que influyen en el desenvolvimiento de los negocios. Muchas empresas optan por el almacenamiento y análisis de la información estadística derivada de sus procesos de negocio como la clave del éxito o el fracaso en el proceso de toma de decisiones. En este contexto surgen los almacenes de datos, una tecnología que permite el almacenamiento de datos no volátiles a largo plazo y posee técnicas para convertirlos en información relevante.

Este trabajo se basa en la implementación de un mercado de datos para el análisis de las estadísticas de bienes materiales derivadas de los procesos de importación y exportación en Cuba. Tomando como referencia las necesidades existentes en el departamento de análisis estadístico del Ministerio del Comercio Exterior y la Inversión Extranjera (MINCEX).

El mercado de datos será el responsable de extraer y asegurar los datos estadísticos desde diversas fuentes. Responderá a las necesidades de análisis de la información a diferentes niveles, con un sistema de visualización con estructuras predefinidas, vistas de análisis, reportes y gráficos.

Todas las herramientas utilizadas pertenecen a la familia del software libre. La implementación se realizó con la versión 3.6 comunitaria de la suite de Pentaho. Este trabajo representa una primera iteración que seguirá desarrollándose en el centro de desarrollo de la facultad regional Mártires de Artemisa (FRA). Está incluido en la construcción del almacén de datos para las estadísticas del MINCEX.

## Índice de contenidos

<b>INTRODUCCIÓN .....</b>	<b>1</b>
<b>CAPÍTULO 1. FUNDAMENTACIÓN TEÓRICA .....</b>	<b>9</b>
INTRODUCCIÓN .....	9
1.1 CONCEPTOS ASOCIADOS AL DOMINIO DEL PROBLEMA .....	9
1.2 ALMACENES DE DATOS .....	9
1.2.1 <i>¿Qué es un Almacén de datos?</i> .....	11
1.2.2 <i>Minería de datos</i> .....	14
1.2.3 <i>Mercados de datos</i> .....	14
1.2.4 <i>Modelo Dimensional</i> .....	15
1.2.5 <i>Sistemas de procesamiento analítico en línea</i> .....	16
1.2.6 <i>Esquemas Dimensionales</i> .....	17
1.3 METODOLOGÍA DE DESARROLLO DE SOFTWARE .....	18
1.4 HERRAMIENTAS A UTILIZAR .....	19
1.4.1 <i>Sistema gestor de bases de datos</i> .....	19
1.4.2 <i>Herramienta manejadora de base de datos</i> .....	20
1.4.3 <i>Herramienta de Modelado</i> .....	21
1.4.4 <i>Suites de inteligencia de negocio</i> .....	23
1.5 DOCUMENTACIÓN .....	25
CONCLUSIONES .....	25
<b>CAPÍTULO 2. ANÁLISIS Y DISEÑO DE LA SOLUCIÓN PROPUESTA.....</b>	<b>27</b>
INTRODUCCIÓN .....	27
2.1 DEFINICIÓN DEL NEGOCIO .....	27
2.1.1 <i>Procesos del negocio</i> .....	28
2.2 ÁREAS DE ANÁLISIS .....	28
2.3 REQUERIMIENTOS DEL SISTEMA.....	29

# Índice

---

2.3.1	<i>Requisitos de Información</i> .....	29
2.3.2	<i>Requisitos funcionales</i> .....	30
2.3.3	<i>Requisitos no funcionales</i> .....	31
2.4	CASOS DE USO DEL SISTEMA .....	32
2.5	DISEÑO DE LA SOLUCIÓN.....	33
2.6	MODELO DE DATOS DIMENSIONAL .....	33
2.7	SEGURIDAD DE LOS DATOS .....	37
2.8	ARQUITECTURA DEL SISTEMA .....	38
2.9	ARQUITECTURA DE LA INFORMACIÓN .....	39
2.9.1	<i>Vista de despliegue</i> .....	41
	CONCLUSIONES .....	42
	<b>CAPÍTULO 3. “IMPLEMENTACIÓN DE LA SOLUCIÓN PROPUESTA” .....</b>	<b>43</b>
	INTRODUCCIÓN .....	43
3.1	ESTÁNDAR DE CÓDIGO .....	43
3.2	SUBSISTEMA DE INTEGRACIÓN.....	45
3.2.1	<i>Estrategia de ETL</i> .....	46
3.2.2	<i>Integridad de los datos</i> .....	48
3.3	VISTAS MATERIALIZADAS.....	49
3.4	SUBSISTEMAS DE VISUALIZACIÓN .....	50
3.4.1	<i>Vistas de análisis</i> .....	50
3.4.2	<i>Reportes ad-hoc</i> .....	51
3.4.3	<i>Reportes</i> .....	53
3.4.5	ADMINISTRACIÓN DEL SISTEMA.....	55
	CONCLUSIONES .....	56
	<b>CAPÍTULO 4. “VALIDACIÓN DE LA SOLUCIÓN PROPUESTA” .....</b>	<b>57</b>
	INTRODUCCIÓN .....	57
4.1	PRUEBAS DE SOFTWARE .....	57



# Índice

---

4.2	PRUEBAS FUNCIONALES BASADAS EN CASOS DE USO.....	57
4.3	PRUEBAS DE INTEGRACIÓN. ....	71
4.4	PRUEBA DE LAS ESTRUCTURAS MULTIDIMENSIONALES Y METADATOS. ....	71
4.5	RESULTADOS Y FUNCIONALIDADES OBTENIDAS.....	71
	CONCLUSIONES .....	72
	<b>CONCLUSIONES GENERALES.....</b>	<b>73</b>
	<b>RECOMENDACIONES.....</b>	<b>74</b>
	<b>GLOSARIO DE TÉRMINOS.....</b>	<b>75</b>
	<b>BIBLIOGRAFÍA.....</b>	<b>76</b>

## Índice de ilustraciones

<i>FIGURA 1</i>	<i>ÁREAS DE ANÁLISIS.....</i>	<i>29</i>
<i>FIGURA 2</i>	<i>DIAGRAMA DE CASOS DE USO DEL SISTEMA.....</i>	<i>33</i>
<i>FIGURA 3</i>	<i>HECHO INTERCAMBIO COMERCIAL Y DIMENSIONES ASOCIADAS .....</i>	<i>36</i>
<i>FIGURA 4</i>	<i>HECHO IMPORTACIÓN Y DIMENSIONES ASOCIADAS. ....</i>	<i>36</i>
<i>FIGURA 5</i>	<i>HECHO EXPORTACIÓN Y DIMENSIONES ASOCIADAS.....</i>	<i>37</i>
<i>FIGURA 6</i>	<i>ARQUITECTURA DEL SISTEMA.....</i>	<i>38</i>
<i>FIGURA 7</i>	<i>MAPA DE NAVEGACIÓN.....</i>	<i>40</i>
<i>FIGURA 8</i>	<i>DIAGRAMA DE DESPLIEGUE .....</i>	<i>41</i>
<i>FIGURA 9</i>	<i>REPRESENTACIÓN DEL PROCESO DE ETL .....</i>	<i>46</i>
<i>FIGURA 10</i>	<i>REPRESENTACIÓN DEL PROCESO DE LAS VISTAS MATERIALIZADAS .....</i>	<i>49</i>
<i>FIGURA 11</i>	<i>VISTA DE ANÁLISIS. ....</i>	<i>50</i>
<i>FIGURA 12</i>	<i>CUBOS OLAP.....</i>	<i>51</i>
<i>FIGURA 13</i>	<i>REPORTE AD-HOC.....</i>	<i>52</i>
<i>FIGURA 14</i>	<i>METADATOS. ....</i>	<i>53</i>
<i>FIGURA 15</i>	<i>REPORTE. ....</i>	<i>54</i>
<i>FIGURA 16</i>	<i>GRÁFICO.....</i>	<i>55</i>

## Índice de tablas

# Índice

---

TABLA 1 PLANTILLA PARA CASOS DE PRUEBA. -----	58
TABLA 2 CASO DE PRUEBA “VISUALIZAR LOS INDICADORES DEL BOLETÍN”.-----	62
TABLA 3 CASO DE PRUEBA “VISUALIZAR LOS INDICADORES DE IMPORTACIÓN”.-----	63
TABLA 4 CASO DE PRUEBA “VISUALIZAR LOS INDICADORES DE EXPORTACIÓN”.-----	64
TABLA 5 CASO DE PRUEBA “VISUALIZAR LOS INDICADORES DE INTERCAMBIO COMERCIAL”.-----	64
TABLA 6 CASO DE PRUEBA “VISUALIZAR LOS INDICADORES DEL CUADRE DE CIFRAS”.-----	65
TABLA 7 CASO DE PRUEBA “VISUALIZAR LOS INDICADORES DE TRANSPORTE”.-----	67
TABLA 8 CASO DE PRUEBA “VISUALIZAR LOS INDICADORES DE PROVEEDORES Y CLIENTES”.-----	68
TABLA 9 CASO DE PRUEBA “VISUALIZAR ERRORES EN LAS CIFRAS DE BIENES”.-----	69
TABLA 10 CASO DE PRUEBA “VISUALIZAR LOS INDICADORES DE SERIES DE TIEMPO”.-----	71

# Introducción

---

## Introducción

La constante evolución de la sociedad actual demanda el uso de herramientas y soluciones rápidas y efectivas para los problemas que se generan diariamente. Las tecnologías de la información y las comunicaciones (TIC) están presentes en casi todas las esferas de la humanidad, respondiendo a las necesidades de las personas, instituciones, organizaciones, grandes empresas, entre otros. El uso en aumento de estas tecnologías unido al nivel de desarrollo que ha alcanzado la sociedad ha provocado que las personas necesiten cada vez más obtener e intercambiar información desde cualquier lugar y en cualquier momento.

La información juega un papel importante en el desarrollo de la sociedad moderna. Posee valor social, cultural, político, estratégico, económico, alimenta el quehacer diario de las personas y organizaciones. La cantidad de información relevante y la capacidad de análisis que se pueda tener sobre la misma son factores que influyen en el desenvolvimiento de los negocios.

Muchas empresas optan por el almacenamiento y análisis de la información estadística derivada de sus procesos de negocio como la clave del éxito o el fracaso en el proceso de toma de decisiones. Basados en esta idea y con el desarrollo de las TIC surgen los sistemas de apoyo a la toma de decisiones. Estos sistemas constituyen una tecnología que facilita el análisis de la información requerida para el proceso de la toma de decisiones. Su principal función consiste en proporcionar la mayor cantidad de información relevante en el menor tiempo posible, con el fin de decidir lo más adecuado.

Uno de los sistemas de apoyo a la toma de decisiones, son los llamados almacenes de datos, una tecnología que permite el almacenamiento de datos no

## Introducción

---

volátiles a largo plazo y posee técnicas para convertirlos en información importante. Los almacenes de datos han sido beneficiados con el desarrollo tecnológico mundial. Hoy en día permiten el análisis de la información a distintos niveles, incluso con el uso de gráficos y la creación de modelos predictivos.

Teniendo en cuenta las facilidades que propician el uso de las TIC, Cuba aboga por lograr la informatización de todos los sectores del país, así como del desarrollo de soluciones informáticas encaminadas a resolver diversos problemas. Con el objetivo de darle cumplimiento a esta filosofía, uno de los proyectos más grandes que se han creado en el país es la Universidad de las Ciencias Informáticas (UCI). La UCI cuenta con varios centros vinculados directamente a la producción de software bajo la premisa del software libre y la utilización del mismo en el desarrollo de las aplicaciones en los procesos de informatización.

La FRA, perteneciente a la UCI, ha formado parte del desarrollo de varios proyectos relacionados con el tema de los almacenes de datos, y actualmente pretende brindarle estos beneficios al MINCEX en Cuba.

El MINCEX tiene como función principal; preparar y proponer la política integral del estado y del gobierno, en las actividades de comercio exterior, inversión extranjera y colaboración económica. En el departamento de análisis estadístico del MINCEX se analizan las cifras derivadas de los procesos de importación y exportación de bienes materiales con el objetivo de conocer el estado del comercio exterior del país con el resto del mundo y además proporcionar un medio que apoye el proceso de toma de decisiones en busca de estrategias que permitan el crecimiento de la economía. Dicho departamento cuenta con un sistema de análisis estadístico que depende totalmente del uso de herramientas de escritorio y del factor humano como decisivo en su funcionamiento. Del análisis de dicho sistema

## Introducción

---

se deriva la siguiente **situación problemática**:

La validación de los datos de bienes materiales se realiza manual y al costo de un gran esfuerzo por parte de los especialistas, los cuales necesitan en ocasiones toda una jornada laboral para esto. Los datos manipulados presentan inconsistencia y no son fiables debido a que existen errores que no son subsanados en el proceso de validación. El almacenamiento no es confiable ya que dicho sistema permite alterar la información con facilidad.

El proceso de presentación de los datos, constatado como la clave del análisis estadístico, necesita de varias estructuras para realizar un análisis de lo detallado a lo resumido y viceversa, además, los datos presentados necesitan ser manipulados por especialistas para ser presentados nuevamente. Finalmente, la información solo puede ser accedida desde un solo ordenador con acceso restringido para el resto del personal del departamento.

Todo lo planteado anteriormente dificulta el procesamiento y el análisis de los datos estadísticos de bienes materiales en el departamento de análisis estadístico del MINCEX. Esta problemática conlleva al siguiente **problema de la investigación**: ¿Cómo mejorar el procesamiento y el análisis de los datos estadísticos de bienes materiales en el departamento de análisis estadístico del MINCEX?

El problema planteado anteriormente tiene como **objeto de estudio** los almacenes de datos y el **campo de acción** está enmarcado sobre el procesamiento y análisis de los datos estadísticos de bienes materiales, que son las áreas de conocimiento que recibirán una transformación relevante. Trazándose como **objetivo general** desarrollar un mercado de datos que permita mejorar el procesamiento y análisis de los datos estadísticos de bienes materiales en el departamento de análisis estadístico del MINCEX.

# Introducción

---

## **Hipótesis:**

Si se realiza un mercado de datos mejorará el procesamiento y análisis de los datos estadísticos de bienes materiales en el departamento de análisis estadístico del MINCEX.

Para darle cumplimiento al objetivo general se definen las siguientes **tareas de la investigación:**

1. Establecimiento de los fundamentos teórico-metodológicos para el desarrollo de los almacenes de datos.
2. Caracterización de las metodologías, tecnologías y herramientas a utilizar en el desarrollo del Mercado de datos estadísticas de bienes para el departamento de análisis estadístico del MINCEX.
3. Caracterización del proceso de análisis estadístico de bienes materiales en cuanto a procesamiento y análisis de los datos en el departamento de análisis estadístico del MINCEX.
4. Desarrollo del Mercado de datos estadísticas de bienes para el departamento de análisis estadísticos del MINCEX
5. Validación de la contribución lograda en cuanto a la mejora en el procesamiento y el análisis de los datos mediante la introducción del mercado de datos en un ambiente similar al departamento de análisis estadístico del MINCEX.

## **Métodos de la investigación:**

Para el desarrollo de esta investigación se hizo necesario apoyarse en los siguientes métodos de la investigación.

# Introducción

---

## Métodos teóricos

- ▶ **Analítico-Sintético:** Este método se utilizó en el estudio del estado del arte, facilitando el análisis de las bibliografías existentes y la síntesis de los elementos más importantes enfocados al entorno de la situación problemática.
- ▶ **Hipotético-Deductivo:** Este método permitió a partir de conocimientos adquiridos en los proyectos de almacenes de datos que se han realizado anteriormente, definir hipótesis en la investigación.

## Métodos empíricos:

- ▶ **Entrevistas:** Se utiliza en el proceso de estudio preliminar para obtener la información necesaria acerca de las características y funcionalidades que debe poseer el sistema a implementar. La entrevista se realiza de forma personal utilizando preguntas previamente elaboradas (Ver anexo 1). Como constancia de las mismas se obtienen las minutas de reunión. (Ver anexo 2)
- ▶ **Análisis documental:** Permitió el estudio de documentación referente al análisis estadístico de bienes materiales, para lograr un mejor entendimiento de los procesos.

La **población** está constituida por el departamento de análisis estadístico, siendo la **muestra** dos especialistas del departamento.

**Variables independientes:** Mercado de datos.

**Variables dependientes:** Procesamiento de los datos y análisis de los datos.

## Operacionalización de las variables:

Variables	Dimensiones	Indicadores	Categorías
-----------	-------------	-------------	------------

# Introducción

Independientes			
Mercado de Datos	Almacenamiento	Seguro	Si: Los datos almacenados son históricos y no volátiles.
			No: Los datos almacenados son volátiles.
	Servicio	Interactivo	Si: Permite análisis similares desde diferentes perspectivas.
			No: No permite análisis similares.
		Optimizado	Si: En un servidor estable los tiempos de respuesta son aceptables.
			No: Los tiempos de respuesta depende en gran medida de la calidad del servidor.
Variables dependientes	Dimensiones	Indicadores	Categorías
Procesamiento de los datos	Validación de los datos	Esfuerzo	Alto: La validación de los datos se realiza manualmente.
			Medio: Se realiza manual con la ayuda de herramientas informáticas.
			Bajo: La validación de los datos es automatizada.
		Seguridad	Alta: Los datos son consistentes.
			Baja: Los datos son inconsistentes.
	Almacenamiento	Seguridad	Baja: Los datos almacenados son volátiles.



# Introducción

	de los datos		Alta: Los datos almacenados son históricos.
Análisis de los datos	Presentación de los datos	Interactividad	No: Cuando se necesitan varias estructuras para visualizar los mismos datos de forma detallada y resumida a la vez
			Si: Cuando no se necesitan varias estructuras para visualizar los mismos datos de forma detallada y resumida a la vez.
		Confiabilidad	Baja: Los datos son inconsistentes y necesitan manipulación humana.
	Media: Los datos son consistentes pero necesitan manipulación humana.		
	Alta: Los datos son consistentes y no necesitan manipulación humana.		
	Disponibilidad	Accesibilidad	Baja: Los datos están disponibles en un solo ordenador
Alta: Los datos están disponibles desde varios ordenadores y se puede acceder simultáneamente a ellos.			

El **aporte práctico** de esta investigación consiste en la creación del Mercado de datos estadísticas de bienes para el departamento de análisis estadístico del MINCEX.

## Estructura capitular del documento

### Capítulo 1: “Fundamentación Teórica”:

En este capítulo se realizará un estudio del estado del arte de la investigación, se

# Introducción

---

abordarán los principales conceptos que serán de utilidad para tener un dominio sobre los temas tratados para el desarrollo del mercado de datos, se realizará la caracterización de la metodología seleccionada y las herramientas utilizadas para el tratamiento necesario de la información.

## **Capítulo 2: “Análisis y diseño de la solución propuesta”.**

En este capítulo se describe la propuesta de solución para la situación problemática anteriormente planteada. Se llevará a cabo la definición de los requerimientos del sistema, se diseñarán el diagrama de casos de uso del sistema, el estándar de código y los subsistemas de integración y visualización.

## **Capítulo 3: “Implementación de la solución propuesta”.**

En este capítulo se realizará la implementación de los subsistemas de integración y visualización.

## **Capítulo 4: Validación de la solución propuesta”.**

Se validará el mercado de datos a través de los casos de prueba definidos.

## Capítulo 1. Fundamentación Teórica

### Introducción

La información es vital en la mayoría de las actividades que se desarrollan en las sociedades. Los grandes negocios necesitan una tecnología que les permita predecir los problemas, evaluar nuevos procedimientos y resultados; basándose fundamentalmente en el análisis de la información. Como una buena alternativa para resolver este problema surgen los almacenes de datos.

En este capítulo se refleja un estudio del estado actual de los almacenes de datos a nivel internacional y nacional, se abordan los principales conceptos que serán de utilidad para tener un dominio sobre los temas tratados, y se realiza la caracterización de la metodología seleccionada y las herramientas utilizadas para el tratamiento necesario de la información.

### 1.1 Conceptos asociados al dominio del problema

#### .Base de datos

*Se define una base de datos como una serie de datos organizados y relacionados entre sí, los cuales son recolectados y explotados por los sistemas de información de una empresa o negocio en particular. (Pérez Damián, 2008)*

### 1.2 Almacenes de datos

El uso de los almacenes de datos aumenta considerablemente en sectores donde existe un gran volumen de transacciones y de información acumulada a través de los años. Ejemplo de estos sectores son:

- ▶ *Sistemas de marketing: Investigación comercial, segmentación de mercados, identificación de necesidades no cubiertas, definición de la estrategia de*

*canales de comercialización y distribución, definición de la estrategia de promoción y atención al cliente, seguimiento de acciones comerciales, apoyo al canal de venta con información cualificada*

- ▶ *Análisis de riesgo financiero: Análisis, monitorización de riesgos contra límites, simulación de escenarios históricos, análisis de rentabilidad, establecimiento y seguimiento de límites.*
- ▶ *Control de Gestión: Análisis de desviaciones, realización de reportes.*
- ▶ *Recursos Humanos: Planificación de incorporaciones, gestión de carreras profesionales, asignación de recursos a proyectos alternativos.*
- ▶ *Turismo: Cadenas hoteleras, agencias de viajes.*
- ▶ *Salud: Análisis de datos clínicos, combinación y creación de informes para datos diferenciados, realización de estudios complejos.(Dataprix, 2009)*

## **.A nivel Internacional**

Existentes innumerables ejemplos de aplicaciones de los almacenes de datos a nivel internacional. *En el mundo empresarial, Sybase y Sun han unido fuerzas para desarrollar una arquitectura de referencia para el almacén de datos de la empresa, la arquitectura desarrollada abrió la posibilidad de la creación y desempeño del Almacén de Datos más grande del mundo. Esta solución suple las necesidades de: simplicidad, flexibilidad, escalabilidad, administración y protección de la inversión (Silicon News, 2007)*

*El almacén de datos SIE (Sistema de Información Estadística) está estructurado en base a módulos temáticos, en los que se agrupa información sobre la Comunidad Autónoma de Castilla y León para un mejor conocimiento de la misma. (SIE, 2008)*

El almacén de datos del Banco Central Europeo está dirigido a un amplio espectro de usuarios (analistas financieros, periodistas, empresas, economistas,

investigadores, etc...) *Proporciona los datos macroeconómicos en línea en la zona del euro: Inflación, los tipos de cambio, costos laborales, el desempleo, productividad, déficit público, deuda pública, balanza de pagos, posición de inversión internacional, datos sociodemográficos, información fiscal, etc.* (Vidal Juan, 2011)

## **.En Cuba**

La economía cubana aprovecha la tecnología de los almacenes de datos desde hace varios años. La adopción de esta ha proporcionado nuevas herramientas para la toma de decisiones. Los almacenes de datos permiten que la información sea consultable, medible y comparable, debido a ello se incrementa el número de empresas que han escogido estos sistemas para almacenar su información. Algunos de los almacenes de datos desarrollados en Cuba en los diferentes sectores:

- ▶ Almacén del sistema de información hospitalaria: facilitar la gestión de información en ésta área de las instituciones hospitalarias.
- ▶ Almacén de datos para la red nacional de genética médica: facilita el apoyo a la toma de decisiones.
- ▶ Almacén de datos operacionales para la corporación CIMEX: facilite el análisis de la información y la toma de decisiones gerenciales.
- ▶ Almacén de datos estadísticos de la ONE: Desarrollo de la capa de visualización del mercado de datos demografía.

### **1.2.1 ¿Qué es un Almacén de datos?**

Existen dos importantes conceptos acerca de los mismos, el de Ralph Kimball y el de Bill Inmon.

Según Inmon *un almacén de datos es una colección de carácter temático, integrado, no volátil y variante en el tiempo en apoyo de la toma de decisiones de la administración.* (SIS KLE, 2009)

Según Kimball *un almacén de datos es una copia de las transacciones de datos específicamente estructurada para la consulta y el análisis.* (Gestión, 2011)

Un almacén de datos puede definirse como una colección de datos provenientes de múltiples y diversas fuentes que proporciona un medio para el análisis de la información y brinda apoyo al proceso de la toma de decisiones. Sobre este contexto son varios los paradigmas que se tienen en consideración a nivel mundial.

## **.Paradigma Bill Inmon**

Bill Inmon ve la necesidad de transferir la información de los diferentes sistemas transaccionales (OLTP) de las organizaciones a un lugar centralizado donde los datos puedan ser utilizados para el análisis. Insiste además en que ha de tener las siguientes características:

- ▶ *Orientado a temas: Los datos en la base de datos están organizados de manera que todos los elementos de datos relativos al mismo evento u objeto del mundo real queden unidos entre sí.*
- ▶ *Integrado: La base de datos contiene los datos de todos los sistemas operacionales de la organización, y dichos datos deben ser consistentes.*
- ▶ *No volátil: La información no se modifica ni se elimina, una vez almacenado un dato, éste se convierte en información de sólo lectura, y se mantiene para futuras consultas.*
- ▶ *Variante en el tiempo: Los cambios producidos en los datos a lo largo del tiempo quedan registrados para que los informes que se puedan generar reflejen esas variaciones.* (Espinosa Roberto, 2010)

## **.Paradigma Ralph Kimball**

*Un almacén de datos es un conglomerado de todos los mercados de datos dentro de una empresa, siendo una copia de los datos transaccionales estructurados de una forma especial para el análisis, de acuerdo al Modelo Dimensional, que incluye las dimensiones de análisis y sus atributos, su organización jerárquica, así como los diferentes hechos de negocio que se quieren analizar. Por un lado se tienen tablas para representar las dimensiones y por otro lado tablas para los hechos. Los diferentes mercados de datos están conectados entre sí por la llamada estructura de bus, que contiene los elementos anteriormente citados a través de las dimensiones conformadas y permiten que los usuarios puedan realizar consultas sobre los diferentes mercados de datos, pues este bus contiene los elementos en común que los comunican. (Espinosa Roberto, 2010)*

Los almacenes de datos deben tener las siguientes características:

- ▶ El contenido del almacén de datos debe ser entendible: Los datos deben ser intuitivos y obvios para los usuarios del negocio.
- ▶ La información del almacén de datos debe ser consistente: Los datos del almacén deben ser ensamblados cuidadosamente desde una variedad de fuentes pertenecientes a la organización, limpiados y con calidad garantizada. La información de un proceso de negocio debe coincidir con la información de otro.
- ▶ El almacén de datos debe ser resistente al cambio: Las necesidades de los usuarios, las condiciones del negocio, los datos y la tecnología están sujetos a cambios. El almacén debe ser diseñado para evitar estos posibles cambios.
- ▶ Debe ser un bastión seguro que proteja la información: Como mínimo un

almacén contiene información sensible acerca de lo que se está vendiendo, los precios, principales clientes. Esta información es potencialmente dañina en manos equivocadas.

- ▶ Debe ser la base para la mejora del proceso de toma de decisiones: El almacén debe tener datos correctos. Solo existe una verdadera salida desde un almacén de datos; las decisiones que se toman basadas en las evidencias del almacén de datos.
- ▶ Los usuarios del negocio deben aceptar el almacén una vez que haya sido desarrollado: no importa que se haya construido una solución elegante utilizando los mejores productos y tecnologías, si el usuario no acepta el almacén y continua usándolo al menos seis meses, entonces habrá fallado las pruebas de aceptación.

## 1.2.2 Minería de datos

*La minería de datos, es el conjunto de técnicas y tecnologías que permiten explorar grandes bases de datos, de manera automática o semiautomática, con el objetivo de encontrar patrones repetitivos, tendencias o reglas que expliquen el comportamiento de los datos en un determinado contexto.*

*Básicamente, la minería de datos surge para darle solución a problemas de predicción, clasificación y segmentación. Con este fin, hace uso de prácticas estadísticas y, en algunos casos, de algoritmos de búsqueda próximos a la inteligencia artificial y a las redes neuronales. (Sinnexus, 2011)*

## 1.2.3 Mercados de datos

*Un mercado de datos es una base de datos departamental, especializada en el almacenamiento de los datos de un área de negocio específica. (Soto Lauro, 2008)*



*Un mercado de datos es una versión especial de un almacén de datos. Contienen una visión de datos operacionales que ayudan a decidir sobre estrategias de negocio basadas en el análisis de tendencias y experiencias pasadas. (Cavsi, 2011)*

Se caracteriza por disponer la estructura óptima de datos para analizar la información al detalle desde todas las perspectivas que afecten a los procesos de dicho departamento. Un mercado de datos puede ser alimentado desde los datos de un almacén de datos, o integrar por sí mismo un compendio de distintas fuentes de información.

Se resume de las anteriores definiciones que un mercado de datos es una alternativa de solución a los problemas al igual que un almacén de datos, porque presentan las mismas características en cuanto a diseño y al modo de construcción. La diferencia entre estos es que un mercado de datos se enfoca en proporcionar información de un área de negocios específica y el almacén lo hace a nivel corporativo.

#### **1.2.4 Modelo Dimensional**

*El modelo dimensional es esencialmente un modelo lógico de los requerimientos del negocio. El modelo de datos dimensional es el concepto de diseño más utilizado para construir sistemas de almacenamiento de datos. La idea fundamental del modelo es que los datos de negocio pueden ser representados como un tipo de cubo de datos. El modelo dimensional utilizado en el modelado de almacenes de datos, organiza y presenta los datos definiendo dimensiones.” (Oporto Samuel, 2008).*

El modelo dimensional consta de relaciones entre tablas nombradas hechos que almacenan valores cuantitativos llamados medidas y tablas específicas nombradas

dimensiones la cual contiene niveles o jerarquías.

## **.Elementos que integran el modelo de datos dimensional:**

### ▶ **Dimensiones**

Una tabla de dimensiones almacena un conjunto de valores que están relacionados a una dimensión particular. *“Las tablas de dimensiones almacenan los valores que se utilizan en las tablas de hechos.”*(Microsoft, 2011)

### ▶ **Jerarquías**

Son grupos de atributos que representan niveles de agrupación en las dimensiones.

### ▶ **Tablas de Hechos**

*Es la tabla central en un esquema dimensional. Es en ella donde se almacenan las mediciones numéricas del negocio. Estas medidas se hacen sobre el grano, o unidad básica de la tabla. Cada medida es tomada de la intersección de las dimensiones que la definen.* (Proyecto datawarehousing/BI, 2010)

### ▶ **Medida**

*Es un atributo numérico de un hecho que representa el comportamiento del negocio relativo a la dimensión.* (Oporto Samuel, 2008).

## **1.2.5 Sistemas de procesamiento analítico en línea**

OLAP (del inglés On-Line Analytical Processing). *Es una solución utilizada en inteligencia empresarial (BI del inglés Business Intelligence) cuyo objetivo es agilizar la consulta de grandes cantidades de datos. Para ello utiliza estructuras multidimensionales, o cubos OLAP, que contienen datos resumidos de grandes bases de datos.* (SIS KLE, 2009)

Existen diferentes tipos de sistemas OLAP:

### **.OLAP multidimensional (MOLAP)**

*Almacena los datos en forma dimensional para realizar el acceso a la información almacenada de forma más rápida y efectiva utilizándose en depósito donde el tiempo en la velocidad de respuesta es crítico, 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 almacén de datos. (Ortiz Julio, 2009)*

## **.OLAP relacionales (ROLAP)**

*Los datos son almacenados en filas y columnas de forma relacional. Este modelo presenta los datos a los usuarios en forma de dimensiones de negocio. Con el fin de ocultar las estructuras de almacenamiento y presentar los datos dimensionalmente es creada la semántica de las etiquetas de los metadatos. Estos metadatos también son almacenados en tablas relacionales. (Ortiz Julio, 2009)*

## **.OLAP Híbrido (HOLAP)**

*Como su nombre lo indica, es un híbrido entre los métodos ROLAP y MOLAP. Permite almacenar una parte de los datos como en un sistema MOLAP y el resto como en uno ROLAP. (Ortiz Julio, 2009)*

Teniendo en cuenta las características de los sistemas OLAP y al almacenar los datos de la solución propuesta en bases de datos relacionales, se escoge el sistema ROLAP para el desarrollo del mercado de datos.

## **1.2.6 Esquemas Dimensionales**

### **.Esquema estrella**

*Deriva su nombre del hecho que su diagrama forma una estrella, con puntos radiales desde el centro. El centro de la estrella consiste de una o más tablas de hechos, y las puntas de la estrella son las dimensiones. Este modelo entonces,*

*resulta ser asimétrico, pues hay una tabla dominante en el centro con varias conexiones a las otras tablas. Las tablas de dimensiones tienen sólo la conexión a la tabla de hecho y ninguna más. (Proyecto datawarehousing/BI, 2010)*

## **.Esquema en copo de nieve<sup>1</sup>**

Es una estructura algo más compleja que el esquema en estrella. Se da cuando alguna de las dimensiones se implementa con más de una tabla de datos. La finalidad es normalizar las tablas y así reducir el espacio de almacenamiento al eliminar la redundancia de datos; pero tiene la contrapartida de generar 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. En el copo de nieve una dimensión es representada en varias tablas. (Proyecto datawarehousing/BI, 2010)

Para el desarrollo de la solución propuesta se escoge el esquema en copo de nieve ya que responde a las características que presenta el negocio, las cuales imposibilitan el uso de un esquema en estrella.

## **1.3 Metodología de desarrollo de software**

Las metodologías de desarrollo de software establecen un conjunto de pasos y procedimientos para el proceso de desarrollo de software, con el fin de hacerlo predecible y eficiente. En el área de los almacenes de datos las metodologías más utilizadas son las que definen Ralph Kimball y Bill Inmon.

---

<sup>1</sup> Un **copo de nieve** se da cuando alguna de las dimensiones se relaciona con una o más tablas además de la tabla de hecho.

Para el desarrollo de la solución propuesta se escogió la metodología de Kimball, tomando de guía el estudio del paradigma y la visión del autor sobre la construcción de los almacenes de datos. Kimball es una metodología ascendente que plantea la creación de mercados de datos independientes orientados a temas o procesos de un área de negocios específico de una empresa y vislumbra el almacén de datos como la unión de todos los mercados de datos de la propia empresa.

## 1.4 Herramientas a utilizar

Las herramientas que se utilizaran para crear la solución propuesta pertenecen a la familia de software libre, aprovechando las facilidades del código abierto<sup>2</sup> para sus futuros cambios y adaptaciones.

### 1.4.1 Sistema gestor de bases de datos

*Un sistema gestor de bases de datos (SGBD) es una colección de programas cuyo objetivo es servir de interfaz entre la base de datos, el usuario y las aplicaciones. Se compone de un lenguaje de definición de datos, de un lenguaje de manipulación de datos y de un lenguaje de consulta. Un SGBD permite definir los datos a distintos niveles de abstracción y manipular dichos datos, garantizando la seguridad e integridad de los mismos. (Cavsi, 2008)*

Entre los sistemas de gestión de base de datos más utilizados se encuentran:

#### **.MySQL**

*Es un sistema gestor de base de datos relacional, multihilos y multiusuario con más de seis millones de instalaciones (Oracle Corporation, 2011). MySQL incluye el conjunto más completo de características avanzadas, herramientas de gestión y*

---

<sup>2</sup> **Código abierto** es el software distribuido y desarrollado libremente.

apoyo técnico para lograr los más altos niveles de escalabilidad, seguridad, fiabilidad y tiempo de actividad. Reduce el riesgo, costo y complejidad en el desarrollo, implementación y administración de aplicaciones críticas de negocio MySQL.

## **.PostgreSQL**

*PostgreSQL es un poderoso sistema de gestión de base de datos objeto relacional de código abierto. Cuenta con más de 15 años de desarrollo activo y una arquitectura probada que se ha ganado una sólida reputación de fiabilidad, integridad y corrección de datos. Es compatible con los principales sistemas operativos, incluyendo Windows, Mac OS, Solaris, Linux, entre otros. Tiene soporte completo para claves foráneas, uniones, vistas, disparadores y procedimientos almacenados. Incluye los más importantes tipos de datos del lenguaje SQL: integer, numérico, boolean, char, varchar, date, interval y timestamp. Permite la gestión de diferentes usuarios, como también los permisos asignados a cada uno de ellos. (PostgreSQL Global DevelopmentGroup, 2011)*

Se escogió como SGBD PostgreSQL 9.1 debido a que incorpora avanzadas características en materia de seguridad, soporte de aplicaciones, seguimiento y control, rendimiento y almacenamiento de datos.

### **1.4.2 Herramienta manejadora de base de datos**

*El Sistema de administración de bases de datos (DBMS del inglés Database management system) es un conjunto de servicios (aplicaciones de software) para administrar bases de datos, permite fácil acceso a los datos, acceso a la información por parte de múltiples usuarios y la manipulación de los datos encontrados en la base de datos. (Kioskea, 2011)*

## **.PGADMIN III**

*PgAdmin III es la aplicación gráfica para PostgreSQL más completa y popular bajo licencia de código abierto. Es una herramienta multiplataforma. Está diseñada para responder a las necesidades de todos los usuarios, desde escribir consultas SQL simples hasta desarrollar bases de datos complejas. La interfaz gráfica soporta todas las características de PostgreSQL y facilita enormemente la administración.*  
(Ubuntu, 2008)

Incluye las siguientes características:

- ▶ Soporte para PostgreSQL 9.0.
  - ▶ Nueva interfaz de copia de seguridad
  - ▶ Opciones de restauración
  - ▶ Manejo de conexiones
  - ▶ Soporte para nuevo nivel de disparadores<sup>3</sup>.
  - ▶ SQL editor recuerda automáticamente las consultas que ya fueron ejecutadas y permite alternarlas y volverlas a ejecutar.
- (PostgresOnLineJournal, 2011)

Se utilizó el PgAdmin en su versión 1.12 teniendo en cuenta las características anteriormente expuestas.

### **1.4.3 Herramienta de Modelado**

*Las herramientas de modelado de datos permiten realizar un diseño de datos nuevo o comprender un modelo existente .El diseño se realiza a través de los diagramas entidad-relación y permite mejorar el diseño de los sistemas, la*

---

3 Un **disparador** (trigger) en una base de datos, es un procedimiento que se ejecuta cuando se cumple una condición establecida al realizar una operación.

*eficiencia en la programación y otorgar a la administración un mejor soporte en la documentación.* (Case, 2011). Dentro de las herramientas de modelado se encuentran:

## **.Visual Paradigm for UML**

*Es una herramienta de modelado visual para todo tipo de diagramas UML (Lenguaje de Modelado Unificado). Esta herramienta está diseñada con el objetivo de la creación de diseños basados en el paradigma de la programación orientada a objetos. Es muy utilizada por analistas, arquitectos de software, diseñadores y desarrolladores con el fin de realizar sus diseños y modelos durante el desarrollo de software.* (Visual Paradigm, 2010).

## **.Xmind**

Se utiliza para la arquitectura de la información, es una herramienta de código abierto para desarrollar mapas mentales. *XMind es una gran herramienta de intercambio de ideas, proporciona un modo especial de edición de tormenta de ideas que permite organizar y fijar ideas innovadoras. El diseño intuitivo de XMind y sus potentes características permiten centrarte en su trabajo y compartirlo fácilmente con los demás. Se puede exportar los mapas en una variedad de formatos populares, tales como PDF, Word y Power Point* (XMindLtd, 2010).

## **.Power Architect**

*Recuerda el origen de cada columna y es capaz de generar automáticamente las asignaciones de origen a destino de los datos. Proporcionará una visión completa de todas las estructuras de base de datos necesaria y agilizar todos los aspectos de su diseño de almacenamiento de datos.* (Ortiz Valmaseda Marcos, 2009).

Para apoyar la documentación se utiliza Visual Paradigm 6.4, Xmind 3.2.1 y Power



Architect 3.0.

#### **1.4.4 Suites de inteligencia de negocio**

Una suite de inteligencia de negocio está provista de herramientas que asisten el análisis y la presentación de los datos. Existen disímiles suites, algunas de ellas son:

##### **.Jaspersoft**

*Es un conjunto de herramientas muy poderosas que permiten a las organizaciones generar información basada en sus datos de administración para la evaluación y toma diaria de decisiones. (Disytel, 2011)*

##### **.Business Objects**

*Las herramientas de Business Objects permiten definir con facilidad entornos visuales para explorar jerarquías amplias y relaciones en sus aplicaciones, descubrir relaciones y analizar tendencias en datos tabulares, así como visualizar horizontes a largo plazo.(Sap, 2011)*

##### **.Pentaho**

*Está compuesto por un conjunto de herramienta poderosa para ganar visibilidad en la salud de los departamentos, del negocio y de los procesos de negocio, ofrece una plataforma la cual incluye un servidor, herramientas clientes y tecnologías que posibilitan el espectro entero del BI. (Linalis, 2011)*

Se escogió la suite desarrollada por la comunidad de Pentaho en su versión 3.6 porque además de ser gratis, es fácil de personalizar con estándares abiertos, interactúa con los informes y permite visualizar y analizar sus datos e integrar los datos desde diversas fuentes.

La suite de Pentaho está integrada por las siguientes herramientas:

## **.Pentaho MetaData Editor**

El Metadatos es una característica de la plataforma BI diseñada para facilitar a los usuarios acceder a información en términos de negocio. *Muestra cómo crear un reporte ad-hoc con la interfaz web para reportes ad-hoc, con la posibilidad de elegir elementos específicos que deseen que se muestren en el reporte.* (Pentaho, 2009)

## **.Pentaho Report Designer (PRD)**

*Es una aplicación de escritorio que proporciona un entorno de diseño visual para crear reportes. Los reportes se pueden ejecutar y guardar de forma local o publicarlos en un servidor de BI de Pentaho para permitir a muchas personas a acceder y programar la ejecución de informes. PRD está dirigido a usuarios con experiencia y poder, que están familiarizados con los conceptos y fuentes de datos utilizadas.* (Pentaho, 2010)

## **.Pentaho bisever-ce**

*Proporciona la arquitectura y la infraestructura necesaria para crear soluciones de inteligencia de negocio. Proporciona servicios básicos como la autenticación, registro, auditoría, servicios web entre otros. Incluye un motor de solución que integra la presentación de reportes y análisis. Incluye a Mondrian que es un servidor OLAP que permite a los usuarios de negocio analizar grandes cantidades de datos en tiempo real e incluye también reportes ad-hoc que extiende la capacidad de los usuarios finales para la creación de informes a partir de plantillas pre-configuradas y con la ayuda de un asistente de creación.*(Pentaho, 2010)

## **.Pentaho schema workbench**

Es la herramienta gráfica que permite la construcción de los esquemas de Mondrian, y además permite publicarlos al servidor BI para que puedan ser

utilizados en los análisis por los usuarios de la plataforma.

*Workbench ofrece un entorno de desarrollo para bases de datos. Permite comparar y sincronizar esquemas de bases de datos y el contenido de la base de datos. (MySQL, 2011)*

## **.Pentaho data integration**

*Es uno de los componentes del proyecto Pentaho, también se le conoce como Kettle. Se trata de un motor de integración de datos al que se puede acceder utilizando una interfaz gráfica para definir trabajos y transformaciones. El Kettle soporta la ejecución de los trabajos en equipos convencionales, pero también cuenta con funciones para el procesamiento distribuido. (Johan, 2011)*

## **1.5 Documentación**

### **.OpenOffice**

*Es un programa de software libre multiplataforma y multilingüe totalmente compatible con las versiones de documentos creados con Microsoft Office. OpenOffice permite crear documentos, hojas de cálculo, generar y utilizar macros, corrección ortográfica en español, generación de páginas html y documentos pdf desde cualquier aplicación de OpenOffice, utilización de plantillas, conexión a base de datos, ahorro de dinero, una sociedad más libre. (Ciberaula, 2010)*

## **Conclusiones**

En este capítulo se realizó un estudio de las tendencias, tecnologías y herramientas actuales para el diseño e implementación de los almacenes de datos. Después de este se determinó utilizar el Xmind para el diseño de mapas mentales, como herramienta CASE Visual Paradigm para realizar el diseño del mercado de datos,

# Capítulo 1

---

PostgreSQL como SGBD, PGAdmin como DBMS, Power Architect para el modelado de datos, para el BI se escogió las herramientas que proporcionan la Suite de Pentaho, como editor de gráfico el Open Flash Chart y para la documentación Open Office, todas ellas pertenecientes a la familia de software libre. Además se definió como metodología a utilizar: Arquitectura multidimensional de Ralph Kimball y se puntualizaron los principales componentes de la arquitectura de un almacén de datos.

### **Capítulo 2. Análisis y diseño de la solución propuesta.**

#### **Introducción**

En este capítulo se describe la propuesta de solución para la situación problemática anteriormente planteada. En el mismo se define el negocio, se identifican requerimientos del sistema y se diseñan casos de uso y el modelo dimensional.

#### **2.1 Definición del negocio**

En el departamento de análisis estadístico del MINCEX se analizan las cifras de importación y exportación de bienes materiales con el objetivo de conocer el estado del comercio exterior del país con el resto del mundo. De esta manera se proporciona un medio que apoye el proceso de toma de decisiones en busca de estrategias que permitan el crecimiento de la economía.

Las cifras referentes a los procesos de importación y exportación de bienes materiales de Cuba, son suministradas por la Aduana General de la República (AGR) en un archivo DBF<sup>4</sup>. Dicho archivo se exporta a una base de datos de Access que contiene los nomencladores necesarios para el análisis estadístico.

Los datos recopilados se someten a un conjunto de validaciones que garantizan integridad de los mismos y dan paso a la elaboración de reportes dirigidos al personal encargado de la toma de decisiones. El proceso de validación de los datos puede durar toda una jornada laboral. Los especialistas deben verificar el valor de las cifras suministradas y reparar la falta de nomencladores.

El departamento de análisis estadístico suministra información a la mayoría de los departamentos del MINCEX y a muchas de las empresas del país. La información

---

<sup>4</sup> DBF formato para archivos de base de datos.

suministrada se elabora en estructuras o reportes preelaborados en el propio Access, no obstante, la mayor parte de los pedidos de información se realizan sobre documentos Word, Excel y presentaciones de Power Point.

### 2.1.1 Procesos del negocio

La propuesta de solución está encaminada a los procesos identificados en el análisis estadístico de las cifras de bienes. Cada proceso marca una pauta dentro del desarrollo de la aplicación, principalmente en el diseño de las estructuras dimensionales.

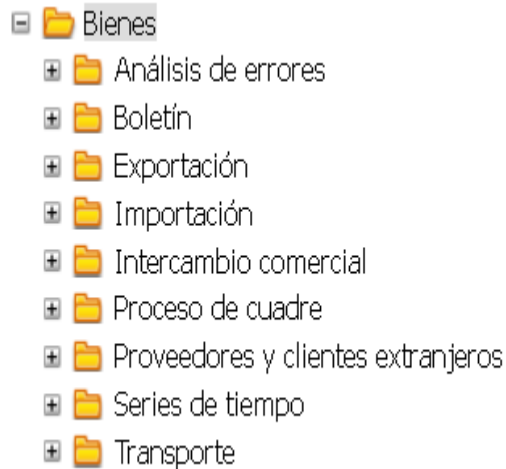
**Importación:** Es la actividad económica que suministra mayor cantidad de cifras en el negocio. Refleja la compra u obtención de productos desde otros países para uso o consumo nacional. Es producto de análisis estadístico para las empresas importadoras de Cuba y para la mayoría de los departamentos del MINCEX.

**Exportación:** Es la actividad económica que refleja la venta de productos hacia otros países. Es producto de análisis estadístico para las empresas exportadoras de Cuba y para la mayoría de los departamentos del MINCEX.

**Intercambio comercial:** Es la terminología utilizada por los especialistas del MINCEX para valorar el comportamiento del intercambio de bienes entre productos y países con respecto a Cuba.

### 2.2 Áreas de Análisis

El análisis de la información se deriva de los procesos del negocio pero se organiza por áreas bien definidas. Dichas áreas responden a la agrupación de la información según su propósito y en ocasiones agrupan estructuras que responden a varios de los procesos del negocio. Se corresponden al análisis de los indicadores específicos de:



*Figura 1 Áreas de análisis*

### **2.3 Requerimientos del Sistema**

#### **2.3.1 Requisitos de Información**

Los requisitos de información representan los pedidos de información o consultas realizadas por los clientes. Se trata de 206 requisitos distribuidos en las áreas de análisis de la siguiente manera:

- ▶ Intercambio comercial: 11 requisitos.
- ▶ Exportación: 54 requisitos
- ▶ Importación: 63 requisitos.
- ▶ Proveedores y clientes extranjeros: 10 requisitos.
- ▶ Transporte: 13 requisitos.
- ▶ Proceso de cuadro: 6 requisitos.
- ▶ Boletín: 19 requisitos.
- ▶ Series de tiempo: 22 requisitos.

- ▶ Análisis de Errores: 8 requisitos.

### 2.3.2 Requisitos funcionales

Los requisitos funcionales son capacidades o condiciones que el sistema debe cumplir para satisfacer las expectativas del cliente.

- ▶ Autenticar usuario
- ▶ Crear usuario.
- ▶ Modificar usuario.
- ▶ Visualizar usuario.
- ▶ Eliminar usuario.
- ▶ Crear rol.
- ▶ Modificar rol.
- ▶ Visualizar rol.
- ▶ Eliminar rol.
- ▶ Crear una 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 bienes
- ▶ Cargar hechos de bienes



### 2.3.3 Requisitos no funcionales

Los requisitos no funcionales son propiedades o cualidades que el producto debe tener, y que lo hacen atractivo, usable, rápido o confiable. Para este sistema se definieron los requisitos siguientes:

#### **.Requisitos de apariencia o interfaz externa**

- ▶ Un diseño simple y flexible, permitiendo que no sean necesarias muchas instrucciones para utilizar el sistema.
- ▶ Los reportes estadísticos deben contar con una interfaz simple que facilite la interacción usuario-aplicación.
- ▶ Los gráficos serán con los colores establecidos por el MINCEX ajustándose a los estándares establecidos de un buen diseño.
- ▶ La interfaz web debe contener el logotipo del MINCEX.

#### **.Requisitos de software**

- ▶ Por parte del servidor:
  1. Servidor de BI de Pentaho 3.6.
  2. Sistema gestor de bases de datos PostgreSQL 9.0.
  3. Máquina virtual de java 6.0 o superior.
  4. Herramienta de integración de datos Kettle 4.1 o superior.
- ▶ Por parte del cliente:
  1. Navegador web que permita el uso de JavaScript y de flash.

#### **.Requisitos de usabilidad.**

- ▶ Para utilizar el sistema es necesario poseer conocimientos elementales de computación y sobre el ambiente web en sentido general.
- ▶ El usuario debe tener conocimiento en análisis y tratamiento de la información.

- ▶ Debe existir al menos una persona que domine el trabajo con la herramienta “Kettle” para realizar o modificar las transformaciones para la integración de datos.

### **.Requisitos de seguridad.**

- ▶ Garantizar que la información sea accedida únicamente por las personas que pertenecen al rol indicado para esta actividad.

### **.Requisitos de hardware.**

- ▶ Por parte del servidor:
  1. Memoria RAM 3GB o superior.
  2. Capacidad superior a los 60 GB.
- ▶ Por parte del cliente:

Una computadora con al menos 256 MB de memoria RAM.

### **.Requisitos de restricciones de diseño.**

- ▶ El sistema debe ser legible, con una navegación sencilla.

### **.Requisitos de disponibilidad.**

- ▶ El sistema necesita estar conectado directamente a un dispositivo de red y estar disponible las 24 horas del día.

## **2.4 Casos de uso del sistema**

Los casos de uso del sistema representan gráficamente a los procesos y su interacción con los actores, el especialista de bienes, el especialista de estadísticas y el técnico informática.

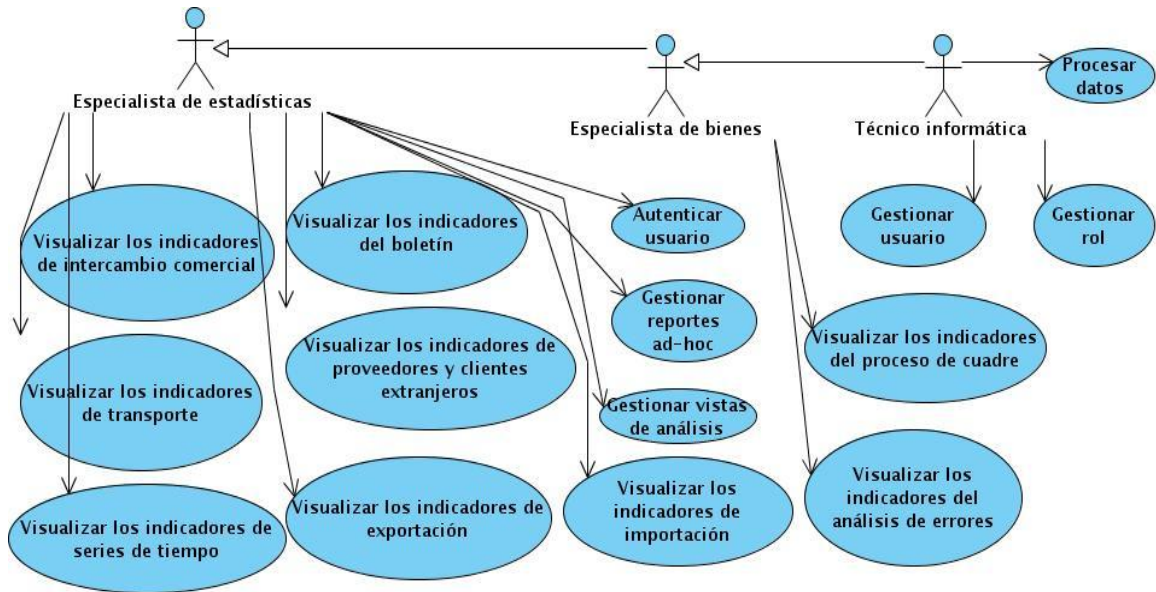


Figura 2 Diagrama de casos de uso del sistema

### 2.5 Diseño de la solución

Para el diseño de la propuesta de solución constituye un hito importante la efectiva definición y descripción de los procesos de negocio, sus dimensiones, medidas y granularidad. Todos estos elementos conllevan a la elaboración de los modelos dimensionales que serán transformados en estructuras de almacenamiento para el posterior análisis de la información suministrada.

### 2.6 Modelo de datos dimensional

Se utilizó el modelo en copo de nieve para modelar la solución propuesta ya que fue imposible utilizar el modelo en estrella para representar las organizaciones internacionales.

#### .Dimensiones

## Capítulo 2

---

Se identificaron 16 dimensiones que tienen relación de uno a muchos con las tablas de hecho, exceptuando las dimensiones país y organizaciones internacionales donde existe una relación de muchos a muchos, formando un modelo en copo de nieve.

- ▶ Empresa (dim\_empresa): La dimensión describe las características de la empresa, incluyendo el tipo de empresa y el organismo al que pertenece. Está compuesta por: el nombre, el tipo de la empresa, el organismo y el REUP.
- ▶ Organizaciones internacionales (dim\_organizaciones\_internacionales): La dimensión contiene el nombre de las organizaciones internacionales.
- ▶ País (dim\_pais): La dimensión describe las características del país, regiones, subregiones, la región policome y el país de interés.
- ▶ Tiempo (dim\_tiempo): La dimensión describe los valores bajo los cuales puede clasificarse la información atendiendo a los intervalos de tiempo. Está compuesta por: año, trimestre, mes y fecha.
- ▶ Arancel (dim\_arancel): La dimensión arancel describe las características del producto, así como sus clasificaciones según varios códigos. Está compuesta por: ministros, grupo de importación, grupo de exportación, clasificación, categoría económica, comité, sección, capítulo, partida y subpartida.
- ▶ Condición de entrega (dim\_condicion\_entrega): La dimensión contiene los nombres de las condiciones de entrega.
- ▶ Unidad de medida (dim\_um): La dimensión contiene los nombres de las unidades de medidas y las descripciones de las mismas.
- ▶ Acuerdo (dim\_acuerdo): La dimensión contiene los nombres de los

acuerdos.

- ▶ Transporte (dim\_transporte): La dimensión contiene los nombres de los transportes.
- ▶ Declaración de mercancía (dim\_dm): La dimensión contiene el nombre y el código de las declaraciones de mercancía.
- ▶ Lugar de embarco (dim\_lugar\_embarco): La dimensión contiene el nombre del lugar de embarco.
- ▶ Aduana (dim\_aduana): La dimensión describe los nombres de las aduanas.
- ▶ Condición de pago (dim\_condicion\_pago): La dimensión describe las condiciones de pago.
- ▶ Régimen definitivo (dim\_regimen\_definitivo): La dimensión describe los regímenes definitivos.
- ▶ Cliente proveedor (dim\_cliente\_proveedor): La dimensión describe las características del cliente-proveedor. Contiene el nombre y el código del cliente y del proveedor.

### .Tablas de hechos

Este modelo está compuesto por 3 tablas de hecho; Intercambio, Importación y Exportación. Por cada tabla de hecho se confeccionan dos vistas materializadas<sup>5</sup> las cuales van a almacenar los datos de las cifras de bienes para el año actual y los últimos dos años respectivamente.

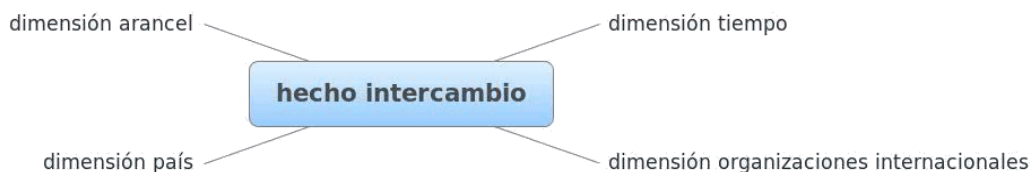
- ▶ Intercambio comercial (hech\_intercambio)

Almacena la información del proceso de intercambio comercial y contiene las

---

<sup>5</sup> Una **vista materializada** corresponde a una tabla que almacena física y automáticamente los datos resultantes de una consulta.

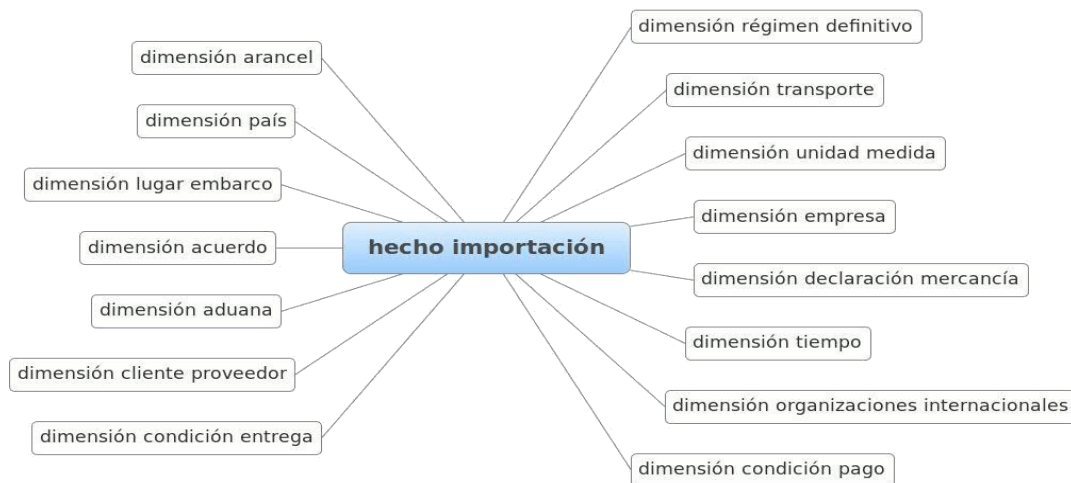
medidas valor intercambio, valor exportado y valor importado.



**Figura 3 Hecho intercambio comercial y dimensiones asociadas**

► Importación (hech\_importacion)

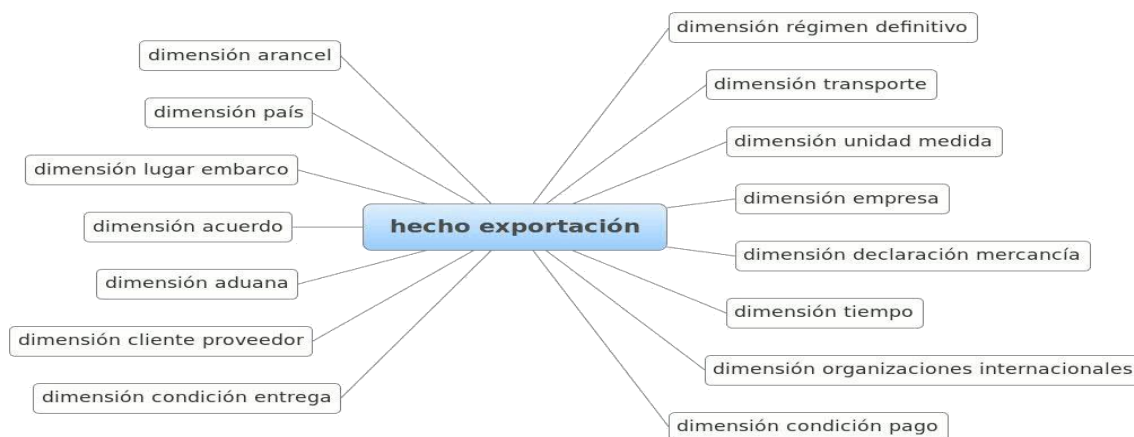
Almacena la información del proceso de importación y contiene las medidas; valor importado, valor factura, flete, precio unitario, seguro, peso bruto, peso neto, otros gastos, sacrificio cuc, sacrificio mn, sacrificio usd, servicios cuc, servicios mn, servicios usd, arancel usd, arancel cuc, arancel mn, cantidad bultos, cantidad artículos, base impuestos, número declaración mercancía, número adicional, año y código adicional.



**Figura 4 Hecho importación y dimensiones asociadas.**

► Exportación (hech\_exportacion)

Almacena la información del proceso de exportación y contiene las medidas: valor exportado, valor factura, flete, precio unitario, seguro, peso bruto, peso neto, otros gastos, sacrificio cuc, sacrificio mn, sacrificio usd, servicios cuc, servicios mn, servicios usd, arancel usd, arancel cuc, arancel mn, cantidad bultos, cantidad artículos, base impuestos, número declaración mercancía, número adicional, año y código adicional.



**Figura 5 Hecho exportación y dimensiones asociadas.**

### 2.7 Seguridad de los datos

La seguridad de los datos se basa en la definición de roles y permisos, controlando que el usuario autenticado solo pueda acceder a la información reglamentada. Los roles definidos son:

- ▶ Técnico informática.
- ▶ Especialista de estadísticas.
- ▶ Especialista de bienes.

## 2.8 Arquitectura del sistema

La arquitectura del sistema define las estructuras del sistema y la relación entre cada elemento o componente. En la solución propuesta se utilizó una arquitectura ascendente centrada en el uso de mercados de datos, siguiendo las pautas de la metodología de Kimball.



Figura 6 Arquitectura del sistema.

### .Fuentes operacionales

Bases de datos y archivos provenientes de los sistemas operacionales de donde se extraerán las cifras de bienes materiales. Consta de dos componentes:

- ▶ Archivos en formato DBF con los indicadores de los procesos de importación y exportación.
- ▶ Base de datos de Access con los nomencladores necesarios para el análisis de los indicadores.

### .Área de estacionamiento

Representación simbólica de los procesos de extracción, transformación y carga (ETL por sus siglas en inglés). Está constituida por los archivos de la herramienta



de integración de datos y todas las tablas de almacenamiento temporal necesarias para ejecutar las transformaciones. Es decir, en esta área se almacenan los datos temporalmente y luego se les aplica un conjunto de transformaciones para posteriormente almacenarlos en el área de presentación.

### **. Área de presentación**

Estructuras de almacenamiento de los datos. Son todas las tablas del mercado de datos. Dichas tablas almacenan permanentemente las cifras del comercio de bienes materiales. 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. Este elemento es una representación simbólica del servidor de BI o alguna otra herramienta que facilite el análisis de la información almacenada. Contiene todos los archivos de las estructuras predefinidas para el análisis de las cifras de bienes, ya sea reportes, gráficos o vistas de análisis.

## **2.9 Arquitectura de la información**

La arquitectura de la información refleja un mapa del sistema. Cada elemento del mapa define un esquema de navegación a seguir por el usuario para encontrar las estructuras predefinidas para el análisis de las cifras del comercio de bienes materiales.

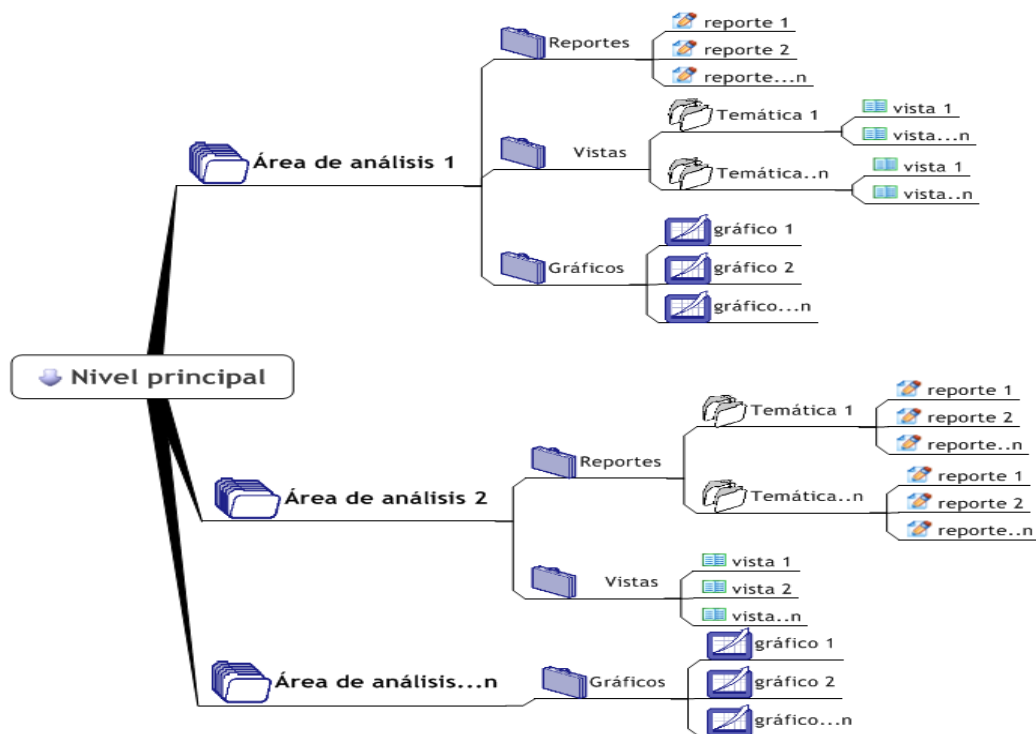



Figura 7 Mapa de navegación


**Área de análisis:** Responde a la agrupación de la información según su propósito cuyo criterio depende de las necesidades del departamento de estadísticas.

**Libro de trabajo:** Agrupa los reportes generados dentro de las áreas de análisis. Solo existen tres tipos de libros de trabajo, reportes, gráficos y vistas de análisis.

**Temática:** Se utilizan en caso de que un libro de trabajo no permita la organización completa de la información, ya que dentro del mismo pueden existir otros criterios de organización. Las temáticas se agrupan dentro de los libros de trabajo.

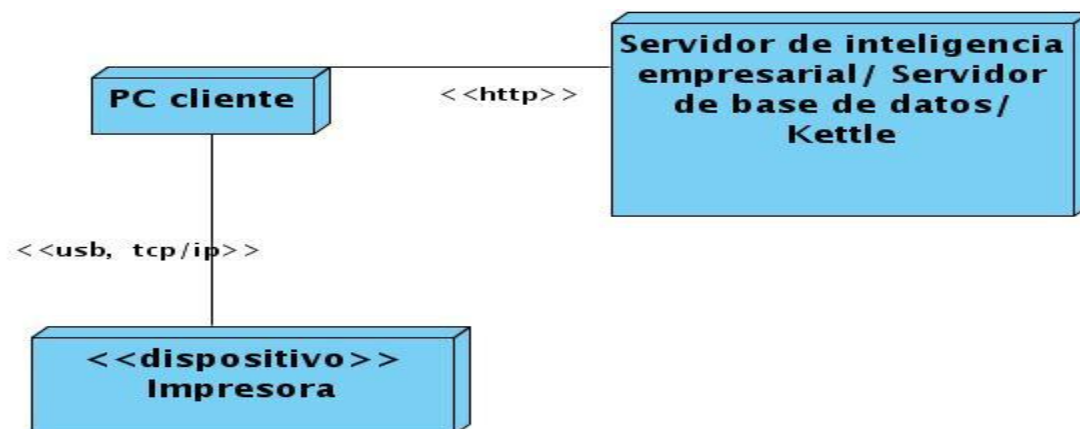
 **Reporte:** Unidad básica utilizada para mostrar la información almacenada.

 **Vista de Análisis:** Unidad básica utilizada para mostrar la información almacenada. Permite agrupar y detallar la información para los indicadores escogidos.

 **Gráfico:** Unidad básica utilizada para mostrar la información almacenada. Permite visualizar cambios en los indicadores a través de colores y tamaños.

### 2.9.1 Vista de despliegue

La vista de despliegue otorga una apreciación visual de la ubicación física de los componentes de la aplicación y los protocolos de comunicación.



*Figura 8 Diagrama de despliegue*

#### **.PC cliente**

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

## Capítulo 2

---

### **.Servidor**

El servidor es un Dual Core con 2 GB de RAM y 500 GB de disco duro. Trabaja con el sistema operativo Ubuntu Server 10.4 y en él se almacena el Mercado de datos estadísticas de bienes para el departamento de análisis estadístico del MINCEX.

### **.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**

En el desarrollo de este capítulo se brinda una breve descripción del negocio, destacando los aspectos más importantes a tener en cuenta para el desarrollo del Mercado de datos estadísticas de bienes para el departamento de análisis estadístico del MINCEX. Se explican elementos del análisis, arquitectura y diseño de la solución propuesta.

## Capítulo 3. “Implementación de la solución propuesta”.

### Introducción

En este capítulo se describen los elementos fundamentales de la implementación de los subsistemas de integración y visualización del mercado de datos. Se explicará el proceso de extracción, transformación y carga, la creación de vistas materializadas, cubos OLAP, metadatos y el proceso de visualización de los datos.

### 3.1 Estándar de Código

El estándar de código se aplica en todos los elementos de la implementación. Está presente en el diseño de las tablas físicas, en el desarrollo de las vistas materializadas y en la implementación de los subsistemas de integración y visualización. Es importante señalar que durante la implementación no se utilizan caracteres extraños ni tildes.

- ▶ 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\_estadisticas

  - 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]

Disparadores.....fc\_tgr\_[nombre tabla]

Particionamiento.....fc\_prt\_[nombre tabla]

Replicación.....fc\_rp\_[nombre tabla]

► Vistas

Vistas.....v\_ [nombre]

► Índices

Índices.....ldx\_ [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, \*

### 3.2 Subsistema de integración

A través del proceso de ETL se gestionan datos obtenidos de múltiples fuentes, con el fin de recolectarlos, limpiarlos y cargarlos en el mercado de datos. Este proceso constituye una pauta importante para la integración de los datos en una única fuente, el mercado de datos. Dicho proceso se divide de la siguiente manera:

#### **.Extracción**

El inicio de la estrategia ETL consiste en adquirir o sustraer los datos brutos desde los sistemas de origen.

#### **.Transformación**

En esta fase se aplican una serie de procedimientos sobre los datos extraídos basados principalmente en las reglas del negocio. Algunos datos, requirieren una pequeña manipulación, en otros es necesario aplicar algunas transformaciones específicas.

#### **.Carga**

La fase de carga es el momento donde los datos provenientes de la fase anterior son insertados en las tablas del mercado de datos.



*Figura 9 Representación del proceso de ETL*

### 3.2.1 Estrategia de ETL

El proceso ETL se realiza de modo específico para las tablas de dimensiones y las de hechos. Los datos de las dimensiones se almacenan temporalmente en tablas de configuración. Acto seguido se extraen los datos de estas tablas y se realizan transformaciones específicas sobre ellos. Posteriormente dichos datos serán insertados en las respectivas tablas de dimensiones.

Algunas de las acciones que se realizan en la transformación de estos datos según se requiera en cada caso son las siguientes:

1. Búsqueda en base de datos. Se verifica coincidencia entre datos que se relacionen de la tabla de entrada y una tabla especificada para obtener nuevos datos de la misma.
2. Selecciona/Renombrar. Se utiliza en la transformación para fijar los campos luego de cada acción importante y antes de cargar los datos para renombrar los campos garantizando la coincidencia con los nombres de los campos de la tabla destino. Permiten tener un control de los datos que van quedando



luego de cada transformación. Es el componente que más se utiliza en todo el proceso.

Los datos de los hechos requieren un nivel más alto de transformación. En los hechos no se almacenan valores vacíos, además las llaves deben coincidir exactamente con los identificadores de las dimensiones. Por otro lado, es necesario llevar a cabo un conjunto de validaciones en correspondencia con las reglas del negocio, garantizando así la integridad de los datos.

Por tanto, en el caso de importación y exportación se debe realizar una búsqueda por cada una de las dimensiones con las que estos se relacionan. Además, es necesario el uso de filtros y sentencias de validación durante el proceso de transformación de los datos.

Los datos del hecho de intercambio comercial se cargan para una tabla temporal, utilizando la misma secuencia de ETL de importación y exportación. Luego dichos datos se someten a un conjunto de validaciones y se agrupan para posteriormente ser almacenados en la base de datos del mercado de datos.

Durante el proceso de ETL realizado sobre los hechos se llevan a cabo las transformaciones efectuadas en las dimensiones, pero también se utilizan otras como:

1. Ordenar filas. Se especifican los campos por los cuales se quiere ordenar.
2. Filas únicas. Se especifican los campos que no se pueden repetir en las de una fila. Para que esta transformación sea efectiva, es necesario ordenar los datos con anterioridad
3. Añadir constante. Se añade una nueva constante que se necesite utilizar en las transformaciones. Dicha constante puede tomar valores diferentes en dependencia de cada secuencia de transformación.

4. Valor JavaScript. Se validan datos y permite modificar valores de entrada. Es un componente que utiliza el lenguaje JavaScript y permite introducir condiciones, asignaciones y hasta ciclos durante el proceso de transformación.
5. Filtrar filas. Permite validar datos, posee condicionales y de acuerdo con el resultado toma caminos diferentes. Su mayor uso tuvo lugar en la validación de datos vacíos. Con la ayuda de este componente es posible almacenar las filas incorrectas en archivos diferentes y de fácil acceso para los clientes, facilitando su rápida corrección.

### 3.2.2 Integridad de los datos

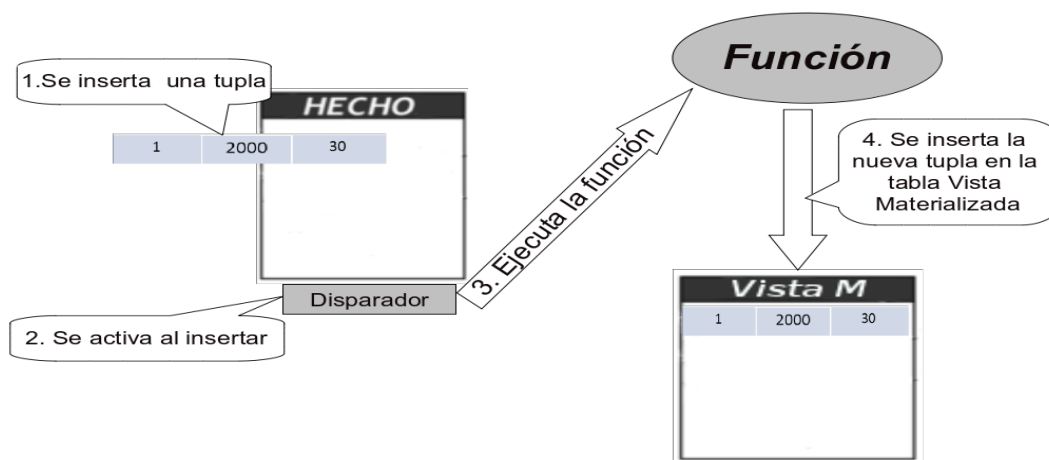
Durante el proceso de ETL, existen un conjunto de reglas o condiciones que se deben tener en cuenta para evitar posibles errores, garantizando así la integridad de los datos.

- ▶ La numeración de la declaración de mercancía no puede repetirse para fechas o países diferentes.
- ▶ El código de proveedores debe tener 4 dígitos (la parte numérica).
- ▶ El código del país debe tener 2 dígitos (la parte numérica).
- ▶ El código de la subpartida debe tener 8 dígitos (la parte numérica).
- ▶ No puede existir régimen de importación definitivo (código 4000) para país proveedor Cuba.
- ▶ No puede existir régimen de exportación definitivo (código 1000) para país cliente Cuba.
- ▶ En el caso de importaciones el país proveedor no puede ser Estados Unidos.
- ▶ En el caso de exportaciones el país cliente no puede ser Estados Unidos.

### 3.3 Vistas materializadas

Las vistas materializadas se utilizaron como estrategia para agilizar las consultas que se realizan sobre las bases de datos. Es importante señalar que PostgreSQL no facilita la creación de este tipo de vistas, además su utilización conlleva al aumento del tamaño de almacenamiento. No obstante, el resultado obtenido es satisfactorio, además; las vistas se pueden construir y mantener con el uso de disparadores y funciones disparadoras<sup>6</sup>.

Las funciones realizadas controlan y mantienen el subconjunto de datos almacenado en cada una de las vistas materializadas. Los disparadores son los encargados de activar cada una de las funciones, en dependencia de la acción que se lleve a cabo sobre los datos de las tablas de hecho. Como ejemplo de este proceso, se muestra la siguiente figura:



**Figura 10 Representación del proceso de las vistas materializadas**

<sup>6</sup> Las **funciones disparadoras** son funciones enlazadas a operaciones sobre los datos. Ejecutan un tipo de consulta específico en dependencia de la acción realizada sobre una tupla.

### 3.4 Subsistemas de visualización

Los subsistemas de visualización se implementaron sobre las capacidades del servidor de BI de Pentaho en su versión 3.6. Se implementó el sistema ROLAP con la ayuda del Mondrian y de la interfaz definida para la creación de reportes ad-hoc. Además se crearon un conjunto de reportes y gráficos que están disponibles y accesibles desde el servidor de BI.

#### 3.4.1 Vistas de análisis

Las vistas de análisis se pueden considerar como el elemento más fuerte para el análisis de la información. Permiten detallar y agrupar la información con una facilidad espontánea y se sustenta de las jerarquías y niveles definidos en cubos de OLAP. Además permiten graficar y exportar a Excel.



The screenshot shows a software interface for an analysis view. The title bar reads 'Nueva vista de aná.'. Below the title bar is a toolbar with various icons for navigation and actions. The main area displays a table with three columns: 'Trimestre\_mes\_fecha', 'Empresa', and 'Medidas'. The 'Medidas' column is expanded to show 'Valor\_cif'. The table data is as follows:

Trimestre_mes_fecha	Empresa	Medidas
Tiempo	Empresa	41.435.662
+ TRIM I	+ Empresa	8.724.043
+ TRIM II	+ Empresa	2.959.453
+ abril	+ Empresa	1.931.281
2010-04-29	+ Empresa	1.931.281
	EMPRESA DE ASISTENCIA Y SERVICIOS	1.754.984
	MOVITEL, S.A.	176.297
+ junio	+ Empresa	927.993
+ mayo	+ Empresa	100.179
+ TRIM III	+ Empresa	29.752.166

Figura 11 Vista de análisis.

#### .Cubos OLAP

Los cubos OLAP se confeccionaron en un archivo XML<sup>7</sup> con el uso de la

<sup>7</sup> XML es un lenguaje de etiquetado extensible muy simple que sirve para estructurar, almacenar e

herramienta Schema-Workbench. Cada cubo OLAP representa una tabla de hecho, sus medidas y las dimensiones con que se relaciona. Cada dimensión dentro del cubo se compone de jerarquías y niveles. Los niveles responden a las características de las dimensiones mientras que las jerarquías son agrupaciones de niveles. Una misma dimensión puede poseer varias jerarquías.



*Figura 12 Cubos OLAP.*

### 3.4.2 Reportes ad-hoc

El servidor de BI de pentaho 3.6 facilita la creación de reportes ad-hoc con una interfaz entendible y manejable. De esta manera el usuario final del producto puede

---

intercambiar información.

realizar disímiles consultas y reportes que no se hayan implementado en la solución propuesta. Además, para realizar reportes ad-hoc solo se requiere un mínimo de conocimientos acerca de tablas, relaciones y consultas sobre bases de datos.

### *Indicadores de empresa*

Trimestre	Empresa nombre	Valor fob
TRIM I	ETECSA	46.295
TRIM I	ENERGOIMPORT	537.484
TRIM II	ENERGAS S.A.	226.846
TRIM III	MCV COMERCIAL, S.A.	61.565
TRIM III	GCATE S.A.	1.100.176
TRIM II	ENERGAS S.A.	46.827
TRIM III	MCV COMERCIAL, S.A.	288.881
TRIM I	EMIDICT	73.917
TRIM II	ENERGOIMPORT	11.154
TRIM II	ENERGOIMPORT	1.561.624
TRIM III	ENERGAS S.A.	26.875.405
TRIM II	MCV COMERCIAL, S.A.	288.881
TRIM III	IBF IMPORT - EXPORT	10.362
TRIM I	ETECSA	37.777
TRIM II	ENERGOIMPORT	52.726
TRIM I	ETECSA	32.406
TRIM II	MOVITEL, S.A.	146.848
TRIM III	ETECSA	88.146
TRIM I	CORPORACION COPEXTEL, S.A.	97.914
TRIM I	CLUPET / SHERRIT GORDON	3.412.587
TRIM I	ETECSA	61.638
TRIM III	MOVITEL, S.A.	26.101
TRIM II	EMPRESA DE ASISTENCIA Y SERVICIOS	1.754.984
TRIM III	ENERGOIMPORT	890.565
TRIM III	ETECSA	138.885
TRIM I	EMPRESA COMC PARA LOS PROD. COMBIOMED	37.956

mayo 04, 2012

**Figura 13 Reporte ad-hoc.**

### **.Metadatos**

Para el correcto funcionamiento del diseñador de reportes ad-hoc se crearon un conjunto de metadatos. Dichos metadatos responden a la manera en que el usuario percibe la información, como la desea nombrar o agrupar según los criterios y necesidades del MINCEX. De esta manera, se elimina la frontera entre el lenguaje de los desarrolladores del sistema y el de los usuarios finales para el manejo y entendimiento de los datos.

Los metadatos son elaborados con la ayuda del editor de metadatos de pentaho.

## Capítulo 3

Su estructura se divide en tres partes principales, tablas del negocio, relaciones y vistas del negocio. Un mismo archivo de metadatos puede ser elaborado sobre varios idiomas.

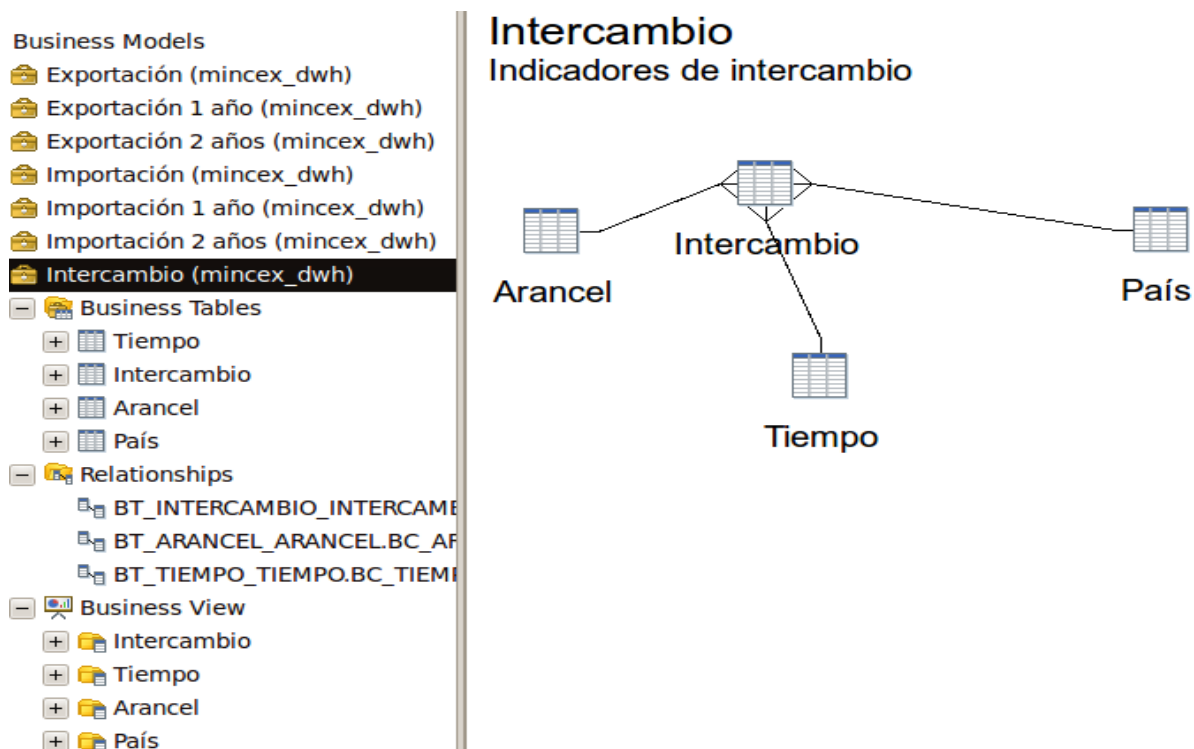


Figura 14 Metadatos.

### 3.4.3 Reportes

Los reportes son documentos preelaborados para la presentación de los datos en el mercado de datos. Garantizan una fuerte interacción usuario-aplicación, de forma tal que un usuario pueda consultar un mismo reporte en momentos diferentes u obtener diferentes resultados de un mismo reporte, en dependencia del pedido realizado.



**Estadística, MINCEX**

**Valor exportado de países por Proveedor y País cliente**

Secreto  
Ejemplar No:  
Hoja No:

**Total Valor Exportado:** 48.039,53

<b>País cliente</b>	<b>Cliente</b>	<b>Valor Exportado</b>
REINO DE BELGICA	BE0181	37,956
REPUBLICA FEDERATIVA DE BRASIL	BR0809	37,777
REPUBLICA FEDERATIVA DE BRASIL	BR0809	61,638
REPUBLICA FEDERATIVA DE BRASIL	BR0809	88,146
REPUBLICA FEDERAL DE ALEMANIA*	CA0785	7.392,494
REPUBLICA FEDERAL DE ALEMANIA*	CA0785	3.412,587
REPUBLICA ARGENTINA	CA1074	226,846
REPUBLICA DOMINICANA	CA1074	46,827
REPUBLICA ARGENTINA	CA1074	655,351
EMIRATOS ARABES UNIDOS	CA1074	26.875,405

**Fecha actual:** 26 abril 2012

*Cifras preliminares emitidas para el MINCEX*  
Fuente: Oficina Nacional de Estadísticas e Información  
CONFECCIONADOS \_\_\_\_\_ HOJAS \_\_\_\_\_  
DISTRIBUCION DE ENVIO: \_\_\_\_\_  
EJECUTOR. MECA: \_\_\_\_\_  
INF. \_\_\_\_\_

**Figura 15 Reporte.**

### 3.4.4 Gráficos

Los gráficos constituyen un elemento primordial en el análisis de la información. Permite el rápido análisis del comportamiento de la información y se fundamenta en el contacto visual con el usuario. El usuario puede diferenciar los cambios en la información suministrada a través de tamaños y colores, con solo un mínimo de esfuerzo mental.

En la solución propuesta los gráficos presentados fueron definidos con antelación, dependiendo de los criterios del cliente. Solo se usaron gráficos de barra y se pueden encontrar en el área de análisis series de tiempo.



Valor exportado por trimestre en el año actual

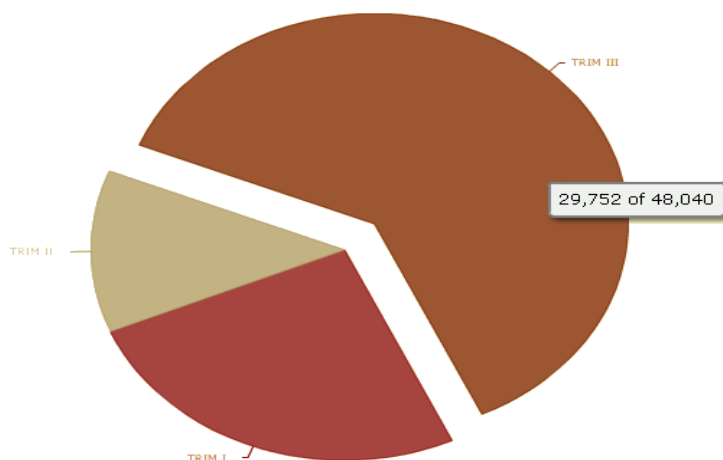


Figura 16 Gráfico.

### 3.4.5 Administración del sistema

La administración se lleva a cabo con la consola de administración del sistema. En dicha aplicación se realizan un conjunto de configuraciones de elementos a tener en cuenta para garantizar la seguridad de la información y facilitar el despliegue de la aplicación o la emigración de la información hacia diferentes bases de datos.

#### .Conexiones JDNI

Las estructuras de visualización funcionan conectándose directamente a la base de datos o usando el nombre de una conexión JDNI. La conexión JDNI se define en la consola de administración. De esta manera, si se cambia la base de datos de lugar solo se debe configurar los atributos de la conexión y las estructuras de visualización continuarán de manera eficiente.

#### .Usuarios y roles

Otra de las funcionalidades más importante de la consola de administración es la

de establecer usuarios y roles. Los usuarios y roles creados se utilizan en las estructuras del servidor de BI para garantizar o denegar acceso a cada una de ellas, creando así un esquema de seguridad muy eficiente.

### **Conclusiones**

En el desarrollo de este capítulo se brindó una breve explicación de la implementación del mercado de datos y los subsistemas de integración y visualización. Se explicaron algunos elementos correspondientes a la administración de la solución propuesta.

### Capítulo 4. “Validación de la solución propuesta”.

#### Introducción

En este capítulo se describen las pruebas realizadas al mercado de datos mediante casos de pruebas, garantizando así u mejor procesamiento y análisis de los datos.

#### 4.1 Pruebas de software

El proceso de pruebas constituye el instrumento adecuado para determinar la calidad de un producto de software. En este proceso se ejecutan pruebas dirigidas a componentes del software o al sistema de software en su totalidad, con el objetivo de medir el grado en el que este cumple con los requerimientos. Además se puede utilizar este para depurar posibles errores que se hayan cometido en la implementación del software.

#### 4.2 Pruebas funcionales basadas en casos de uso

Para la validación del mercado de datos se diseñaron 14 casos de prueba que se corresponden con los casos de uso definidos en el negocio. A continuación se presentan los casos de prueba más significativos.

<i>&lt;Nombre del caso de prueba&gt;</i>	
<b>Escenario:</b> <i>&lt;Nombre del escenario&gt;</i>	
<b>Subescenario</b>	<i>&lt;Nombre del subescenario (solo en caso de que el caso de uso tenga subescenarios )&gt;</i>
<b>Descripción</b>	<i>&lt;Breve descripción del subescenario o escenario&gt;</i>
<b>Perfiles de análisis</b>	<i>&lt;Dimensiones que permitirán el análisis del comportamiento de los indicadores a medir&gt;</i>
<b>Indicadores a medir</b>	<i>&lt;Medidas que serán analizadas en el escenario o subescenario&gt;</i>

## Capítulo 4

<b>Flujo central</b>	<Pasos a seguir para realizar una petición al sistema>
<b>Respuesta del sistema</b>	<Breve descripción de la respuesta que ofrece el sistema para la petición realizada>
<b>Resultados obtenidos</b>	<Satisfactorios o no satisfactorios>

**Tabla 1 Plantilla para casos de prueba.**

<b>Visualizar los indicadores del boletín</b>	
<b>Escenario:</b> Visualizar reportes.	
<b>Subescenario</b>	Visualizar reportes de exportación.
<b>Descripción</b>	Muestra los reportes del boletín para el proceso de exportación de bienes materiales.
<b>Perfiles de análisis</b>	Tiempo, régimen definitivo, arancel, proveedores y clientes, país, empresa.
<b>Indicadores a medir</b>	Peso bruto, valor exportado.
<b>Flujo central</b>	Se autentifica y selecciona la carpeta de exportación en el lateral izquierdo superior de la aplicación, siguiendo el siguiente esquema de navegación: boletín/reportes/exportación. Se selecciona el reporte deseado en el lateral izquierdo inferior.
<b>Respuesta del sistema</b>	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.
<b>Resultados obtenidos</b>	Satisfactorios.
<b>Subescenario</b>	Visualizar reportes de importación.
<b>Descripción</b>	Muestra los reportes del boletín para el proceso de importación de bienes materiales.
<b>Perfiles de análisis</b>	Tiempo, régimen definitivo, arancel, proveedores y clientes, país, empresa, transporte.
<b>Indicadores a medir</b>	Peso bruto, valor importado.

## Capítulo 4

<b>Flujo central</b>	Se autentifica y selecciona la carpeta de importación en el lateral izquierdo superior de la aplicación, siguiendo el siguiente esquema de navegación: boletín/reportes/importación. Se selecciona el reporte deseado en el lateral izquierdo inferior.
<b>Respuesta del sistema</b>	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.
<b>Resultados obtenidos</b>	Satisfactorios.
<b>Subescenario</b>	Visualizar reportes de intercambio comercial.
<b>Descripción</b>	Muestra los reportes del boletín para el proceso de intercambio comercial de bienes materiales.
<b>Perfiles de análisis</b>	Tiempo, régimen definitivo, arancel, proveedores y clientes, país, empresa, transporte.
<b>Indicadores a medir</b>	Valor intercambio.
<b>Flujo central</b>	Se autentifica y selecciona la carpeta de importación en el lateral izquierdo superior de la aplicación, siguiendo el siguiente esquema de navegación: boletín/reportes/intercambio. Se selecciona el reporte deseado en el lateral izquierdo inferior.
<b>Respuesta del sistema</b>	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.
<b>Resultados obtenidos</b>	Satisfactorios.
<b>Escenario:</b> Visualizar vistas de análisis.	
<b>Subescenario</b>	Visualizar vistas de análisis de exportación.
<b>Descripción</b>	Muestra las vistas del boletín para el proceso de exportación de bienes materiales.
<b>Perfiles de análisis</b>	Tiempo, régimen definitivo, arancel, proveedores y clientes, país,

## Capítulo 4

	empresa.
<b>Indicadores a medir</b>	Peso bruto, valor exportado.
<b>Flujo central</b>	Se autentifica y selecciona la carpeta de exportación en el lateral izquierdo superior de la aplicación, siguiendo el siguiente esquema de navegación: boletín/vistas de análisis/exportación. Se selecciona la vista deseada en el lateral izquierdo inferior.
<b>Respuesta del sistema</b>	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.
<b>Resultados obtenidos</b>	Satisfactorios.
<b>Subescenario</b>	Visualizar vistas de análisis de importación.
<b>Descripción</b>	Muestra las vistas del boletín para el proceso de importación de bienes materiales.
<b>Perfiles de análisis</b>	Tiempo, régimen definitivo, arancel, proveedores y clientes, país, empresa, transporte.
<b>Indicadores a medir</b>	Peso bruto, valor importado.
<b>Flujo central</b>	Se autentifica y selecciona la carpeta de importación en el lateral izquierdo superior de la aplicación, siguiendo el siguiente esquema de navegación: boletín/vistas de análisis/importación. Se selecciona la vista deseada en el lateral izquierdo inferior.
<b>Respuesta del sistema</b>	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.
<b>Resultados obtenidos</b>	Satisfactorios.
<b>Subescenario</b>	Visualizar vistas de análisis de intercambio comercial.
<b>Descripción</b>	Muestra las vistas del boletín para el proceso de intercambio comercial

## Capítulo 4

	de bienes materiales.
<b>Perfiles de análisis</b>	Tiempo, arancel, país.
<b>Indicadores a medir</b>	Valor intercambio.
<b>Flujo central</b>	Se autentifica y selecciona la carpeta de intercambio en el lateral izquierdo superior de la aplicación, siguiendo el siguiente esquema de navegación: boletín/vistas de análisis/intercambio. Se selecciona la vista deseada en el lateral izquierdo inferior.
<b>Respuesta del sistema</b>	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.
<b>Resultados obtenidos</b>	Satisfactorios.
<b>Escenario:</b> Visualizar gráficos.	
<b>Subescenario</b>	Visualizar gráficos de exportación
<b>Descripción</b>	Muestra los gráficos del proceso de exportación de bienes materiales en el boletín.
<b>Perfiles de análisis</b>	Tiempo, régimen definitivo, arancel, proveedores y clientes, empresa, país.
<b>Indicadores a medir</b>	Valor exportado
<b>Flujo central</b>	Se autentifica y selecciona la carpeta de exportación en el lateral izquierdo superior de la aplicación, siguiendo el siguiente esquema de navegación: boletín/gráficos/exportación. Se selecciona el gráfico deseado en el lateral inferior izquierdo.
<b>Respuesta del sistema</b>	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.
<b>Resultados obtenidos</b>	Satisfactorios.
<b>Subescenario</b>	Visualizar gráficos de importación

## Capítulo 4

<b>Descripción</b>	Muestra los gráficos del proceso de importación de bienes materiales en el boletín.
<b>Perfiles de análisis</b>	Tiempo, régimen definitivo, arancel, proveedores y clientes, empresa, país, condición de pago, unidad de medida
<b>Indicadores a medir</b>	Valor importado, flete, seguro
<b>Flujo central</b>	Se autentifica y selecciona la carpeta de importación en el lateral izquierdo superior de la aplicación, siguiendo el siguiente esquema de navegación: boletín/gráficos/importación. Se selecciona el gráfico deseado en el lateral inferior izquierdo.
<b>Respuesta del sistema</b>	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.
<b>Resultados obtenidos</b>	Satisfactorios.
<b>Subescenario</b>	Visualizar gráficos de intercambio comercial
<b>Descripción</b>	Muestra los gráficos del proceso de intercambio comercial de bienes materiales en el boletín.
<b>Perfiles de análisis</b>	Tiempo, régimen definitivo, arancel, proveedores y clientes, empresa, país, condición de pago, unidad de medida
<b>Indicadores a medir</b>	Valor importado, valor exportado, valor intercambio, balanza comercial.
<b>Flujo central</b>	Se autentifica y selecciona la carpeta de intercambio comercial en el lateral izquierdo superior de la aplicación, siguiendo el siguiente esquema de navegación: boletín/gráficos/intercambio. Se selecciona el gráfico deseado en el lateral izquierdo inferior.
<b>Respuesta del sistema</b>	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.
<b>Resultados obtenidos</b>	Satisfactorios.

*Tabla 2 Caso de prueba "Visualizar los indicadores del boletín".*



## Capítulo 4

<b>Visualizar los indicadores de importación</b>	
<b>Escenario:</b> Visualizar reportes.	
<b>Descripción</b>	Muestra los reportes del proceso de importación de bienes materiales.
<b>Perfiles de análisis</b>	Tiempo, régimen definitivo, transporte, condición de pago, arancel, proveedores y clientes, empresa, país, unidad de medida, condición de entrega.
<b>Indicadores a medir</b>	Valor importado, flete, seguro, cantidad de artículos
<b>Flujo central</b>	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.
<b>Respuesta del sistema</b>	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.
<b>Resultados obtenidos</b>	Satisfactorios.

*Tabla 3 Caso de prueba “Visualizar los indicadores de importación”.*

<b>Visualizar los indicadores de exportación</b>	
<b>Escenario:</b> Visualizar reportes.	
<b>Descripción</b>	Muestra los reportes del proceso de exportación de bienes materiales.
<b>Perfiles de análisis</b>	Tiempo, régimen definitivo, condición de pago, arancel, proveedores y clientes, empresa, país, unidad de medida, condición de entrega.
<b>Indicadores a medir</b>	Valor exportado, cantidad de artículos
<b>Flujo central</b>	Se autentifica y selecciona la carpeta de reportes en el lateral izquierdo superior de la aplicación, siguiendo el siguiente

## Capítulo 4

	esquema de navegación: exportación/reportes. Se selecciona el reporte deseado en el lateral izquierdo inferior.
<b>Respuesta del sistema</b>	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.
<b>Resultados obtenidos</b>	Satisfactorios.

**Tabla 4 Caso de prueba “Visualizar los indicadores de exportación”.**

<b>Visualizar los indicadores de intercambio comercial</b>	
<b>Escenario:</b> Visualizar reportes.	
<b>Descripción</b>	Muestra los reportes del proceso de intercambio comercial de bienes materiales.
<b>Perfiles de análisis</b>	Tiempo, arancel, país, unidad de medida.
<b>Indicadores a medir</b>	Valor del intercambio.
<b>Flujo central</b>	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.
<b>Respuesta del sistema</b>	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.
<b>Resultados obtenidos</b>	Satisfactorios.

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

<b>Visualizar los indicadores del cuadro de cifras</b>	
<b>Escenario:</b> Visualizar reportes.	
<b>Subescenario</b>	Visualizar vistas de análisis de importación

## Capítulo 4

<b>Descripción</b>	Muestra las vistas de cuadro de cifras del proceso importación de bienes materiales.
<b>Perfiles de análisis</b>	Tiempo, régimen.
<b>Indicadores a medir</b>	Cantidad de artículos, valor importado.
<b>Flujo central</b>	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.
<b>Respuesta del sistema</b>	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.
<b>Resultados obtenidos</b>	Satisfactorios.
<b>Subescenario</b>	Visualizar vistas de análisis de exportación
<b>Descripción</b>	Muestra las vistas de cuadro de cifras del proceso exportación de bienes materiales.
<b>Perfiles de análisis</b>	Tiempo, régimen.
<b>Indicadores a medir</b>	Cantidad de artículos, valor exportado.
<b>Flujo central</b>	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.
<b>Respuesta del sistema</b>	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.
<b>Resultados obtenidos</b>	Satisfactorios.

*Tabla 6 Caso de prueba "Visualizar los indicadores del cuadro de cifras".*

## Capítulo 4

<b>Visualizar los indicadores de transporte</b>	
<b>Escenario:</b> Visualizar reportes.	
<b>Subescenario</b>	Visualizar reportes de exportación
<b>Descripción</b>	Muestra los reportes del transporte de bienes materiales en el proceso de exportación.
<b>Perfiles de análisis</b>	Tiempo, régimen definitivo, transporte, arancel, país.
<b>Indicadores a medir</b>	Valor exportado, peso bruto, flete, seguro.
<b>Flujo central</b>	Se autentifica y selecciona la carpeta de exportación en el lateral izquierdo superior de la aplicación, siguiendo el siguiente esquema de navegación: transporte/reportes/exportación. Se selecciona el reporte deseado en el lateral izquierdo inferior.
<b>Respuesta del sistema</b>	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.
<b>Resultados obtenidos</b>	Satisfactorios.
<b>Subescenario</b>	Visualizar reportes de importación
<b>Descripción</b>	Muestra los reportes del transporte de bienes materiales en el proceso de importación.
<b>Perfiles de análisis</b>	Tiempo, régimen definitivo, transporte, arancel, país.
<b>Indicadores a medir</b>	Valor importado, peso bruto, flete, seguro.
<b>Flujo central</b>	Se autentifica y selecciona la carpeta de importación en el lateral izquierdo superior de la aplicación, siguiendo el siguiente esquema de navegación: transporte/reportes/importación. Se selecciona el reporte deseado en el lateral izquierdo inferior.
<b>Respuesta del sistema</b>	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

## Capítulo 4

	definidos.
<b>Resultados obtenidos</b>	Satisfactorios.

*Tabla 7 Caso de prueba “Visualizar los indicadores de transporte”.*

<b>Visualizar los indicadores de proveedores y clientes.</b>	
<b>Escenario:</b> Visualizar reportes.	
<b>Subescenario</b>	Visualizar reportes de exportación
<b>Descripción</b>	Muestra los reportes del comercio entre clientes y proveedores para el proceso de exportación de bienes materiales.
<b>Perfiles de análisis</b>	Tiempo, régimen definitivo, arancel, país, cliente/proveedor, empresa.
<b>Indicadores a medir</b>	Cantidad de artículos, valor exportado.
<b>Flujo central</b>	Se autentifica y selecciona la carpeta de exportación en el lateral izquierdo superior de la aplicación, siguiendo el siguiente esquema de navegación: proveedores y clientes/reportes/exportación. Se selecciona el reporte deseado en el lateral izquierdo inferior.
<b>Respuesta del sistema</b>	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.
<b>Resultados obtenidos</b>	Satisfactorios.
<b>Subescenario</b>	Visualizar reportes de importación
<b>Descripción</b>	Muestra los reportes del comercio entre clientes y proveedores para el proceso de importación de bienes materiales.
<b>Perfiles de análisis</b>	Tiempo, régimen definitivo, arancel, país, cliente/proveedor, empresa.
<b>Indicadores a medir</b>	Cantidad de artículos, valor importado.
<b>Flujo central</b>	Se autentifica y selecciona la carpeta de importación en el lateral izquierdo superior de la aplicación, siguiendo el siguiente esquema de navegación: proveedores y clientes/reportes/importación. Se selecciona el reporte deseado en el lateral izquierdo inferior.

## Capítulo 4

<b>Respuesta del sistema</b>	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.
<b>Resultados obtenidos</b>	Satisfactorios.

*Tabla 8 Caso de prueba “Visualizar los indicadores de proveedores y clientes”.*

<b>Visualizar errores en las cifras de bienes</b>	
<b>Escenario:</b> Visualizar vistas de análisis.	
<b>Subescenario</b>	Visualizar vistas de análisis de importación
<b>Descripción</b>	Muestra las vistas de análisis que pudieran reflejar posibles errores en las cifras de bienes materiales.
<b>Perfiles de análisis</b>	Empresa, tiempo, arancel, proveedores y clientes, país, régimen definitivo.
<b>Indicadores a medir</b>	Valor importado.
<b>Flujo central</b>	Se autentifica y selecciona la carpeta de importación en el lateral izquierdo superior de la aplicación, siguiendo el siguiente esquema de navegación: análisis de errores/vistas de análisis/importación. Se selecciona la vista deseada en el lateral izquierdo inferior.
<b>Respuesta del sistema</b>	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.
<b>Resultados obtenidos</b>	Satisfactorios.
<b>Subescenario</b>	Visualizar vistas de análisis de exportación
<b>Descripción</b>	Muestra las vistas de análisis que pudieran reflejar posibles errores en las cifras de bienes.
<b>Perfiles de análisis</b>	Empresa, tiempo, arancel, proveedores y clientes, país, régimen

## Capítulo 4

	definitivo.
<b>Indicadores a medir</b>	Valor exportado.
<b>Flujo central</b>	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.
<b>Respuesta del sistema</b>	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.
<b>Resultados obtenidos</b>	Satisfactorios.

*Tabla 9 Caso de prueba “Visualizar errores en las cifras de bienes”.*

<b>Visualizar los indicadores de series de tiempo</b>	
<b>Escenario:</b> Visualizar reportes.	
<b>Subescenario</b>	Visualizar reportes de exportación
<b>Descripción</b>	Muestra los reportes del proceso de exportación de bienes materiales en series de tiempo.
<b>Perfiles de análisis</b>	Tiempo, régimen definitivo, arancel, proveedores y clientes, empresa, país.
<b>Indicadores a medir</b>	Valor exportado
<b>Flujo central</b>	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.
<b>Respuesta del sistema</b>	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

## Capítulo 4

	definidos.
<b>Resultados obtenidos</b>	Satisfactorios.
<b>Subescenario</b>	Visualizar reportes de importación
<b>Descripción</b>	Muestra los reportes del proceso de importación de bienes materiales en series de tiempo.
<b>Perfiles de análisis</b>	Tiempo, régimen definitivo, arancel, proveedores y clientes, empresa, país, condición de pago, unidad de medida
<b>Indicadores a medir</b>	Valor importado, flete, seguro
<b>Flujo central</b>	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.
<b>Respuesta del sistema</b>	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.
<b>Resultados obtenidos</b>	Satisfactorios.
<b>Subescenario</b>	Visualizar reportes de intercambio comercial
<b>Descripción</b>	Muestra los reportes del proceso de intercambio comercial de bienes materiales en series de tiempo.
<b>Perfiles de análisis</b>	Tiempo, régimen definitivo, arancel, proveedores y clientes, empresa, país, condición de pago, unidad de medida.
<b>Indicadores a medir</b>	Valor importado, valor exportado, valor intercambio, balanza comercial.
<b>Flujo central</b>	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.
<b>Respuesta del sistema</b>	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.
<b>Resultados obtenidos</b>	Satisfactorios.

*Tabla 10 Caso de prueba "Visualizar los indicadores de series de tiempo".*

Para todos los casos de prueba funcionales basados en casos de uso aplicados los resultados fueron satisfactorios.

### **4.3 Pruebas de integración.**

Se insertaron un número finito de datos a través del proceso de ETL. Se comprobó que el proceso valida todas las reglas del negocio y es capaz de detectar errores de consistencia en la información suministrada por la AGR, tales como llaves huérfanas o valores incorrectos. Por tanto, el resultado arrojado fue satisfactorio.

### **4.4 Prueba de las estructuras multidimensionales y metadatos.**

Se verificó que las estructuras construidas (cubos OLAP y metadatos para reportes ad-hoc) respondan a todo el alcance de información que fue definido en el negocio. Las pruebas realizadas arrojaron resultados satisfactorios, dichas estructuras responden a las necesidades de información y se rigen por el lenguaje definido por el cliente.

### **4.5 Resultados y funcionalidades obtenidas**

El desarrollo de la propuesta de solución se llevó a cabo en cuatro fases. Durante el transcurso de cada fase se realizaron un conjunto de actividades; tales como elaboración de artefactos, capacitación e investigación y estudio de temas relacionados con el desarrollo de almacenes de datos y el proceso de análisis estadístico en general. Cada fase culminada arrojó un resultado práctico o la implementación de una nueva funcionalidad del 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 diseñaron y describieron los casos de uso del sistema. Se establecieron cada uno de los requerimientos funcionales, no funcionales e informativos. 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 cifras de estadísticas de bienes materiales.

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

### **Conclusiones**

En el desarrollo de este capítulo se definen los elementos de validación del mercado de datos. El sistema fue validado mediante casos de prueba y todos arrojaron resultados satisfactorios.

# Conclusiones generales

---

## Conclusiones generales

Concluida la investigación y el desarrollo del sistema y cumplidos a su vez los objetivos planteados así como las tareas de investigación pueden plantearse los resultados alcanzados:

- ▶ Se realizó un estudio detallado sobre los artefactos que se construyen en el desarrollo del mercado de datos.
- ▶ Se desarrolló la primera iteración del mercado de datos estadísticas de bienes para el departamento de análisis estadístico del MINCEX.
- ▶ Se obtuvieron mejoras en el procesamiento de los datos, los cuales son tratados en un proceso de ETL con una herramienta de avanzada que realiza la validación de forma automática y es capaz de detectar y corregir errores garantizando que los datos sean consistentes y propiciando un almacenamiento seguro e histórico.
- ▶ Se fortalece el análisis de los datos debido a que son consistentes, confiables, su presentación no requiere de manipulación humana y se realiza con diferentes estructuras de visualización que incluye el uso de gráficos.
- ▶ Las estadísticas de bienes materiales pueden analizarse desde cualquier ordenador siempre y cuando el usuario cuente con los privilegios necesarios para acceder a la información suministrada.

## Recomendaciones

---

### **Recomendaciones**

Se recomienda la implementación de modelos predictivos para el mercado de datos desarrollado. Los modelos predictivos permitirán trazar nuevas estrategias para el comercio de bienes materiales en Cuba. Es imprescindible además, comenzar el desarrollo de la segunda iteración del sistema implementado y enriquecerlo con el uso de juegos de gráficos y plantillas de reportes atractivos e interesantes.

Se recomienda además, la integración de los mercados de datos estadística de bienes y estadística de servicios con el objetivo de consolidar el almacén de datos para las estadísticas del MINCEX.

## Glosario de términos

---

### Glosario de términos

**SQL:** Lenguaje estructurado de consulta

**BI:** Inteligencia de negocio

**ONE:** Oficina Nacional de Estadísticas

**MINCEX:** Ministerio del Comercio Exterior y la Inversión Extranjera

**JDBC:** Conector de java para base de datos

**UML:** Lenguaje unificado de modelado

**PC:** Computadora personal

**DML:** Lenguaje de manipulación de datos

**AA:** Áreas de análisis

**MD:** Mercado de datos

**PRD:** Diseñador de reportes de Pentaho

**CIMEX:** Corporación de Importaciones y Exportaciones

**OLTP:** Procesamiento de transacciones en línea

# Bibliografía

---

## Bibliografía

**Almacenes de datos (Datawarehouse)** [En línea]. Disponible en:

<http://www.rhernando.net/modules/tutorials/doc/bd/dw.html>

**Balteus. Open System Architect, 2011.** Disponible en:

<http://balteus.blogspot.com/2011/04/powerdesigner-free-alternatives-or.html>

**Carrandi Hernández, Yenisley;Guerrero Barea, Adnier.** Módulo Almacén del Sistema de Información Hospitalaria alas HIS .UCI.Ciudad Habana, 2009.

**Case.** Herramientas de Modelado, 2011. Disponible en:

<http://www.cyta.com.ar/biblioteca/bddoc/bdlibros/proyectoinformatico/libro/c5/c5.htm>

**Cavsi.** ¿Qué es un Sistema Gestor de Bases de Datos o SGBD?, 2011. Disponible en: <http://www.cavsi.com/preguntasrespuestas/que-es-un-sistema-gestor-de-bases-de-datos-o-sgbd/>

**Ciberaula.**OpenOffice.org La suite ofimática Libre, 2010. Disponible en:

<http://linux.ciberaula.com/articulo/openoffice/>

**Contreras Martinez, Yaneisy.** Diseño e Implementación de un almacén de datos para la red nacional de Genética Médica .UCI.Ciudad Habana, 2010.

**Dataprix.**Tipos de aplicaciones donde utilizar técnicas de Data Warehouse, 2009.

Disponible en:<http://www.dataprix.com/book/export/html/90>

**Díaz Morales, Yoel.** Almacén de datos estadísticos de la ONE: Desarrollo de la capa de visualización del mercado de datos demografía.UCI. Ciudad Habana, 2010.

**Díaz Morales, Yoel.** Almacén de datos estadísticos de la ONE: Desarrollo de la capa de visualización del mercado de datos demografía.UCI. Ciudad Habana, 2010.Pág 9

**Disytel.** Jaspersoft, 2011. Disponible en:

## Bibliografía

---

<http://www.disytel.com/capacitacion/cursos-jaspersoft-bi>

**DosIdeas.** Apache Derby, 2009. Disponible en:

[http://www.dosideas.com/wiki/Apache\\_Derby](http://www.dosideas.com/wiki/Apache_Derby)

**Escobar Domínguez, René Raydel.** Desarrollo de un Almacén de Datos para el Control del Consumo de Portadores Energéticos en la Oficina Nacional de Estadística. UCI.Ciudad Habana, 2010.

**Espinosa, Roberto** .Kimball vs Inmon. Ampliación de conceptos del Modelado Dimensional, 2010. Disponible en: <http://churriwifi.wordpress.com/2010/04/19/15-2-ampliacion-conceptos-del-modelado-dimencional/>

**Gestión.** Que es un almacén de datos?, 2011. Disponible en:

<http://www.gestion.org/documental/gestion-del-conocimiento/4891/que-es-data-warehouse.html>

**González Ochoa, Darián.** Diseño e Implementación de un Almacén de Datos Operacionales para la Corporación CIMEX. UCI. Ciudad Habana, 2010.

**Gopac.** ¿Qué es Inteligencia de Negocios?, 2007. Disponible en:

<http://www.gopac.com.mx/v3/gopacbi/quees.asp>

**Johan.** Pentaho Data Integration (Integracion de Datos Pentaho), 2011. Disponible en: <http://www.inqbation.com/pentaho-data-integration/>

**Kimball, Ralph.**The Datawarehouse Toolkit Olap Solutions.New York, Wiley Computer Publishing, 2002.Pág 29

**Kioskea.** Bases de datos, 2011. Disponible en:

<http://es.kioskea.net/contents/bdd/bddintro.php3>

**Linalis.**Pentaho BI Suite, 2011. Disponible en:

<http://www.linalis.com/home/pentaho-bi-suite-bootcamp-es>

**Microsoft.** Tablas de dimensiones, 2011. Disponible en:

<http://msdn.microsoft.com/es-es/library/ms244714%28v=vs.80%29.aspx>

# Bibliografía

---

**MySQL**.Workbench, 2011.Disponible en:

<http://www.mysql.com/products/workbench/>

**Oporto, Samuel**. Modelo Dimensional, 2008. Disponible en:

[http://www.wiphala.net/courses/dwh/.../class\\_32\\_physical\\_model.ppt](http://www.wiphala.net/courses/dwh/.../class_32_physical_model.ppt)

**Oracle Corporation**, MySQL Technical Specifications, 2011 Disponible en:

<http://www.mysql.com/>

**Ortiz Sierra, Julio Ernesto**. Diseño e Implementación de un Mercado de Datos para la Oficina Nacional de Estadísticas. UCI. Ciudad Habana, 2009.

**Pentaho**. Business Intelligence Open Source: Fundamentos de PentahoMetadata Editor, 2009. Disponible en: <http://www.demosdesoftware.com/videos/65/business-intelligence/pentaho-business-intelligence-open-source-fundamentos-de-pentaho-metadata-editor/business-intelligence-codigo-abierto-open-source-pentaho>

<http://www.demosdesoftware.com/videos/65/business-intelligence/pentaho-business-intelligence-open-source-fundamentos-de-pentaho-metadata-editor/business-intelligence-codigo-abierto-open-source-pentaho>

**Pérez Valdés, Damián**. ¿Qué son las bases de datos?, 2008. Disponible en:

<http://www.maestrosdelweb.com/principiantes/%C2%BFque-son-las-bases-de-datos>

**PostgreSQL Global Development Group**, PostgreSQL.2011 Disponible en:

<http://www.postgresql.org/about/>

**Proyecto datawarehousing/BI**. Almacenando Datos, 2010. Disponible en:

<http://proyectopentahodw.wordpress.com/2010/05/04/modelo-dimENSIONAL/>

**Sap**. Herramientas de Business Objects de visualización avanzada, 2011. Disponible en:

<http://www.sap.com/spain/solutions/sapbusinessobjects/large/intelligenceplatform/bi/dashboard-visualization/advanced-visualization/index.epx>

**SIE**. Sistema de Información Estadística, 2008.Disponible en:

[http://www.jcyl.es/sie/sas/broker?\\_PROGRAM=mddbpgm.v2.indexv2.scl&\\_SERVICE=sasweb&\\_DEBUG=0&menu=index](http://www.jcyl.es/sie/sas/broker?_PROGRAM=mddbpgm.v2.indexv2.scl&_SERVICE=sasweb&_DEBUG=0&menu=index)



## Bibliografía

---

**Sinnexus.**Datamining, 2011. Disponible en:

[http://www.sinnexus.com/business\\_intelligence/datamining.aspx](http://www.sinnexus.com/business_intelligence/datamining.aspx)

**Silicon News.** Actualidad Almacenamiento, 2007. Disponible en:

<http://www.siliconnews.es/2007/09/11/almac-n-datos-verde-m-s-grande/>

**SIS KLE.** Bill Inmon, 2009. Disponible en: <http://kle.sisorg.com.mx/articulo02.html>

SIS KLE. Ralph Kimball, 2009. Disponible en:

<http://kle.sisorg.com.mx/articulo03.html>

**SQL PowerArchitect.** SQL PowerArchitect herramienta de modelado de datos, 2011. Disponible en: <http://www.sqlpower.ca/page/architect>

**SQLMaestro.**PostgreSQLCode Factory, 2011. Disponible en:

<http://www.sqlmaestro.com/products/postgresql/codefactory/download/>

**Thompson,Ivan.** Definición de Marketing, 2009. Disponible en:

<http://www.marketing-free.com/articulos/definicion-marketing.html>

**Ubuntu.**PgAdmin\_III, 2008. Disponible en:

[http://www.guia-ubuntu.org/index.php?title=PgAdmin\\_III](http://www.guia-ubuntu.org/index.php?title=PgAdmin_III).

**Valmaseda, Marcos Ortiz.** Lista general de PostgreSQL en español, 2009. Disponible en:

<http://comments.gmane.org/gmane.comp.db.postgresql.spanish/37586>

**Viaclinica.**Hospital episodio estadísticas (HES), 2009. Disponible en:

[http://viaclinica.com/article.php?pmc\\_id=1698491](http://viaclinica.com/article.php?pmc_id=1698491)

**Visual Paradigm.** Boost productivity with Innovative and Intuitive Technologies, 2010. Disponible en: <Http://www.visual-paradigm.com>

XMind Ltd. ¿Que es Xmind?, 2008. Disponible en: <http://www.xmind.net/>

### **Anexo 1: Preguntas de las entrevistas**

1. ¿Cuáles son los objetivos de su organización?
2. ¿Qué se está tratando de resolver? ¿Cuáles son las prioridades de la entidad que deben resolverse?
3. ¿Cómo obtiene habitualmente los datos? ¿Cuál es el proceso de obtención de los mismos? ¿Qué hace con la información una vez que la obtiene?
4. ¿Existen categorías de la información dentro del negocio? ¿Es posible dividir las actividades que se realizan en el negocio por sector, rama, departamento, dirección según el tipo de información que se maneja?
5. ¿Cuál es la información o los tipos de reportes que actualmente se puede obtener en el negocio?
6. ¿Cuáles son las claves del negocio en riesgos actuales? ¿Qué previenes al conocer los objetivos del negocio? ¿Cuál es el impacto de este conocimiento en la organización? ¿Existen aspectos externos que puedan afectar el curso de la información?
7. ¿Con qué frecuencia se obtiene información?
8. ¿Qué cantidad de información en general se maneja?
9. Cantidad de reportes (de cada sistema internos de ellos)
10. Complejidad de los reportes (cruces de variables).
11. ¿Cómo usted sabe que está haciendo bien su trabajo? ¿Cómo usted identifica los problemas o como los soluciona?
12. ¿Cuáles son las métricas para medir el éxito? ¿Cómo obtienes los indicadores (métricas) para monitorizar los procesos?

13. ¿Qué criterios se tienen en cuenta para analizar los indicadores?
14. ¿Qué análisis le gustaría realizar sobre esos indicadores? ¿Existen posibilidades de mejora a su método/ proceso actual?
15. ¿Con qué frecuencia se debe analizar?
16. ¿Existen eventos externos que es interesante analizar sobre esos indicadores (actividades en la universidad, días feriados, etc.)?
17. ¿Existen cortes de información o peticiones más comunes? ¿Qué tipo de análisis a la medida suele realizar con más frecuencia?
18. ¿Cómo le gustaría que se le presentaran los datos en la pantalla? (gráficos, tablas, indicadores en la parte superior, etc.)
19. Describa sus fuentes de datos (u otra clase de dimensión).
20. ¿Cuánto se requiere de información histórica?

### **Almacenes (Para informáticos)**

21. Frecuencia mínima de captación. (Tiempo).
22. Nivel mínimo de detalle de la captación disponible.
23. Indicadores y elementos utilizados para su cálculo.
24. Relaciones jerárquicas entre los elementos captados.
25. ¿Usan clasificadores o nomencladores para los elementos descriptivos identificados? (Cuando los elementos descriptivos sean identificados es necesario clasificarlos para cargar las dimensiones).
26. ¿Cuándo se agrupan los aspectos mantiene su significado? (Validar que los aspectos o medidas incluidas en la tabla de hechos mantienen su significado)

cuando se aplica alguna función de agregación). Un contraejemplo para este caso puede ser el salario promedio, no se pueden sumar promedios.

27. Período de tiempo en que los datos se consideran útiles para los análisis estadísticos. (Buscando definir los diferentes niveles de detalle y el grupo de datos activos que gestionará el repositorio).
28. ¿Qué particiones lógicas se hacen en los análisis? (Busca detectar estrategia para optimizar consultas utilizando consultas paralelas en el proceso de recuperación de datos y reducir tanto como sea posible los cuellos de botella para utilizar los componentes de hardware como pueden ser discos duros y/o infraestructura de comunicaciones).
29. ¿Qué herramientas se utilizan en los análisis? ¿Estás soportan algún estándar de bases de datos? (Importante para integrar la solución con componentes).

### **ETL (Para informáticos)**

30. ¿Cuentan con sistemas operacionales para gestionar la información, qué tipo de estructura utilizan?
31. ¿Cuál es la ubicación física de dicha estructura?
32. ¿Con qué disponibilidad y garantía de acceso cuentan dichas estructuras de almacenamiento de información? ¿Cómo puede accederse a ellos? ¿Personas con acceso a la información?
33. ¿Resulta disponible toda la información necesaria para el proyecto?
34. Descripción detallada de las estructuras utilizadas en el almacenamiento.
35. ¿En qué forma brindará la información sobre los datos almacenados en su empresa?

36. ¿Cuáles son los horarios que puede brindar la mayor cantidad de información sin afectar su funcionamiento?
37. ¿Existe algún tipo de codificación en los datos que requiera modificación previa para hacerlo entendible? (datos encriptados, en otro idioma que implique una traducción, etc.).
38. ¿Qué hacer con los campos identificados con errores, a quién contactar en estos casos para definir la decisión a tomar?
39. ¿Qué hacer en caso de identificar valores inexistentes o nulos, a quién contactar en estos casos para definir el tratamiento que se deberá llevar a cabo (No cargarlos, calcularlos, asignar valores teniendo en cuenta un determinado criterio, etc.)?
40. ¿Con qué frecuencia la empresa genera información?
41. ¿Desde qué fecha la empresa contiene la información digitalizada?

### **BI (Para clientes y usuarios)**

42. ¿A qué tipo de personas está dirigido este producto?
43. ¿Qué espera el usuario del producto, aun cuando no tenga definido la necesidad de información?
44. Grupos de usuario o Roles
45. ¿Todos los trabajadores acceden a la misma información?
46. ¿Cuáles son los reportes de interés?
47. ¿Tienen definidas reglas de navegación?

## Anexo 2: Minutas de reunión

Minuta de reunión #1			
<b>Autor</b>	Yamila Monteagudo Sánchez	<b>Fecha</b>	3/10/2011
<b>Lugar</b>	Departamento de análisis estadístico	<b>Hora inicio</b>	2:00 pm
<b>Proyecto/Grupo</b>	Mercado de datos estadísticas de bienes para el departamento de análisis estadístico del MINCEX	<b>Hora terminación</b>	3:00 pm
<b>Asunto</b>	Reportes de bienes		
<b>Asistentes</b>	Elena Morales y Zaili Ortega		
<b>Ausentes</b>	-		
<b>Actividades</b>	La especialista nos muestra los posibles reportes y salidas de la información.		
<b>Acuerdo</b>	Entrega un documento con los posibles reportes.		
<b>Tarea</b>	<b>Responsable</b>	<b>Fecha de cumplimiento</b>	
	Mario	5/10/2011	

Minuta de reunión #2			
<b>Autor</b>	Yamila Monteagudo Sánchez	<b>Fecha</b>	6/10/2011
<b>Lugar</b>	Departamento de análisis estadístico	<b>Hora inicio</b>	9:00 am
<b>Proyecto/Grupo</b>	Mercado de datos estadísticas de bienes para el departamento de análisis estadístico del MINCEX	<b>Hora terminación</b>	10:30 am
<b>Asunto</b>	Levantamiento de información		
<b>Asistentes</b>	Elena Morales y Zaili Ortega		
<b>Ausentes</b>	-		
<b>Actividades</b>	La especialista nos explica cuáles son sus salidas mensuales y trimestrales (Boletín).		

## Anexos

<b>Acuerdo</b>	Entregarnos un documento con sus cierres mensuales y trimestrales (Boletín).	
<b>Tarea</b>	<b>Responsable</b>	<b>Fecha de cumplimiento</b>
	Elena Morales	En la semana.

Minuta de reunión #3			
<b>Autor</b>	Yamila Monteagudo Sánchez	<b>Fecha</b>	11/10/2011
<b>Lugar</b>	Departamento de análisis estadístico	<b>Hora inicio</b>	2:00 pm
<b>Proyecto/Grupo</b>	Mercado de datos estadísticas de bienes para el departamento de análisis estadístico del MINCEX	<b>Hora terminación</b>	3:00 pm
<b>Asunto</b>	Reportes de bienes		
<b>Asistentes</b>	Elena Morales y Zaili Ortega		
<b>Ausentes</b>	-		
<b>Actividades</b>	La especialista revisa los posibles reportes.		
<b>Acuerdo</b>	Entrega un documento con los posibles reportes.		
<b>Tarea</b>	<b>Responsable</b>	<b>Fecha de cumplimiento</b>	
	Yamila	13/09/2011	

Minuta de reunión #4			
<b>Autor</b>	Yamila Monteagudo Sánchez	<b>Fecha</b>	14/10/2011
<b>Lugar</b>	Departamento de análisis estadístico	<b>Hora inicio</b>	9:00 am
<b>Proyecto/Grupo</b>	Mercado de datos estadísticas de bienes para el departamento de análisis estadístico del MINCEX	<b>Hora terminación</b>	10:30 am
<b>Asunto</b>	Levantamiento de información		

## Anexos

<b>Asistentes</b>	Elena Morales	
<b>Ausentes</b>	-	
<b>Actividades</b>	La especialista nos explica los posibles usuarios a la aplicación.	
<b>Acuerdo</b>	Entregarnos un documento con los posibles usuarios.	
<b>Tarea</b>	<b>Responsable</b>	<b>Fecha de cumplimiento</b>
	Elena Morales	En la semana.

Minuta de reunión #5			
<b>Autor</b>	Yamila Monteagudo Sánchez	<b>Fecha</b>	17/10/2011
<b>Lugar</b>	Departamento de análisis estadístico	<b>Hora inicio</b>	2:00 pm
<b>Proyecto/Grupo</b>	Mercado de datos estadísticas de bienes para el departamento de análisis estadístico del MINCEX	<b>Hora terminación</b>	3:00 pm
<b>Asunto</b>	Reportes de bienes		
<b>Asistentes</b>	Elena Morales		
<b>Ausentes</b>	-		
<b>Actividades</b>	La informática muestra el cuestionario a las especialista		
<b>Acuerdo</b>	Entregar el cuestionario.		
<b>Tarea</b>	<b>Responsable</b>	<b>Fecha de cumplimiento</b>	
	Elena Morales	20/10/2011	

### Anexo 3: Avals de calidad



### Aval de Calidad de Software

El grupo de Calidad de Software del Centro de Desarrollo de la Facultad Regional "Mártires de Artemisa" conformado por:

- **Asesora de Calidad:** Ing. Maidel Ojeda Cruz
- **Asesor de Tecnología:** Ing. Domma Moreno Dager
- **Especialista de Calidad:** Ing. Yenisleidy Rodríguez Martínez

emite el presente **Aval de Calidad de Software** en colaboración con los especialistas del Centro de Desarrollo a: Yamila Monteagudo Sánchez, como resultado satisfactorio de su desempeño en la tareas asociadas al proyecto: Almacén de Datos para el MINCEX.

Para emitir el presente aval se valoraron un conjunto de elementos evaluados de manera individual teniendo en cuenta los parámetros de calidad de software del proyecto. A continuación se presenta los resultados en cada uno de los aspectos valorados:

Elementos evaluados	Resultado
Validar la coincidencia entre requisitos informativos y estructuras de visualización.	Satisfactorio
Validar el correcto funcionamiento de las estructuras de visualización.	Satisfactorio
Validar el cumplimiento de la arquitectura de la información.	Satisfactorio
Validar estructura y funcionamiento de los metadatos y reportes ad-hoc.	Satisfactorio
Validar estructura y funcionamiento de los esquemas de mondrían.	Satisfactorio
Verificar que los reportes brindaran información coherente.	Satisfactorio
Validar el funcionamiento interno de las vistas materializadas.	Satisfactorio
Validar el correcto funcionamiento de los procesos de ETL.	Satisfactorio
Validar el tratamiento de errores en los procesos de ETL.	Satisfactorio
Generación de todos los artefactos de la metodología Kimball.	Satisfactorio
Cumplimiento de las plantillas establecidas para cada artefacto.	Satisfactorio
Ortografía, concordancia y redacción de cada uno de los artefactos.	Satisfactorio
Efectividad de los casos de pruebas definidos en el proyecto	Satisfactorio

  
Ing. Maidel Ojeda Cruz  
Asesora de Calidad de SW



  
Msc. Yamila Vigil Regalado  
Directora Centro de Desarrollo

### Aval de Calidad de Software

El grupo de Calidad de Software del Centro de Desarrollo de la Facultad Regional "Mártires de Artemisa" conformado por:

- **Asesora de Calidad:** Ing. Maidel Ojeda Cruz
- **Asesor de Tecnología:** Ing. Domma Moreno Dager
- **Especialista de Calidad:** Ing. Yenisleidy Rodríguez Martínez

emite el presente **Aval de Calidad de Software** en colaboración con los especialistas del Centro de Desarrollo a: Dianelys Alonso Martínez, como resultado satisfactorio de su desempeño en la tareas asociadas al proyecto: Almacén de Datos para el MINCEX.

Para emitir el presente aval se valoraron un conjunto de elementos evaluados de manera individual teniendo en cuenta los parámetros de calidad de software del proyecto. A continuación se presenta los resultados en cada uno de los aspectos valorados:

Elementos evaluados	Resultado
Validar la coincidencia entre requisitos informativos y estructuras de visualización.	Satisfactorio
Validar el correcto funcionamiento de las estructuras de visualización.	Satisfactorio
Validar el cumplimiento de la arquitectura de la información.	Satisfactorio
Validar estructura y funcionamiento de los metadatos y reportes ad-hoc.	Satisfactorio
Validar estructura y funcionamiento de los esquemas de mondrian.	Satisfactorio
Verificar que los reportes brindaran información coherente.	Satisfactorio
Validar el funcionamiento interno de las vistas materializadas.	Satisfactorio
Validar el correcto funcionamiento de los procesos de ETL.	Satisfactorio
Validar el tratamiento de errores en los procesos de ETL.	Satisfactorio
Generación de todos los artefactos de la metodología Kimball.	Satisfactorio
Cumplimiento de las plantillas establecidas para cada artefacto.	Satisfactorio
Ortografía, concordancia y redacción de cada uno de los artefactos.	Satisfactorio
Efectividad de los casos de pruebas definidos en el proyecto	Satisfactorio

  
Ing. Maidel Ojeda Cruz  
Asesora de Calidad de SW



  
Msc. Yamila Vigil Regalado  
Directora Centro de Desarrollo