

**CREACION DE UNA DATAWAREHOUSE A UNA PYME TENIENDO EN CUENTA EL  
CONCEPTO DE INTELIGENCIA DE NEGOCIOS**

**CARLOS ALBERTO CARDONA VILLALBA  
VIVIANA AREVALO CASTAÑEDA**

**UNIVERSIDAD EAFIT  
ESCUELA DE INGENIERÍA  
DEPARTAMENTO DE INGENIERÍA DE SISTEMAS  
MEDELLÍN  
2012**

**CREACION DE UNA DATAWAREHOUSE A UNA PYME TENIENDO EN CUENTA EL  
CONCEPTO DE INTELIGENCIA DE NEGOCIOS**

**CARLOS ALBERTO CARDONA VILLALBA  
VIVIANA AREVALO CASTAÑEDA**

**TRABAJO DE GRADO**

**ASESOR:  
Ingeniera Maribel Bayer Mejía  
Departamento de Ingeniería de Sistemas  
Universidad EAFIT**

**UNIVERSIDAD EAFIT  
ESCUELA DE INGENIERÍA  
DEPARTAMENTO DE INGENIERÍA DE SISTEMAS  
MEDELLÍN  
2012**

**Nota de aceptación:**

---

---

---

---

---

---

---

---

**Presidente del jurado**

---

**Jurado**

---

**Jurado**

**Medellín, Mayo 15 de 2012**

## **AGRADECIMIENTOS**

Expresamos nuestros agradecimientos en primera instancia a Dios, quien hizo posible este logro en nuestras vidas, dándonos la sabiduría, dirección y confianza, a nuestros padres, por su apoyo incondicional y la confianza depositada, para ayudarnos a alcanzar nuestros sueños, a nuestros amigos cercanos, quienes con su apoyo moral nos dieron fuerzas para seguir adelante y animo para no desfallecer.

A todas aquellas personas que hicieron posible la realización de esta investigación, a nuestra asesora Maribel Bayer por su paciencia y apoyo. Y a la Universidad EAFIT por permitirnos espacios como estos, para el desarrollo de nuestras capacidades intelectuales e investigativas.

## 4. GLOSARIO

**OLTP:** Procesamiento de Transacciones En Línea (OnLine Transaction Processing) es un tipo de sistemas que facilitan y administran aplicaciones transaccionales, usualmente para entrada de datos y recuperación y procesamiento de transacciones.

**Base de datos relacional:** Es una base de datos que cumple con el modelo relacional, el cual es el modelo más utilizado en la actualidad para modelar problemas reales y administrar datos dinámicamente.

**Modelo relacional:** Es un modelo de datos basado en la lógica de predicado y en la teoría de conjuntos. Es el modelo más utilizado en la actualidad para modelar problemas reales y administrar datos dinámicamente.

**Base de datos transaccional:** Es una base de datos orientada a la realización de transacciones; es decir, la información se procesa descomponiéndola en forma unitaria e indivisible.

## TABLA DE CONTENIDO

	Pág.
LISTADO DE TABLAS.....	vii
LISTADO DE FIGURAS .....	viii
LISTADO DE ANEXOS .....	9
INTRODUCCION .....	<b>¡Error! Marcador no definido.</b>
1. CONTEXTO DE TIC EN LAS PYMES.....	<b>¡Error! Marcador no definido.</b>
2. INTELIGENCIA DE NEGOCIOS.....	<b>¡Error! Marcador no definido.</b>
2.1 Beneficios de implementar Inteligencia de Negocios o BI.....	<b>¡Error! Marcador no definido.</b>
	<b>definido.</b>
2.2 Historia de Inteligencia de Negocios o BI .....	<b>¡Error! Marcador no definido.</b>
3. DATAWAREHOUSE .....	<b>¡Error! Marcador no definido.</b>
3.1 Arquitectura del Datawarehouse.....	<b>¡Error! Marcador no definido.</b>
3.2 Herramientas de consulta y análisis .....	<b>¡Error! Marcador no definido.</b>
4. CREAR DATAWAREHOUSE E INTELIGENCIA DE NEGOCIOS.....	<b>¡Error!</b>
	<b>Marcador no definido.</b>
4.1. Fases para Crear Inteligencia de Negocios .....	<b>¡Error! Marcador no definido.</b>
4.2. Beneficios de realizar Inteligencia de Negocios .....	<b>¡Error! Marcador no definido.</b>
4.3. Herramientas de Inteligencia de Negocios.....	<b>¡Error! Marcador no definido.</b>

4.3.1.	Herramientas On Demand, SaaS (Software-as-a-service) ...	<b>¡Error! Marcador no definido.</b>
4.3.1.1.	Ventajas y características el Modelo SaaS.....	<b>¡Error! Marcador no definido.</b>
4.3.1.2.	Desventajas del Modelo SaaS .....	<b>¡Error! Marcador no definido.</b>
4.3.1.3.	Casos de éxitos, herramientas OnDemand.....	<b>¡Error! Marcador no definido.</b>
4.3.1.4.	Herramientas BI bajo Modelos SaaS .....	<b>¡Error! Marcador no definido.</b>
4.3.1.4.1.	Litebi .....	<b>¡Error! Marcador no definido.</b>
4.3.1.4.2.	Microstrategy .....	<b>¡Error! Marcador no definido.</b>
4.3.1.4.3.	Panorama Business Intelligence Software .....	<b>¡Error! Marcador no definido.</b>
5.	ANALISIS SECTOR HOTELERO DE MEDELLÍN..	<b>¡Error! Marcador no definido.</b>
5.1.	Estado del arte.....	<b>¡Error! Marcador no definido.</b>
5.2.	Comportamiento de la Oferta Hotelera.....	<b>¡Error! Marcador no definido.</b>
5.2.1.	Nuevos proyectos hoteleros .....	<b>¡Error! Marcador no definido.</b>
5.3.	Comportamiento de la demanda hotelera.....	<b>¡Error! Marcador no definido.</b>
5.4.	Proyección de la hotelería .....	<b>¡Error! Marcador no definido.</b>
6.	CASO PRACTICO: HOTEL CASA VICTORIA.....	<b>¡Error! Marcador no definido.</b>
6.1.	Descripción del Hotel Casa Victoria.....	<b>¡Error! Marcador no definido.</b>
6.2.	Implementación de Datawarehouse para el Hotel Casa Victoria...	<b>¡Error! Marcador no definido.</b>
6.2.1.	Fase 1: Dirigir y Planear – Realizar Análisis de los requerimientos .....	<b>¡Error! Marcador no definido.</b>
6.2.1.1.	Identificar preguntas .....	<b>¡Error! Marcador no definido.</b>

6.2.1.2.	Identificar indicadores y perspectivas.....	<b>¡Error! Marcador no definido.</b>
6.2.1.3.	Modelo Conceptual .....	<b>¡Error! Marcador no definido.</b>
6.2.2.	Fase 2: Recolección de información - Análisis de los OLTP ...	<b>¡Error! Marcador no definido.</b>
6.2.2.1.	Conformar indicadores.....	<b>¡Error! Marcador no definido.</b>
6.2.2.2.	Establecer correspondencias .....	<b>¡Error! Marcador no definido.</b>
6.2.3.	Fase 3: Procesamiento de datos – Creación de un Modelo lógico .....	<b>¡Error! Marcador no definido.</b>
6.2.3.1.	Tipo de Modelo Lógico del DW .....	<b>¡Error! Marcador no definido.</b>
6.2.3.2.	Tablas de dimensiones y hechos .....	<b>¡Error! Marcador no definido.</b>
6.2.4.	Fase 4: Analisis y producción .....	<b>¡Error! Marcador no definido.</b>
6.2.4.1.	Carga Inicial .....	<b>¡Error! Marcador no definido.</b>
6.2.1.	Fase 5: Difusión – Cubos multidimensionales.....	<b>¡Error! Marcador no definido.</b>
6.2.2.	Generación de reportes.....	<b>¡Error! Marcador no definido.</b>
6.2.2.1.	Generación y análisis de reportes generados sobre la DW ..	<b>¡Error! Marcador no definido.</b>
4.	CONCLUSIONES.....	<b>¡Error! Marcador no definido.</b>
5.	GLOSARIO .....	<b>¡Error! Marcador no definido.</b>
6.	BIBLIOGRAFIA .....	<b>¡Error! Marcador no definido.</b>
7.	ANEXOS .....	<b>¡Error! Marcador no definido.</b>

## LISTADO DE TABLAS

<u>Tabla 1 Oferta de habitaciones de los hoteles que reportan a SITUR, Fuente: SITUR</u> .....	41
<u>Tabla 2 Apertura de nuevos hoteles en el 2009, Fuente: SITUR</u> .....	42
<u>Tabla 3 Apertura de nuevos hoteles en el 2010, Fuente: SITUR</u> .....	43

## LISTADO DE FIGURAS

Figura 1 Datos, información y conocimiento ..... **¡Error! Marcador no definido.**

Figura 2 Ciclo de vida de las aplicaciones de soporte a la decisión .....**¡Error! Marcador no definido.**

Figura 3 - Fases de la inteligencia de negocios ..... **¡Error! Marcador no definido.**

Figura 4 – Esquema de modelo conceptual..... **¡Error! Marcador no definido.**

Figura 5 – Perspectivas y Dimensiones..... **¡Error! Marcador no definido.**

Figura 6 – Ocupación Hotelera con idéntica muestra, años 2009 y 2010.....**¡Error! Marcador no definido.**

Figura 7 – Modelo Conceptual de necesidades del proceso alojamientos. ....**¡Error! Marcador no definido.**

Figura 8 – Diagrama de Correspondencia ..... **¡Error! Marcador no definido.**

Figura 9 – Diagrama de Modelo Dimensional ..... **¡Error! Marcador no definido.**

Figura 10 – Actividades de integración de datos ..... **¡Error! Marcador no definido.**

Figura 11 – Obtención de datos OLTP ..... **¡Error! Marcador no definido.**

Figura 12 – Carga de archivo de Fechas de dimensión Tiempo**¡Error! Marcador no definido.**

Figura 13 – Indicadores de medición ..... **¡Error! Marcador no definido.**

Figura 14 – Indicadores de medición y atributos ..... **¡Error! Marcador no definido.**

Figura 15 – Indicadores de medición, atributos y jerarquías**¡Error! Marcador no definido.**

Figura 16 – Análisis de gráfica, Reporte de Cantidad de huéspedes por profesión y principales ciudades capitales en los años 2010 y 2011.**¡Error! Marcador no definido.**

Figura 17 – Análisis de gráfica, Reporte de edad promedio de huéspedes en los años 2009, 2010 y 2011. .... **¡Error! Marcador no definido.**

Figura 18 – Análisis de gráfica, Reporte de alojamientos por tarifa en los años 2009,2010 y 2011. .... **¡Error! Marcador no definido.**

Figura 19 – Análisis de gráfica, Reporte de total de alojamientos por sector económico y ciudad en los años 2009 y 2010. .... **¡Error! Marcador no definido.**

## LISTADO DE ANEXOS

No se encuentran elementos de tabla de ilustraciones. Pág.

## INTRODUCCION

Las Tecnologías de información han penetrado de tal manera el mercado empresarial que es una necesidad para cada empresa independiente de su tamaño, el adquirir soluciones informáticas para garantizar no solo la sobrevivencia de la empresa sino también el posicionamiento al enfrentarse a la competencia completamente “armada”.

La adquisición de soluciones informáticas se ha convertido para muchos en un COMODITY, lo que quiere decir, asequible en precio, sin embargo para muchas PYMES, esto no ha sido tan fácil en materia de creación de información realmente valiosa para la toma de decisiones, en muchas ocasiones por desconocimiento de herramientas, por resistencia al cambio o simplemente se escudan en que no cuentan con los recursos económicos suficientes para invertir en este tipo de soluciones.

A través del desarrollo de esta investigación les mostraremos a las PYMES del sector hotelero, cómo se puede generar información a través de una DATAWAREHOUSE y como junto a la aplicación del concepto de INTELIGENCIA DE NEGOCIOS se pueden crear soluciones asequibles y fáciles de manejar, sin necesidad en ocasiones de ser profesionales en el área informática o de una inversión muy grande; se les enseñará el cómo la automatización de toma de decisiones puede dar un giro al negocio y llevarlo un paso delante de su competencia; adelantándose a conocer el mercado y a ofrecer servicios acordes a sus necesidades reales.

Todo lo anterior se mostrará a través de un caso ejemplo aplicado a una Pyme.  
***“Hotel Casa Victoria”***

## 1. CONTEXTO DE TIC EN LAS PYMES

En vista de que se mostrará un panorama global de las TI en las PYMES es necesario que primero se defina que es TIC:

Las Tecnologías de la Información y las Comunicaciones, TIC, son el conjunto de tecnologías que permiten el acceso a la información en forma de voz, imágenes y datos, utilizando la electrónica como tecnología base.

Algunos de los elementos que componen las TIC son los computadores, los servidores, el internet, televisores, cámaras, las redes de comunicación.

Según la OCDE <sup>1</sup>, las TIC son aquellos dispositivos que capturan, transmiten y despliegan datos, información electrónica y que apoyan el crecimiento y desarrollo económico de la industria manufacturera y de servicios.

De acuerdo a Stephen Haag, Maeve Cummings y Donald J. MaCubbrey<sup>2</sup> las tecnologías de información se componen de cualquier herramienta basada en computadora que la gente utiliza para trabajar con información, apoyar a la información y procesar las necesidades de información de una organización. Bajo esta definición ellos incluyen dentro de las TIC a las computadoras personales, Internet, teléfonos móviles, asistentes personas digitales y todo aquel dispositivo similar.

---

<sup>1</sup> OECD Reviewing the ICT sector definition: Issues for discussion. Working party on indicators for the information society. Stockholm, 25-26 April 2002.

<sup>2</sup> HAAG, S., CUMMINGS M., & MCCUBBREY D. J.. *Management information systems for the information age*. (4th Edition). p. 41. New York: McGraw-Hill. 2004

Benjamín, I & Blunt, J.<sup>3</sup> definen a las tecnologías de información como todas las tecnologías basadas en computadora y comunicaciones por computadora, usadas para adquirir, almacenar, manipular y transmitir información a la gente y unidades de negocios tanto internas como externas en una organización.

Es por ello se entiende que el éxito en las PYMES, va muy ligado a la forma como estas se adelantan a su competencia, con nuevas estrategias, nuevos planes de negocios; vivimos en una época en la que la información es la clave para obtener una ventaja competitiva en el mundo de los negocios.

En Colombia, según la Ley para el Fomento de la Micro, Pequeña y Mediana Empresa, Ley 590, las PYMES se clasifican así: <sup>4</sup>

<b>De empresa</b>	<b>Empleados</b>	<b>Activos anuales (\$)</b>
Microempresa	1-10	menos de 501 (SMLMV)
Pequeña Empresa	11-50	501-5000 (SMLMV)
Mediana Empresa	51-200	5001-30000 (SMLMV)

Según investigación realizada por Fundes Colombia <sup>5</sup> en Junio 2007, sólo el 2% de las PYMES en Colombia invierten en tecnología, y por ello es necesario que

---

<sup>3</sup> BENJAMÍN, I & BLUNT, J. *Critical IT Issues: The Next Ten Years*, Sloan Management Review. p. 7-19. 1992

<sup>4</sup> MIPYMES (09-02-2007)

estas empresas aceleren su inversión en Tecnologías de Información y Comunicaciones (TIC) para ser competitivas en el mercado global. Este panorama fue presentado en el marco de la mesa redonda "Tecnología, oportunidad de crecimiento para las Pymes en Colombia", realizada por Cisco y ACOPI.

A continuación algunas cifras de un estudio realizado por FUNDES Colombia sobre las razones y prioridades en la inversión en Tecnologías de Información y Comunicación.

La PYME invierte en TIC's con los siguientes propósitos:

- 30% en tecnologías de la información
- 28% en equipamiento
- 12% para mejoramiento de sus productos
- 10% para el gerenciamiento de su negocio
- 10% para el desarrollo de nuevos productos
- 9% para el mejoramiento de su productividad

Para mantener la competitividad, los gerentes, directivos de las PYMES, requieren además de información rápida y fácil, que sea útil, que proporcione una visión del negocio en cualquier área de la compañía; algunas organizaciones, pasan más tiempo recopilando y preparando la información, que analizándola ó sacando provecho de ésta, y muchas veces no encuentran información valiosa para la toma de decisiones, por lo que acuden a la intuición o la experiencia.

Algunas PYMES que requieren cambios y mejoras en sus procedimientos actuales, proceden a invertir en infraestructura, otras abren sus puertas a nuevos

---

<sup>5</sup> La Fundación para el Desarrollo Sostenible - FUNDES

mercados, pero tan solo una pocas invierten en tecnología de información que les ayuden a estandarizar o automatizar sus procesos.

En cualquier empresa, independiente de su tamaño, se llevan los datos necesarios para su continuo funcionamiento, ya sea de forma manual, computarizada, en sistemas básicos como Excel, o tal vez se cuente con una base de datos operacional<sup>6</sup>. Una de las dificultades a la hora de sacar reportes, es que los datos se encuentran dispersos y es complejo centralizarlos; esta es información que llega tarde a las personas que toman decisiones, lo que incide en la subutilización de sus recursos y capacidades.

Por todo esto, toma importancia la utilización del concepto de Inteligencia de negocios, el cual permite pasar de los datos al conocimiento, conocimiento que utilizado de forma oportuna puede generar el valor agregado a una empresa.

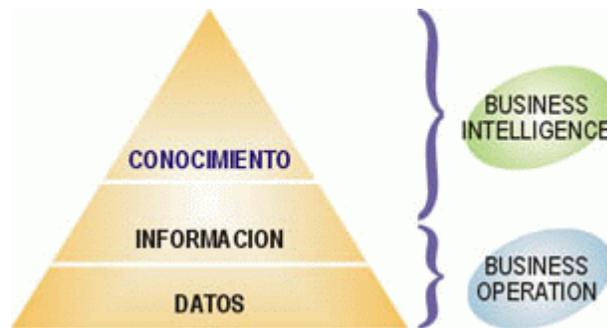
Se debe recordar que los Datos solo son un conjunto de valores que en sí solos no dan valor agregado a quien los tiene. La “información” en cambio permite darle un contexto y un propósito lo que disminuye el nivel de incertidumbre.

---

<sup>6</sup> Bases de datos operacionales: son bases de datos dinámicas, orientadas a almacenar información que es modificada con el tiempo, permitiendo operaciones como actualización y adición de datos, además de las operaciones fundamentales de consulta.

Cuando se mezcla la información con la experiencia adquirida se forma el conocimiento, el cual suele reflejarse en los procesos, prácticas y normas que hacen parte de la estrategia de una empresa.

Figura 1 Datos, información y conocimiento



Fuente: Dato, Información y Conocimiento según DAVENPORT Y PRUSAK (1999)

El concepto Inteligencia de Negocios, nace de la necesidad de gestionar, pero no se puede gestionar lo que no se controla y no se puede controlar lo que no se mide. Su objetivo es proporcionar una visión del negocio a aquellas personas encargadas de la toma de decisiones, permitiendo tomar decisiones relevantes y más rápidas, les permite comprender la manera actual en la que se encuentra la empresa y anticipar estrategias o acciones para dar una buena dirección a la empresa.

## **2. INTELIGENCIA DE NEGOCIOS**

INTELIGENCIA DE NEGOCIOS es un concepto que integra el almacenamiento y el procesamiento de grandes cantidades de datos para transformarlos en conocimiento y en decisiones en tiempo real, a través del análisis y exploración.

Hace hincapié en los procesos de recolectar y utilizar la información para mejorar las operaciones de un negocio, brindando a sus usuarios información clave para llevar a cabo las tareas habituales del negocio.

### **Beneficios de implementar Inteligencia de Negocios o BI**

Las soluciones de BI en el negocio, le permitirá:

- Mejorar la toma de decisiones basado en información más confiable, oportuna, e integrada de múltiples fuentes
- Contar con indicadores de desempeño de todo el negocio
- Fácil de integrar con los sistemas de información utilizados
- Generar reportes relevantes para la empresa
- Crear escenarios con respecto a una decisión

- Compartir información entre diferentes áreas de la organización
- Análisis multidimensionales
- Generar y procesar datos
- Mejorar el servicio al cliente
- Las soluciones de Inteligencia de Negocios, tienen la capacidad de poner en marcha actuaciones preventivas o proactivas en caso de que cambien las tendencias de modelo de contratación o del comportamiento de los clientes

### **Historia de Inteligencia de Negocios o BI**

En un tiempo las organizaciones dependían de sus departamentos de sistemas para proporcionarles reportes, y el acceso a los computadores era muy limitado, luego en los 70"s cambió un poco el panorama cuando llegaron de moda los sistemas basados en servidores; pero estos eran usados principalmente para transacciones de negocios, y los sistemas de información se sobrecargaban y se debía esperar días o semanas para obtener reportes, en caso de necesitar unos distintos a los estándares existentes.

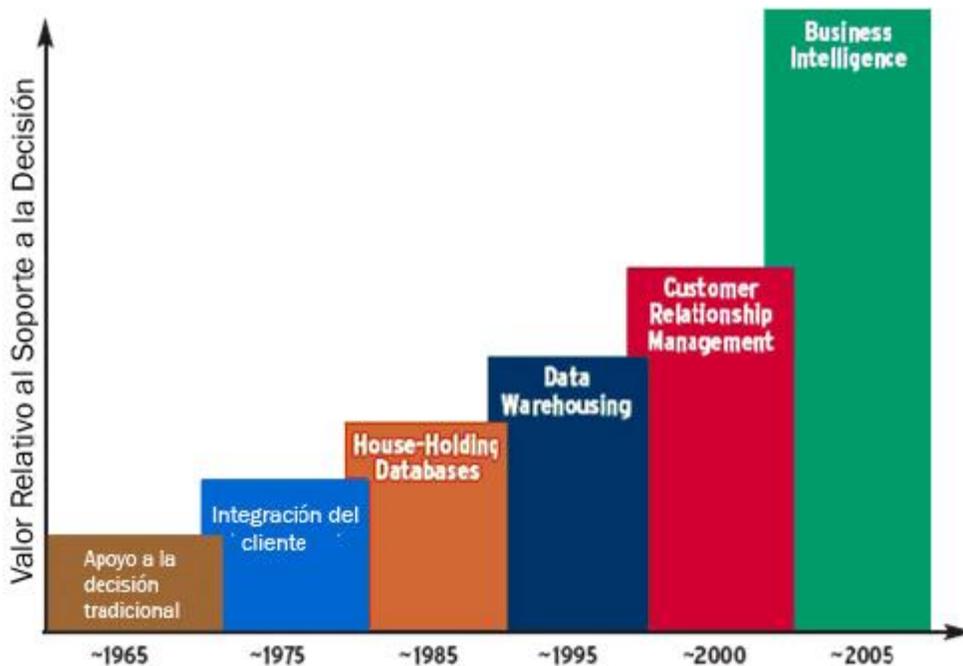
Con el tiempo se desarrollaron los EIS<sup>7</sup> los cuales fueron adaptados para apoyar las necesidades de ejecutivos y administradores. En la figura observamos el

---

<sup>7</sup> Sistema de Información Ejecutiva

desarrollo de lo que hoy se conoce como BI Business Intelligence, y como evolucionaron las aplicaciones relacionadas al soporte de decisiones.

Figura 2 Ciclo de vida de las aplicaciones de soporte a la decisión



Fuente figura: Business Intelligence Road Map (Moss & Atre, 2003).

**BI** tiene sus raíces en los Sistemas de Información Ejecutiva (Executive Information System – EIS) y en los Sistemas para la Toma de Decisiones (Decision Support Systems – DSS).

**SISTEMA DE INFORMACION EJECUTIVA (EIS):** Es una herramienta software que le permite a los gerentes de los negocios un panorama sobre los indicadores principales de rendimiento; permitiendo tomar acciones correctivas en caso que

los rendimientos obtenidos no superen los objetivos trazados. El EIS está basado en el Sistema de Soporte a la Decisión DSS.

El EIS en la mayoría de los casos cuenta con una interfaz gráfica que le permite de manera flexible e intuitiva obtener informes y reportes de forma consolidada para facilitar la monitorización del estado del negocio, permitiendo el manejo de información histórica, suele incluir alertas e informes basados en excepción, así como análisis de tendencias. También es frecuente que permita la domiciliación por correo de los informes más relevantes.

**SISTEMA DE SOPORTE A LA DECISION (DSS):** Un Sistema de Soporte a la Decisión (DSS) es una herramienta de Business Intelligence enfocada al análisis de los datos<sup>8</sup> de una organización.

En muchas ocasiones los informes que se tienen en la organización son estáticos, y no permiten navegar en los datos o manejarlos desde distintas perspectivas.

Los DSS permiten informes dinámicos, flexibles e interactivos, no requiere conocimientos técnicos y es rápido en tiempo de respuesta. Su principal objetivo a diferencia del EIS, es explotar al máximo la información residente en una base de datos corporativa (Datawarehouse o Datamart).

---

<sup>8</sup> Los datos son la mínima unidad semántica, y se corresponden con elementos primarios de información que por sí solos son irrelevantes como apoyo a la toma de decisiones.

### **3. DATAWAREHOUSE**

La base para la generación de Inteligencia de negocio o BI es un Datawarehouse. Un Datawarehouse es un repositorio de información que consolida, integra y centraliza datos generados por la empresa y los almacena mediante una estructura que permite el acceso y exploración con alto rendimiento; facilitando el análisis que permitirá la toma de decisiones estratégicas. En adelante, utilizaremos el termino DW para referirnos al Datawarehouse.

W.H. Inmon Padre del DW dice “El Datawarehouse es una colección de datos orientada al negocio, integrada, variante en el tiempo y no volátil para el soporte del proceso de toma de decisiones de la gerencia”

El DW se diseña para realizar consultas e investigaciones sobre las actividades de la organización y no para los procesos que se realizan en ella. Permite el almacenamiento de datos históricos, organiza y almacena datos que se necesitan para el procesamiento analítico e informático, con el propósito de responder a las preguntas de negocios y brindarle a los usuarios finales una interfaz amigable.

## Arquitectura del Datawarehouse

La información del DW se estructura en una base de datos normalmente desnormalizada o en cubos multidimensionales. Preparan la información para responder a consultas dinámicas.

Los componentes de la arquitectura DW son los siguientes:

- 1. Fuentes de datos:** A partir de la fuente de datos se realiza la captura de datos que se contemplará en el DW, . Estas fuentes de datos pueden ser sistemas operacionales, fuentes externas, OLTP (On line Transaction Processing), archivos de texto, hojas de cálculo. En la actualidad el componente más usado son las fuentes OLTP, las cuales representan toda la información transaccional que genera la empresa en su accionar diario.
- 2. ETL (Extracción, transformación y carga)** son extractores que extraen datos de las diversas fuentes que se requieran, los transforman para resolver posibles problemas de inconsistencia entre ellos y luego se cargan por medio de un integrador en el depósito de datos o DW.

La extracción se realiza teniendo en cuenta las necesidades y requisitos del usuario, explorando las fuentes de OLTP o los archivos de texto y planos que se tengan. Una vez los datos son seleccionados se guardan en un

almacenamiento intermedio para que no interrumpir ni paralizar los OLTP ni el DW.

La transformación se lleva a cabo debido a que muchos datos pueden presentar inconsistencia al momento de integrarlos de esta manera se garantiza que queden con un mismo formato definiendo estándares, como la medida de atributos y convenciones de nombramiento.

Los datos antes de moverse al Datawarehouse, deben ser analizados con el propósito de asegurar su calidad, por el proceso implica no solo la extracción y transformación, sino también la limpieza de los datos, para corregir y depurar errores que puedan tener las fuentes.

Por último los datos son integrados formando la información dentro del Datawarehouse.

Los DW por lo general almacenan altos volúmenes de información, por la que los aplicativos o los usuarios no acceden directamente a ellos, sino que acceden a unas vistas especializadas que son construidas de acuerdo a particulares necesidades de información, por lo general por área en una empresa y son accedidos por diferentes herramientas de consultas analíticas.

## **Herramientas de consulta y análisis**

- 3.2.1. OLAP:** El procesamiento analítico en línea OLAP es la componente más poderosa de los DW, ya que es su motor de consultas especializadas. Esta tecnología es una tecnología de software para análisis en línea,

administración y ejecución de consulta. Permite analizar el negocio desde diferentes escenarios históricos y proyectar el comportamiento y evolución en un ambiente multidimensional. Esto permite deducir tendencias, por medio del descubrimiento de relaciones, que a simple vista no se podrían encontrar sencillamente.

OLAP cuenta con diferentes operadores para explotar la información obtenida como es: Roll-up: nos permite subir las consultas de un nivel de agregación específico a otro más amplio y el Drill-Down es otro operador OLAP que nos permite bajar a un nivel más atómico de los datos.

**3.2.2. ROLAP:** El procesamiento ROLAP, accede a los datos almacenados en un Datawarehouse para proporcionar los análisis OLAP. Los sistemas ROLAP se soportan mejor contra las bases de datos relacionales. En estos sistemas las consultas son más lentas

**3.2.3. MOLAP:** El procesamiento MOLAP usa unas bases de datos multidimensionales para proveer el análisis, permite mayor rendimiento en las consultas y la posibilidad de realizar cálculos complejos, y dentro de sus desventajas es que no puede acceder a datos que no estén en los cubos previamente construidos. En estas estructuras los datos son pre calculados y almacenados en cubos de datos multidimensionales.

**3.2.4. HOLAP:** El procesamiento HOLAP combina al ROLAP Y MOLAP para brindar una solución con las mejores características de ambas: gran desempeño y escalabilidad.

**3.2.5. DATA MINING:** Son herramientas usadas para desempeñar modelado predictivo o para descubrir la relación causa efecto entre dos métricas. Es una técnica para descubrir patrones y relaciones entre abundante cantidades de datos. Los data mining se desarrollan en lenguajes basados en Inteligencia Artificial y utilizan por lo general métodos matemáticos.

En resumen, la forma de operar en un DW es la siguiente:

- Los datos son extraídos desde aplicaciones, base de datos, archivos, etc.
- Los datos se integran, se transforman y se limpian, para luego ser cargados en el DW
- La información del DW, se puede estructurar en una base de datos desnormalizada o en cubos multidimensionales
- Los usuarios acceden al DW utilizando diversas herramientas de consulta, exploración, análisis y reportes.

## 4. CREAR DATAWAREHOUSE E INTELIGENCIA DE NEGOCIOS

### 4.1. Fases para Crear Inteligencia de Negocios

Según la Metodología HEFESTO, el proceso para crear Inteligencia de Negocios está dividido en cinco fases que nos permitirán obtener el conocimiento esperado sobre los datos:

Figura 3 - Fases de la inteligencia de negocios



Fuente: <http://sites.google.com/site/magm33332/hefesto>, versión 1.1 (libro metodología HEFESTO)

**Fase 1: *Dirigir y planear – realizar análisis de los requerimientos:*** Levantar requerimientos de información y generar preguntas que le lleven a alcanzar objetivos reconociendo sus verdaderas necesidades.

Al levantar los requerimientos de información se recomienda realizar preguntas a los usuarios que permitan identificar el objetivo principal del DW, y qué cualidades debe tener para cumplir con este objetivo. Para obtener esta información se pueden utilizar diferentes medios como entrevistas, cuestionarios, encuestas y observaciones. También existen muchas herramientas en las que se pueden realizar formatos de encuestas de manera rápida y fácil como es [encuestafacil.com](http://encuestafacil.com).

Las preguntas a realizar, deben incluir variables del negocio, que se consideren relevantes y que estos datos se encuentren soportados por las bases de datos transaccionales de la empresa.

Una manera de estar seguros que los datos obtenidos son los correctos es verificar que el resultado de ellos, cumpla los objetivos estratégicos planteados por la empresa que se está estudiando.

El resultado del análisis realizado mediante preguntas y entrevistas llevará a reconocer las áreas de la empresa, o los procesos que se incluirán en la construcción del DW que contendrá información para la toma de decisiones estratégicas.

Luego de obtener las preguntas relevantes se deben reconocer **los indicadores y perspectivas** que se utilizarán, los indicadores por lo general son valores numéricos y representan lo que se desea analizar concretamente, por ejemplo: saldos, sumatorias, formulas, entre otros. Las perspectivas son los objetos mediante los cuales se va a analizar los indicadores: por ejemplo: productos, clientes, y uno muy utilizado es el tiempo.

A partir de las perspectivas y los indicadores se construye un modelo conceptual con el que se puede observar el alcance del proyecto. El modelo conceptual se

construye alrededor del proceso seleccionado a trabajar y permite tener una visión de los resultados que se obtendrán.

Figura 4 – Esquema de modelo conceptual



**Fase 2: Recolección de información – Análisis de los OLTP;** Extraer información para encontrar respuesta a las preguntas planteadas, ya sea de fuentes externas o internas.

Se debe revisar las fuentes de datos, y validar que cálculos deben realizarse para obtener los indicadores seleccionados, teniendo en cuenta los siguientes conceptos:

- Hechos que lo componen: con su respectiva formula del cálculo: Hecho1 + hecho2
- Función de sumarización que se utilizará para su agregación.

El siguiente paso es identificar la correspondencia entre el modelo conceptual y las fuentes de datos, en el cual se puedan identificar de donde se obtendrá la información. Teniendo el diagrama entidad relación, se puede realizar esta

correspondencia identificando para cada indicador y perspectiva de donde se obtendrá la información. Es importante identificar la información con la que se cuenta para calcular los indicadores seleccionados, conocer e indagar el significado de cada campo en cada una de las tablas, para luego seleccionar del conjunto de datos, contenidos en las fuentes de información, aquellos que serán de utilidad analizar los indicadores. Por cada perspectiva, se deben seleccionar los campos que serán de utilidad y por cada indicador la fórmula a utilizar.

**Fase 3: *Procesamiento de datos – creación de un modelo lógico;*** Una vez se tiene el modelo conceptual definido, con todos los actores de las diferentes fuentes de información, se debe realizar el modelo lógico para el DW, que represente la estructura de los datos a utilizar para integrar los datos en un formato que permita su análisis.

Existen diferentes tipos de modelos lógicos que se pueden implementar: modelo estrella, copo de nieve o constelación, se recomienda utilizar la que mejor se adapte a los requerimientos.

**Modelo estrella:** el modelo estrella es un modelo de datos simple, que puede ser operado por un usuario final, ya que las consultas no son complejas, las consultas a realizar solo involucran una tabla de hechos y dimensiones que se construyen a través del modelo conceptual realizado, por lo que las uniones que pueden surgir solo serían a dos niveles.

La tabla de hechos es la tabla central de un esquema dimensional (en estrella o constelación) y contiene los indicadores, o medidas numéricas enteros o reales de un DW. La tabla de hechos debe contener las claves subrogadas de las dimensiones, es por ello que se recomienda siempre, antes de definir la tabla de

hechos, realizar el diagrama conceptual y definir el detalle de la información y las fuentes de ellas.

Las tablas de dimensiones, son los elementos que contienen campos que se utilizan para restringir o agrupar los datos en las tablas de hechos, cuando se realizan consultas sobre estos datos mediante un DW. Los datos de las tablas de dimensiones proporcionan un mayor detalle de la información de los datos de la tabla de hechos.

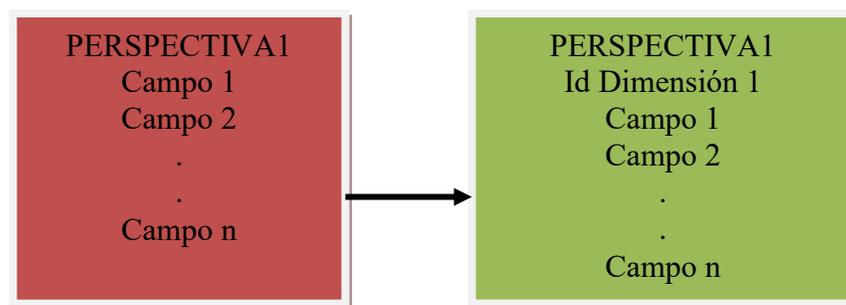
En un modelo estrella, se tiene una tabla por cada dimensión, adicional a la tabla de hechos.

Cada una de las perspectivas definidas en el modelo conceptual representará una tabla de dimensión:

Para crear la tabla de dimensión se realizará de la siguiente forma:

- Elegir un nombre que identifique la dimensión
- Añadir un campo que identifique la clave principal
- Redefinir los nombres de los campos, si los nombres que tenga no son intuitivos.

Figura 5 – Perspectivas y Dimensiones



En cualquiera de los modelos utilizados, como de nieve o estrella, una vez se tienen construidas las tablas de hechos y dimensiones, se deben realizar las uniones correspondientes y la tabla de hechos, debe tener subrogadas las claves de las dimensiones.

**Fase 4: *Análisis y producción***, creación de reportes dinámicos, indicadores, entre otros, para dar respuestas a las preguntas planteadas al inicio del proceso.

En este paso se realiza la carga de los datos a nuestro DW, y se realiza el proceso de limpieza y aseguramiento de la calidad de los datos. Se debe evitar que se ingresen datos duplicados o campos nulos o sin información, al DW solo se deben ingresar datos de interés, por lo cual se deben realizar procesos de ETL que lleven a asegurar los datos a ingresar.

Al realizar la carga de los datos, primero se deben cargar los datos de las tablas de dimensiones, y después la tabla de hechos, teniendo en cuenta las correspondencias.

En los procesos de ETL se deben definir los procesos de la ETL en general y los pasos a realizar en cada uno, partiendo de lo general a lo específico.

1. Definir las tareas que se llevarán a cabo en el proceso a nivel general
2. Especificar cada una de las tareas a realizar que conlleven a la carga de cada una de las dimensiones. Se debe especificar los criterios de la información a cargar, según los datos de interés para el usuario, en el análisis a realizar en el DW, por lo general son consultas SQL.
3. Definir la manera como se realizarán las actualizaciones de la DW y las políticas que se utilizarán.

**Fase 5: Difusión – cubos multidimensionales:** entrega de herramientas que permitan explorar los datos de manera sencilla y veloz.

Los sistemas actuales de Business Intelligence están contruidos con una arquitectura modular que integra todos los componentes en un sistema de inteligencia del negocio.

Los cubos OLAP son bases de datos multidimensionales, consiste en almacenar la información en vectores para permitir un análisis rápido, la desventaja de los cubos OLAP está en que una vez poblada la tabla no se pueden hacer modificaciones en la estructura de los datos , ya que esto implicaría rediseñar el cubo.

Los cubos permiten realizar los informes y reportes que por lo general son complejos, cuando se necesita información observada desde diferentes puntos de vista, sin el cubo, se tendría que realizar un reporte por cada uno de los puntos de vista.

Los cubos multidimensionales están basados en el modelo lógico.

Los cubos multidimensionales están compuestos por:

- **Indicadores:** son obtenidos de la tabla de hechos del modelo lógico, realizando fórmulas para la obtención las sumalizaciones (suma, promedio, etc.) sobre algún hecho, y estos dependen de los atributos o jerarquías que los utilicen.
- **Atributos:** son obtenidos de la tabla de dimensiones. Son los criterios utilizados para analizar los indicadores, estos son los que proporcionaran información de los hechos del cubo.

- Jerarquías: una jerarquía representa una relación lógica entre dos o más atributos.

#### **4.2. Beneficios de realizar Inteligencia de Negocios**

- Toma decisiones mejores y más informadas en menos tiempo.
- Controla los Indicadores claves de rendimiento del negocio
- Obtiene información detallada que afecta el rendimiento
- Permite mejorar los procesos y tomar decisiones con una visión real de la empresa y sus clientes.
- Crea cultura de uso y análisis de la información
- Permite medir el desempeño y tomar acciones correctivas

#### **4.3. Herramientas de Inteligencia de Negocios**

La mayoría de las microempresas que tienen algunas soluciones informáticas, las tienen como sistemas dispersos, donde cada una tiene su propia fuente de datos y formas de representación; esto impide la fácil actualización de información.

***Mientras más integrada este la organización será más fácil obtener la información que se necesite al momento de tomar una decisión.***

Esta necesidad de integración ha permitido que cada vez el número de herramientas de Business Intelligence aumente, cada una de ellas ofreciendo técnicas y herramientas de acuerdo al tipo negocio que se tenga, por ello la importancia de saber escoger el proveedor que se adapte mejor a la necesidad del negocio.

Para poder escoger el proveedor y la herramienta que se desea utilizar para la aplicación de BI, es importante tener un panorama de cómo funcionan cada una de ellas y cuáles son las características principales que podrá usar..

A continuación se presentarán algunas opciones existentes en el mercado.

#### **4.3.1. Herramientas On Demand, SaaS (Software-as-a-service)**

El mercado de BI se ha ido extendiendo cada vez más hacia las pequeñas y medianas empresas, que ven imposible la oportunidad de implantar inteligencia a la toma de decisiones puesto que los costos de implantación y soporte no están a su alcance. Pero actualmente la necesidad presentada por las Pymes ha generado un gran número de herramientas accesibles a estas microempresas que no disponen de mucha inversión en TIC o presupuesto para el montaje de infraestructura que se necesitan para estas aplicaciones.

**SAAS** conocido como „Software on Demand”, es una herramienta que permite acceder a todas la funcionalidades que tiene la aplicación mediante la web. Es una aplicación de software que provee el servicio de hospedaje (almacenamiento) de información que luego puede ser usada por terceros en la web.

El Modelo SAAS incorpora el concepto de *computación nube*<sup>9</sup>. Usa un modelo distribución y comercialización que permite a los usuarios pagar por lo que utilizan; este modelo surge como solución a la problemática de muchas empresas

---

<sup>9</sup> En este tipo de computación todo lo que puede ofrecer un sistema informático se ofrece como servicio de modo que los usuarios puedan acceder a los servicios disponibles en la “nube” de Internet sin conocimientos

medianas y pequeñas quienes sufren las complicaciones de costos inesperado de mantenimiento de las aplicaciones tradicionales de software.

Para la aplicación de este modelo, las empresas no requieren del conocimiento de expertos en BI, una vez iniciado el proceso de implementación en la empresa, es corto el tiempo para poder estar en ejecución por parte de los usuarios finales,

SaaS permite acceder a todas las funcionalidades mediante la web, donde un proveedor de software desarrolla la aplicación y la hospeda en Internet para que terceros puedan utilizarla, con un pago mensual o con otros periodos de pago.

Los SAAS son de gran importancia porque la mayoría de las PYMES no cuentan con grandes departamentos de sistemas, sino que el personal de cada área aprende en cierta forma a manipular los programas que necesita en el día a día. Estos productos son fáciles de usar y no necesitan de gran infraestructura, permiten la posibilidad de tener respuestas a preguntas difíciles, cuando las necesite.

#### **4.3.1.1. Ventajas y características el Modelo SaaS**

- Acceso y administración por medio de la red
- El cliente no necesita un departamento de sistemas especializado,
- La empresa de IT debe garantizar la disponibilidad de la aplicación
- La compañía IT proporciona la seguridad de la aplicación y de los datos procesados en ella.
- El cliente puede usar los sistemas operativos de su preferencia.

- El usuario no debe pagar una licencia perpetua como normalmente ocurre, por lo cual se puede tener por menor tiempo y se puede obtener retorno de manera más rápida.
- No es necesario contar un una infraestructura tecnológica adicional para el soporte de aplicaciones

#### **4.3.1.2. Desventajas del Modelo SaaS**

- La compañía TI podrá tener acceso a los datos, cuando estos no se encuentren encriptados bajo mecanismos de cifrado. Aunque existen leyes como la Ley 24.766 que protege estas labores.
- El usuario no tiene acceso al programa, por lo cual no puede hacer modificaciones
- Los usuarios del cliente dependerán del servicio de internet (ISP) por lo que las operaciones dependerán de que se restablezca el servicio.

#### **4.3.1.3. Casos de éxitos, herramientas OnDemand**

- Cualquiera que utilice los productos de antivirus de McAfee o Symantec para proteger sus PCs, seguramente utiliza las suscripciones y la opción de "Live Update", que representan otro ejemplo del Software On Demand.

#### **4.3.1.4. Herramientas BI bajo Modelos SaaS**

A continuación se presenta una breve descripción de algunas de las aplicaciones de BI que funcionan bajo el modelo SaaS:

#### **4.3.1.4.1. Litebi**

Software BI SaaS / On Demand con cubos OLAP, Reportes, Cuadros de Mando y una Herramienta ETL de integración de datos. Esta plataforma BI permite definir un Datawarehouse completo, o simplemente cargar y analizar datos de un Excel en pocos minutos.

#### **4.3.1.4.2. Microstrategy**

Permite la flexibilidad de acceso a los datos, personalización por el usuario final, tolerancia a fallos.

- Soporte multi-alquiler: permite con una única instancia del software servir a múltiples usuarios permitiendo la personalización de datos para cada usuario y los requisitos de interfaz
- Seguridad integral: se puede integrar con diversos protocolos de autenticación y sistemas de seguridad
- Alta escalabilidad y rendimiento
- Facilidad de uso: sigue los estándares de interfaz familiar incluyendo barras de herramientas, menús, manipulaciones de arrastre...
- Interfaz flexible y personalizable: se puede personalizar cualquier parte de la instancia, permitiendo la internacionalización y dialectos
- Fácil presentación de informes, análisis intuitivo y supervisión continua del negocio

#### **4.3.1.4.3. Panorama Business Intelligence Software**

Panorama SaaS BI Solution permite a los usuarios conectarse a cubos OLAP almacenados en SQL Server Analysis Services y realizar el análisis interactivo y la visualización dentro de Google Apps y Google Docs. Usando el Google Apps permite a los usuarios compartir e interactuar una hoja de cálculo, eliminando la dependencia de Excel.

#### **4.3.2. Herramientas de código abierto**

Son herramientas que han sido desarrolladas utilizando software de libre distribución, una de las mayores ventajas del código libre, es que permite la fácil evolución y desarrollo del mismo, ya que muchos desarrolladores pueden realizar mejoras sobre el código fuente y ayudar a mejorar la herramienta.

Algunas herramientas de BI bajo código abierto son:

##### **4.3.2.1. Pentaho**

Es el líder en cuanto a solución open source (código libre) permite realizar extracción de datos, cuadros de mando, permite gestionar la aplicación y los usuarios de la aplicación.

Dentro de sus beneficios está la comunicación con otros entornos vía Webservice, tiene herramientas de interfaz de usuario, permite crear contenido nuevo, tiene informes predefinidos, crear informes de manera sencilla tiene herramientas de extracción de datos.

Y dentro de sus desventajas esta que se necesita experticia para realizar informes que involucren tablas cruzadas la dificultad para realizar los cuadros de mando y la traducción al español no está al 100% .

## **5. ANALISIS SECTOR HOTELERO DE MEDELLÍN**

### **5.1. Estado del arte**

Para el año de 1999 se registró una de las crisis más altas para la inversión de la industria hotelera, ya que los niveles de ocupación cayeron a un 39.1%, lo que ocasionó grandes pérdidas para los inversionistas hoteleros, sin embargo a partir del año 2000 todo comienza a tomar un rumbo diferente, se observan leves aumentos y hasta el 2006 el incremento representa el 17.3%, equivalente a un 56.4% de ocupación, esto debido a la política de seguridad democrática implantada por el Presidente de la República el señor Álvaro Uribe, que permitió la reactivación económica del país, una excelente infraestructura hotelera y a la diversificación de los productos turísticos.

Cabe destacar que la oferta de establecimientos hoteleros en la ciudad, presenta un crecimiento que corresponde no solo a las campañas que promueven el turismo en la ciudad y otros factores como el trabajo que se ha realizado sobre la seguridad especialmente para los extranjeros que la visitan, sino a otros factores como la reestructuración de la oferta, dirigido al turista de negocios y la exoneración de impuestos por parte del Gobierno Nacional a través de la Ley 788 de 2002 y el Decreto 2755 de 2003. Gracias a esto se incrementó la llegada de nuevos constructores, promotores inmobiliarios, instituciones financieras,

empresas industriales, particulares y otros que ven una oportunidad frente al buen comportamiento de la ocupación hotelera y su potencial crecimiento en un futuro.

## 5.2. Comportamiento de la Oferta Hotelera

Durante el año 2008, en Antioquia la oferta fue de 6,450 habitaciones reportadas ante el Viceministerio de Turismo, de éstas, aproximadamente 4,000 cumplen con los requerimientos necesarios tales como excelencia en el servicio, comodidad e internet (Wi-fi), para atender las necesidades del turista.

La información brindada por los hoteles de la ciudad de Medellín durante los años ha venido cambiando año tras año. SITUR, Sistema de Indicadores Turísticos, realizó un estudio de la oferta de hoteles en los años 2008, 2009 y 2010. Como se observa en la siguiente grafica:

Zona	2008	2009	2010	% VAR (10/09)
Poblado	1.605	2.092	2.678	28,01%
Laureles - Estadio - Belén	500	508	900	77,17%
Centro	379	854	910	6,56%
<b>Total</b>	<b>2.484</b>	<b>3.453</b>	<b>4.488</b>	<b>29,97%</b>

Tabla 1. Oferta de habitaciones de los hoteles que reportan a SITUR, Fuente: SITUR

En la lista de hoteles que tiene el Registro Nacional de Turismo se encuentran un total de 192 establecimientos entre los cuales se encuentran hoteles, apartahoteles, hostería, hostel, residencias, entre otros. De esos 192 establecimientos 79 informan a SITUR, lo cual quiere decir

que la participación de la muestra de SITUR con relación a los hoteles inscritos al Registro Nacional de Turismo es del 41,15%.

### 5.2.1. Nuevos proyectos hoteleros

Apertura de Nuevos Hoteles. Habitaciones Nuevas Medellín, 2009

Zona	Sede	Nº de habitaciones
Centro	Plazuela San Ignacio	98
Poblado	GHL San Diego	78
Poblado	Estelar Milla de Oro	168
Poblado	Art Hotel	54
Poblado	Apartamentos Estelar (Torre Oviedo)	51
Poblado	Las Rosas Hotel Boutique	14
Laureles	Suite Confort	45
Laureles	Hotel Platinum Plaza	16
<b>Total</b>		<b>524</b>

Tabla 2. Apertura de nuevos hoteles en el 2009, Fuente: SITUR

Apertura de Nuevos Hoteles. Habitaciones Nuevas en Medellín, 2010

Zona	Hotel	Nº de habitaciones
Poblado	Best Western Skyplus	34
Poblado	Diez Hotel	115
Poblado	Pacific	168
Guayabal	Puerto Seco Centro Empresarial	120
Poblado	Solers Gardens	144
Poblado	The Charlee (Royal Realty)	42
Conquistadores	33 Conquest Hotel	47
<b>Total</b>		<b>873</b>

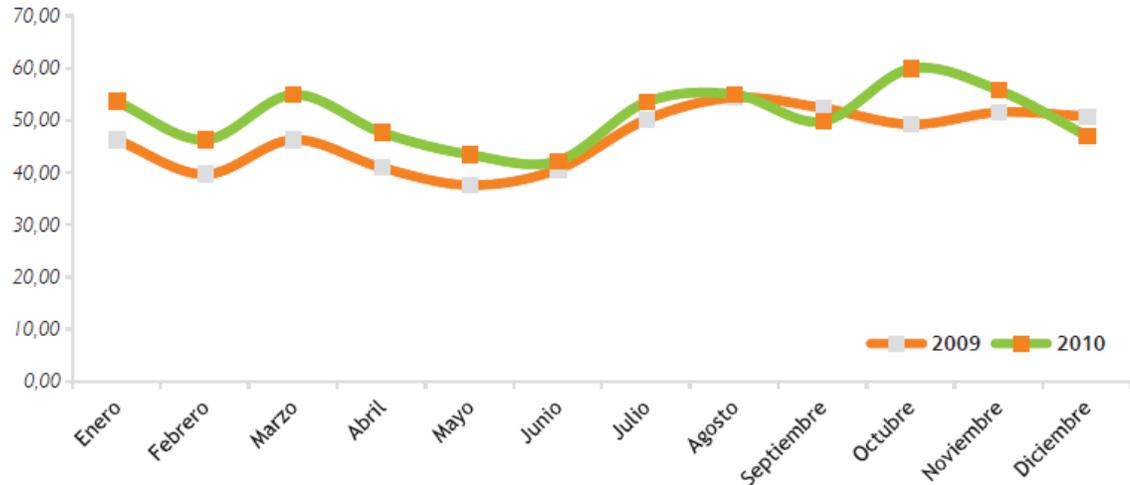
Tabla 3. Apertura de nuevos hoteles en el 2010, Fuente: SITUR

En las anteriores gráficas se puede apreciar el incremento del 2009 al 2010 de un total de 349 habitaciones nuevas a partir de los nuevos proyectos hoteleros en la ciudad de Medellín.

### 5.3. Comportamiento de la demanda hotelera

Para los años 2009 y 2010 se registra un incremento en la ocupación, con un promedio de porcentaje de variación del 4,11%. En la gráfica se puede observar que en casi todos los meses los niveles de ocupación aumentaron considerablemente exceptuando Septiembre y Diciembre que registraron disminuciones en las ocupaciones.

Figura 6 – Ocupación Hotelera con idéntica muestra, años 2009 y 2010.



Fuente: SITUR Sistema de Indicadores Turísticos

#### 5.4. Proyección de la hotelería

El 2011 es un año en el cual el sector hotelero en la ciudad de Medellín tendrá que indagar, planearse y proyectarse a futuro, puesto que el crecimiento de la hotelería tendría que estar medido por el análisis de la demanda actual y la que se espera en los próximos años.

El crecimiento de la infraestructura hotelera, podría representar para aquellos hoteles menos diferenciados y/o no asociados a grandes cadenas, una reducción de precios y bajos niveles de ocupación, de no elevarse los niveles de demanda; ante este incremento en la oferta, se requiere tomar medidas para potenciar y diversificar los atractivos diferenciadores de los hoteles y realizar un esfuerzo notable en la creación de estrategias para la promoción y comercialización de la ciudad, esto con el fin de equilibrar la balanza entre oferta y demanda.

## **6. CASO PRACTICO: HOTEL CASA VICTORIA**

### **6.1. Descripción del Hotel Casa Victoria**

La empresa analizada: Hotel Casa Victoria, es una empresa prestadora de servicios la cual posee clientes nacionales e internacionales, de acuerdo a su volumen de operaciones y número de empleados puede ser considerada como una mediana empresa.

El Hotel Casa Victoria se encuentra ubicado en la ciudad de Medellín, cuenta con aproximadamente 20 empleados distribuidos en toda su estructura organizacional. Dicha estructura está dividida en dos partes; el área administrativa que está conformado por los departamentos comercial, financiero, costos, sistemas, recursos humanos, y el área de servicios generales conformado por las diferentes unidades de servicio como: Restaurante, ama de llaves, recepción.

El departamento comercial está conformado por empleados cuya función principal es estar en contacto frecuentemente con los clientes y suministrarle información oportuna acerca del hotel, dada la competencia con la que se encuentran en el sector hotelero, se escogió esta empresa como caso práctico para aplicarle inteligencia de negocios por medio de una Datawarehouse, con el fin de encontrar estrategias de mercadeo para atrapar al cliente, fidelizarlo a la compañía y ser más competitivo.

Se realizaron entrevistas a diferentes usuarios del área administrativa y de sistemas quienes en la actualidad son los encargados de solicitar y generar los reportes para la toma de decisiones. El objetivo principal a mediano plazo del hotel

es buscar maximizar las utilidades de la empresa y expandirse a nuevos mercados con el fin de conseguir una mayor cantidad de clientes para así poder posicionarse competitivamente. La política que se utiliza para cubrir los objetivos antes mencionados, es la de satisfacer ampliamente las necesidades de los clientes, brindándoles un mejor servicio, confiabilidad, seguridad y promoviendo un excelente ambiente familiar en la empresa, todo eso se logra buscando una mejoría continuamente, teniendo unas buenas estrategias y estando preparados para los futuros cambios que se presentarán en el ámbito hotelero en el país, con el fin de estar a la altura de las exigencias del mercado y así poder conseguir una mejor posición frente a la competencia.

Este caso práctico se basará en la información obtenida por personal administrativo del hotel Casa Victoria.

## **6.2. Implementación de Datawarehouse para el Hotel Casa Victoria**

Para la implementación de la Datawarehouse se usó la Metodología Hefesto.

A continuación se presenta el resumen del proceso de la metodología aplicada al caso práctico.

### **6.2.1. Fase 1: Dirigir y Planear – Realizar Análisis de los requerimientos**

#### **6.2.1.1. Identificar preguntas**

Se realizaron entrevistas a los usuarios de las diferentes áreas administrativas, un ingeniero de sistemas, un administrador y un contador, con el fin de encontrar

necesidades de información y obtener un mayor entendimiento del negocio. Una vez recolectas dichas necesidades se les preguntó cuál era el proceso de negocio que consideraban más importante para el hotel. El proceso fue el de Alojamientos.

Luego se les preguntó que deseaban analizar en el proceso de alojamientos y las respuestas obtenidas fueron:

- Conocer el número de alojamientos,
- Conocer los ingresos recaudados en alojamientos,
- Conocer el porcentaje de ocupación en cada una de las tarifas
- Conocer el porcentaje de alojamiento de hombres y mujeres
- Conocer la edad promedio de los huéspedes en el hotel.

A continuación se les pregunto cuáles serían las variables que consideraban se podrían usar para medir el proceso de negocio escogido. La respuesta fue la siguiente:

- Número de alojamientos en un período determinado
- Valor recaudado de alojamientos en los diferentes sectores económicos de las empresas alojadas en un período determinado
- Número de alojamientos por tarifa en un periodo determinado
- Valor recaudado de alojamientos de las diferentes ciudades de empresas alojadas en período determinado
- Número de huéspedes que hay por profesión y región en un período determinado
- Edades promedio de huéspedes alojados según el sexo

### **6.2.1.2. Identificar indicadores y perspectivas**

Luego de obtener los requerimientos por parte de los usuarios, en el siguiente paso se identificaron los indicadores a medir y las diferentes perspectivas a analizar.

Indicadores identificados:

- Total valor recaudado de alojamientos
- Número de alojamientos
- Cantidad de huéspedes
- Calculo de edades promedio de huéspedes

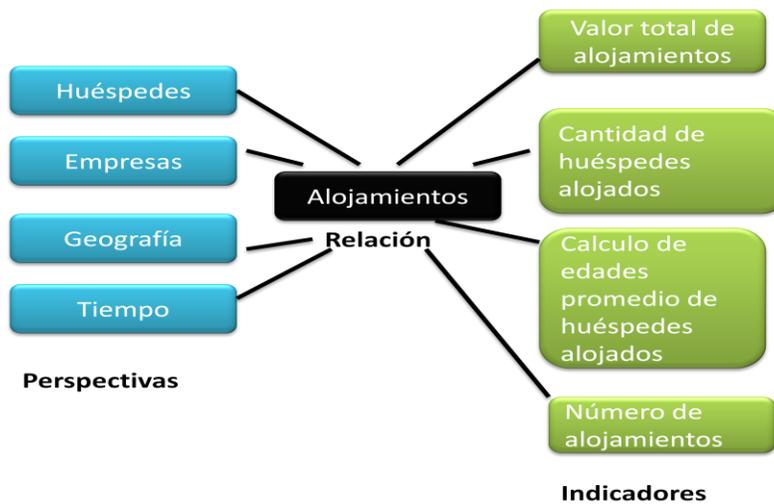
Perspectivas identificadas:

- Huéspedes
- Empresas
- Tiempo
- Geografía

### **6.2.1.3. Modelo Conceptual**

A continuación se muestra el modelo conceptual que surge de los datos encontrados.

Figura 7 – Modelo Conceptual de necesidades del proceso alojamientos.



## 6.2.2. Fase 2: Recolección de información - Análisis de los OLTP

### 6.2.2.1. Conformar indicadores

Los indicadores se calculan de la siguiente manera:

- Valor total de alojamientos

Hechos: Total alojamientos

Función de sumarización: SUM

- Cantidad de huéspedes alojados

Hechos: Cantidad de huéspedes

Función de sumarización: SUM

- Calculo de edades promedio de huéspedes alojados

Hechos: Calculo Edades

Función de promedio: AVERAGE

- Número de alojamientos

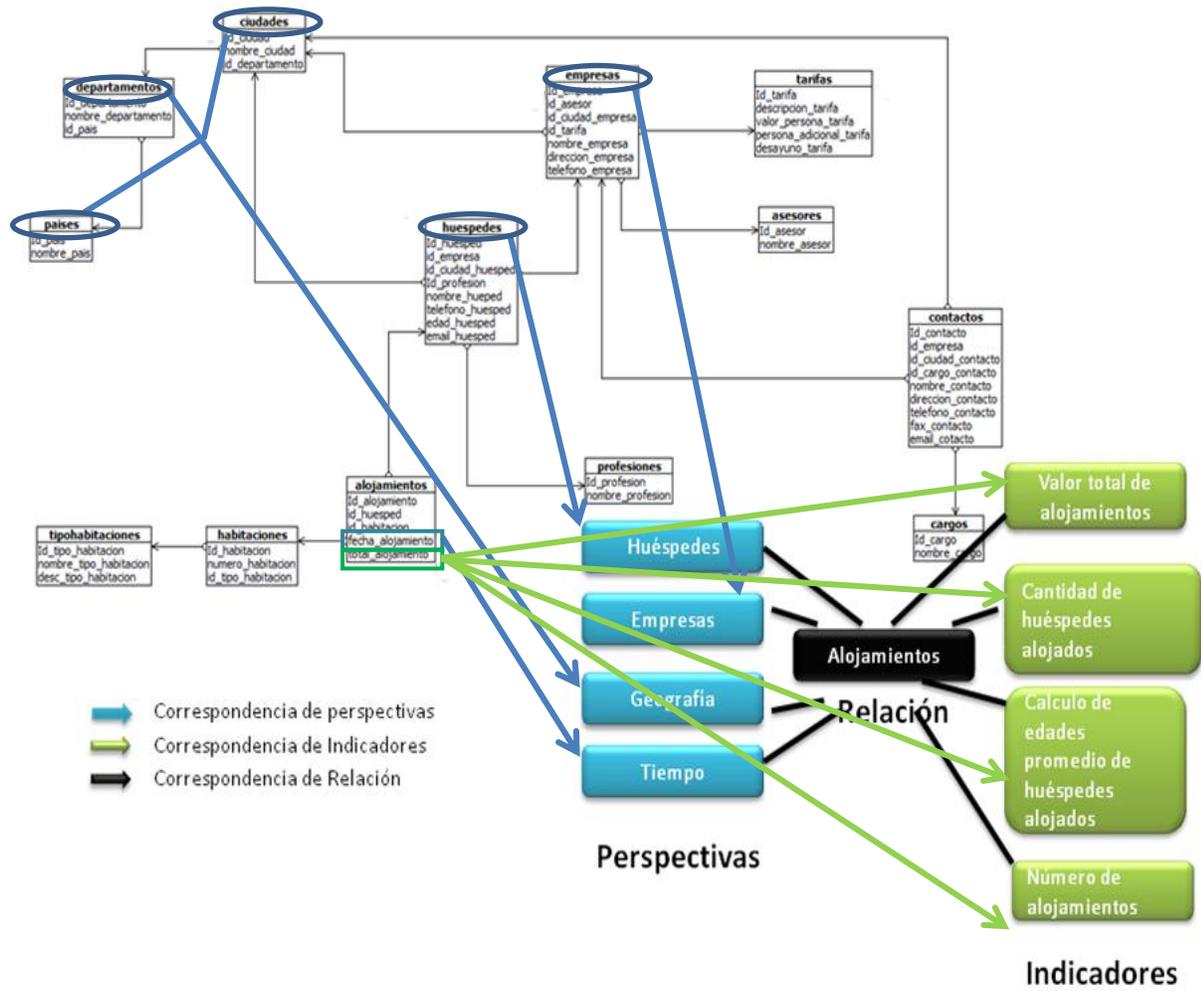
Hechos: Número alojamientos

Función de conteo: COUNT

#### **6.2.2.2. Establecer correspondencias**

Se examinó la base de datos del hotel y con base a las perspectivas e indicadores se realizaron las respectivas correspondencias.

Figura 8 – Diagrama de Correspondencia



### **6.2.3. Fase 3: Procesamiento de datos – Creación de un Modelo lógico**

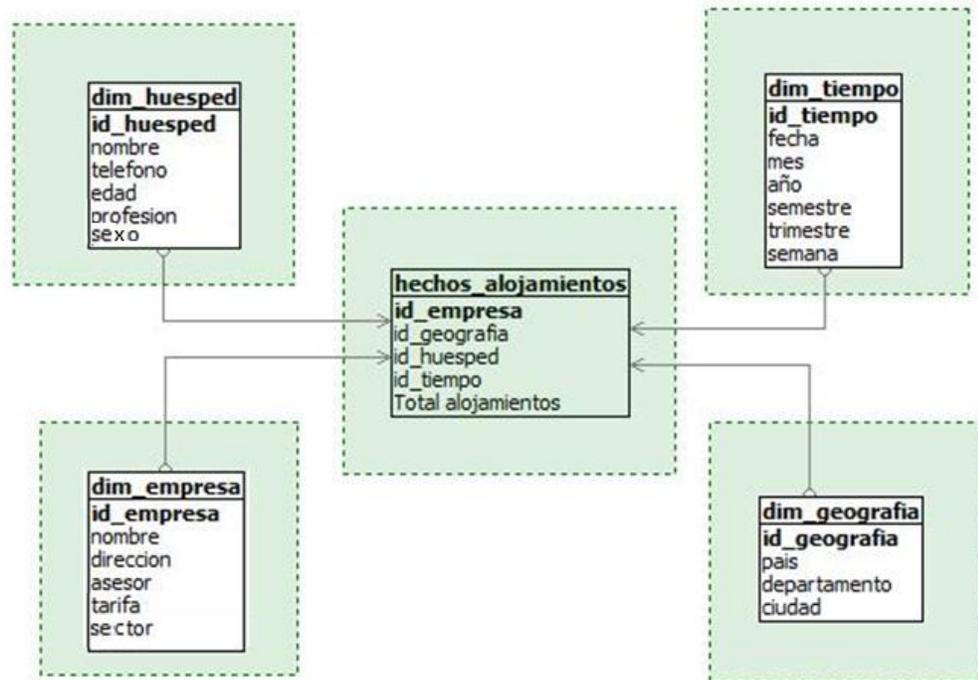
#### **6.2.3.1. Tipo de Modelo Lógico del DW**

Para el modelamiento de tabla de hechos y dimensiones se escogió el modelo estrella debido a que es un modelo entendible, se puede aplicar a cualquier sistema y los tiempos de respuestas de servidor tienden a ser más rápidos porque se recorren menos tablas en los queries.

#### **6.2.3.2. Tablas de dimensiones y hechos**

A continuación se muestra el modelo dimensional que surge del análisis de datos:

Figura 9 – Diagrama de Modelo Dimensional



## 6.2.4. Fase 4: Analisis y producción

### 6.2.4.1. Carga Inicial

Una vez conocidas las tablas de dimensiones y de hechos se procede a hacer la integración y carga de datos.

Las actividades realizadas para este proceso son las siguientes:

- Se inicia la ejecución de los pasos en una fecha de corte definida por los usuarios
- Se establecen las variables Fecha\_Desde y Fecha\_Hasta

- Fecha\_Desde: Se obtiene el valor de la fecha en la que se realizó el primer alojamiento
- Fecha\_Hasta: Se obtiene la fecha actual
- Se cargan las tablas de dimensiones con las consultas SQL y las transformaciones realizadas
- Se carga la tabla de hechos

Figura 10 – Actividades de integración de datos



Para realizar la carga de las tablas de dimensiones y de hechos se obtuvieron los datos del OLTP por medio de consultas SQL

Se realizaron las siguientes transformaciones:

Para el sexo del huésped:

Valor de consulta	Valor Transformado
1	M
2	F

Para los sectores económicos:

Valor de consulta	Valor Transformado
<b>SectAgric</b>	Agrícola
<b>SectConst</b>	Construcción
<b>SectServ</b>	Servicios
<b>SectIndust</b>	Industrial
<b>SecTrans</b>	Transporte
<b>SecMiner</b>	Minero
<b>SecFinan</b>	Financiero

Para las tarifas:

Valor de consulta	Valor Transformado
<b>1</b>	Tarifa_normal
<b>2</b>	Corporativa2
<b>3</b>	Corporativa3
<b>4</b>	Corporativa4
<b>5</b>	Corporativa5
<b>6</b>	Corporativa6

Ejemplo de query de transformación realizado:

```
UPDATE dim_Empresas SET sector =  
CASE sector When „SectAgric“ Then „Agricola“
```

When „SectConst “ Then „Construcción“

When „SectServ“ Then „Servicios“

When „SectIndust“ Then „Industrial“

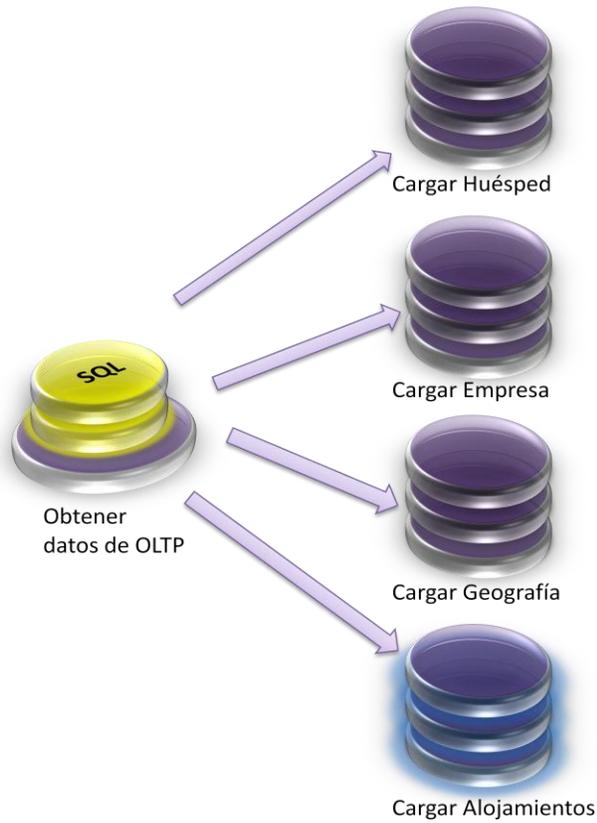
When „SecTrans“ Then „Transporte“

When „SecMiner“ Then „Minero“

When „SecFinan“ Then „Financiero“

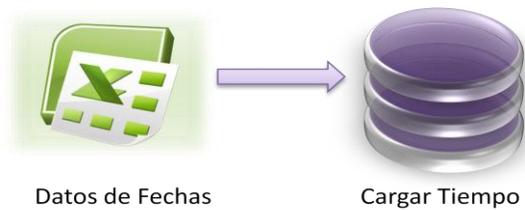
Else „Otros“ END

Figura 11 – Obtención de datos OLTP



Para generar la dimensión Tiempo se creó manualmente datos en un archivo de Excel con las diferentes fechas y luego se exportaron.

Figura 12 – Carga de archivo de Fechas de dimensión Tiempo



### 6.2.1. Fase 5: Difusión – Cubos multidimensionales

La manera para explorar los datos obtenidos en nuestro Datawarehouse será mediante cubos OLAP.

A continuación se creará un cubo multidimensional "Cubo de Alojamientos" y estará basado en el modelo lógico diseñado anteriormente.

Se crearán cuatro indicadores que serán incluidos en el "Cubo de Alojamientos":

- De la tabla de hechos "Alojamientos", se sumará el hecho "Total Alojamientos" para crear el indicador denominado:
  - "Total Alojamiento".

La fórmula utilizada para crear este indicador es la siguiente:

- "Total Alojamiento" =  $SUM(\text{Alojamientos.Total\_alojamientos})$

- De la tabla de hechos "Alojamientos", se sumará el hecho "Cantidad de huéspedes" para crear el indicador denominado:
  - "Cantidad de huéspedes".

La fórmula utilizada para crear este indicador es la siguiente:

- “Cantidad huéspedes” = SUM(Alojamientos.Cantidad\_huespedes).
- De la tabla de hechos “Alojamientos”, se sumará el hecho “Calculo de edades” para crear el indicador denominado:
  - “Cálculo de edades”.

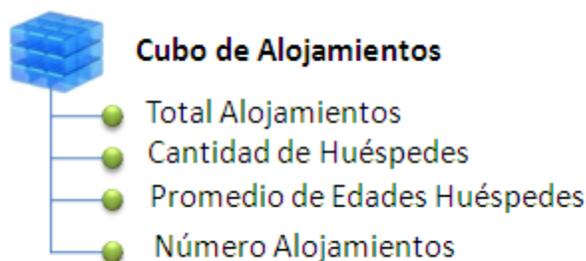
La fórmula utilizada para crear este indicador es la siguiente:

- “Cantidad huéspedes” = AVERAGE(Alojamientos.calculo\_edades).
- De la tabla de hechos “Alojamientos”, se sumará el hecho “Número de alojamientos” para crear el indicador denominado:
  - “Número de alojamientos”.

La fórmula utilizada para crear este indicador es la siguiente:

- “Cantidad huéspedes” COUNT(Alojamientos.Número\_alojamientos).

Figura 13 – Indicadores de medición

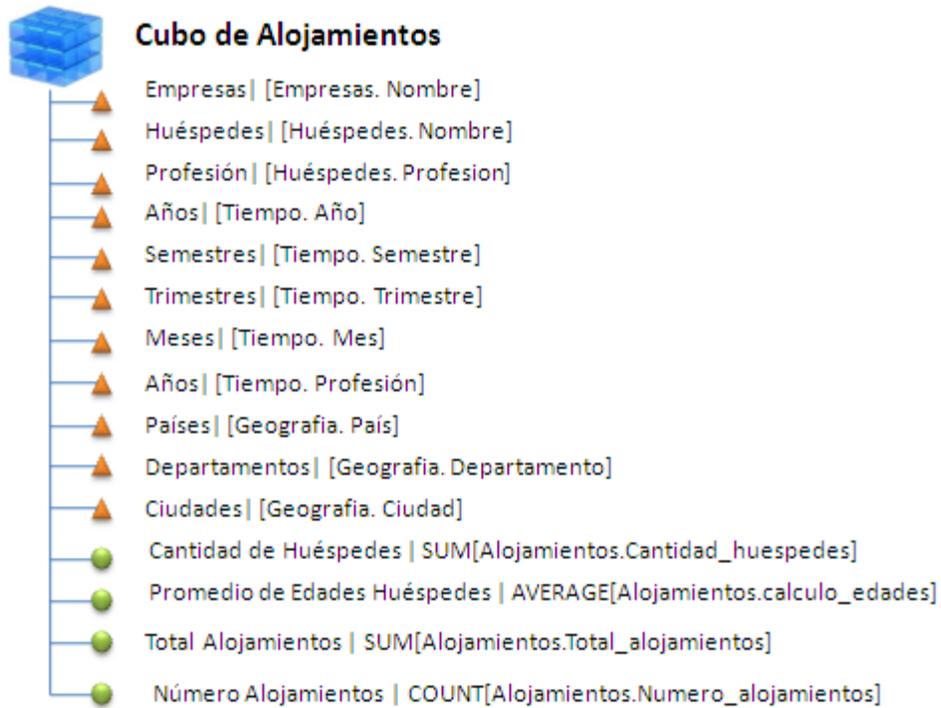


Ahora se crearán y agregarán al cubo seis atributos:

- De la tabla de dimensión “EMPRESA”, se tomará el campo “nombre” para la creación del atributo denominado:
  - “Empresas”.

- De la tabla de dimensión “HUESPÉD”, se tomará el campo “nombre” para la creación del atributo denominado:
  - “Huéspedes”.
- De la tabla de dimensión “HUESPÉD”, se tomará el campo “Profesión” para la creación del atributo denominado:
  - “Profesión”.
- De la tabla de dimensión “TIEMPO”, se tomará el campo “Mes” para la creación del atributo denominado:
  - “Meses”.
- De la tabla de dimensión “TIEMPO”, se tomará el campo “Año” para la creación del atributo denominado: “
  - Años”.
- De la tabla de dimensión “TIEMPO”, se tomará el campo “Semestre” para la creación del atributo denominado:
  - “Semestres”.
- De la tabla de dimensión “TIEMPO”, se tomará el campo “Trimestre” para la creación del atributo denominado:
  - “Trimestres”.
- De la tabla de dimensión “GEOGRAFIA”, se tomará el campo “País” para la creación del atributo denominado:
  - “Países”.
- De la tabla de dimensión “GEOGRAFIA”, se tomará el campo “Departamento” para la creación del atributo denominado:
  - “Departamentos”.
- De la tabla de dimensión “GEOGRAFIA”, se tomará el campo “Ciudad” para la creación del atributo denominado:
  - “Ciudades”.

Figura 14 – Indicadores de medición y atributos



Finalmente se crearán y agregarán tres jerarquías:

Jerarquía de Huéspedes: Compuesta de Huéspedes, Profesiones

Jerarquía de Tiempos: Compuesta de Años, Semestres, Trimestres, Meses

Jerarquía de Geografías: Compuesta de Países, Departamentos, Ciudades

Figura 15 – Indicadores de medición, atributos y jerarquías



## 6.2.2. Generación de reportes

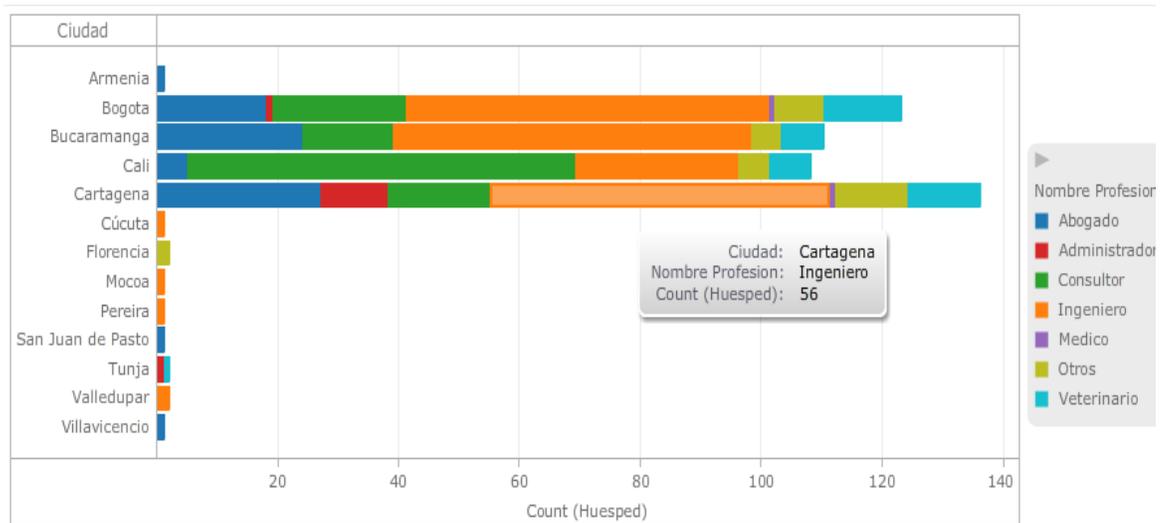
Una vez finalizado el proceso ETL se procede a la generación de reportes.

Los reportes se generaron en la herramienta MicroStrategy Cloud . A continuación se analizan algunos reportes generados que aportan valor agregado al negocio.

### 6.2.2.1. Generación y análisis de reportes generados sobre la DW

En la siguiente gráfica observamos que en los años 2010 y 2011 se ve una clara tendencia mayorista de alojamientos por parte de huéspedes Ingenieros provenientes de las ciudades de Cartagena, Bogotá y Bucaramanga. Con base en esta información para el año 2012, el hotel podría mejorar su infraestructura tecnológica con las comodidades que este tipo de profesional podría necesitar. Como por ejemplo, invertir para mejorar la red inalámbrica, esto con el fin de direccionar de una manera oportuna y rápida el servicio de internet a todos los huéspedes que desarrollan actividades a largas horas de la noche desde sus portátiles.

Figura 16 – Análisis de gráfica, Reporte de Cantidad de huéspedes por profesión y principales ciudades capitales en los años 2010 y 2011.



Teniendo en cuenta que en los meses de mayo, agosto y septiembre se tiene un alto porcentaje de ocupación, es de interés para el hotel conocer la manera de aumentar el nivel de satisfacción de los huéspedes que los visitan y obtener publicidad indirecta. Para esto, el hotel puede desarrollar eventos y actividades orientados a la mujer o al hombre teniendo en cuenta la edad promedio de los huéspedes en dichos meses. Es importante la información histórica de los años anteriores para tomar decisiones basadas en hechos y no en supuestos.

La siguiente gráfica muestra la edad promedio de huéspedes por sexo, alojados en los meses de Mayo, Agosto y Septiembre en los años 2009, 2010 y 2011.

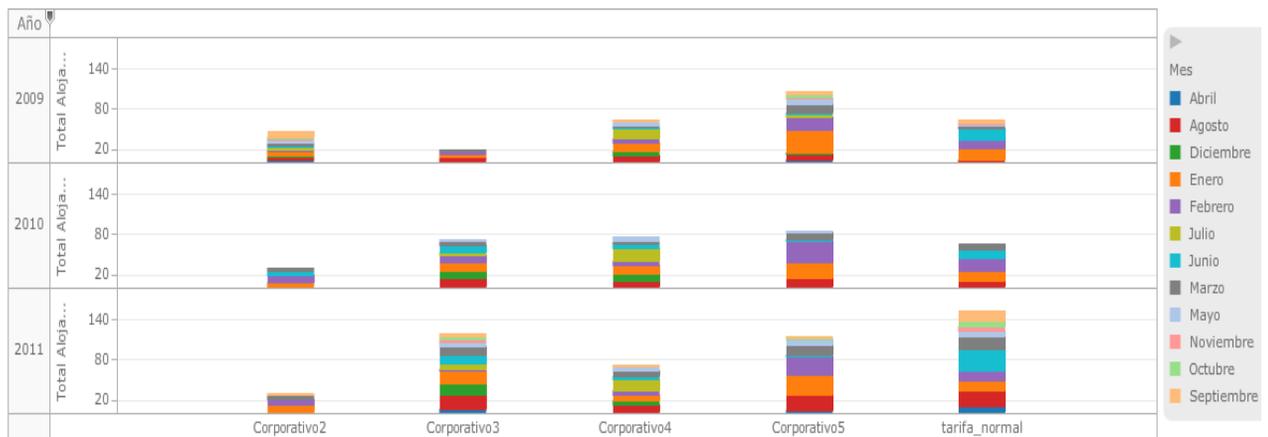
Figura 17 – Análisis de gráfica, Reporte de edad promedio de huéspedes en los años 2009, 2010 y 2011.



Para el Hotel es muy importante saber el comportamiento de ventas de las tarifas en el año, esto con el fin de mejorar y suplir las expectativas de sus diversos

clientes con diferentes alternativas a escoger. Esta gráfica muestra la cantidad de alojamientos por tarifa y mes en los periodos 2009, 2010 y 2011. Esta información le podría servir al Hotel para identificar que tarifas no se están vendiendo para poder actualizarlas ya sea con promociones o reajustando su valor y en que temporadas fortalecer las ventas. En la gráfica se observa que la tarifa Corporativo2 se está vendiendo poco y las tarifas que más se venden son la tarifa corporativa 5 y tarifa normal, además se observa que los meses de enero, febrero y junio son los meses donde más se vende y en los otros meses se observa poca ocupación.

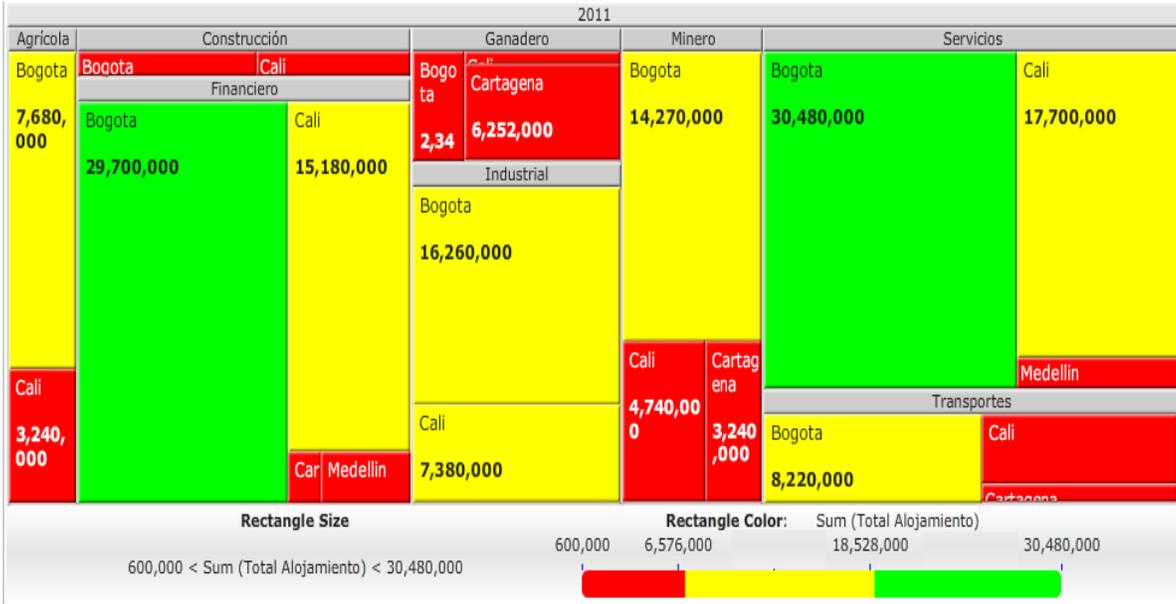
Figura 18 – Análisis de gráfica, Reporte de alojamientos por tarifa en los años 2009,2010 y 2011.



Para el hotel es importante saber que sectores económicos de empresas generan mayor alojamiento en el mes de Diciembre, En la gráfica se observa que la ciudad de Bogotá predomina en el total de alojamientos y que el sector Financiero y de Servicios son en los que más se alojan empresas. Esta información le sirve al hotel para adecuar las habitaciones con los requisitos de cada ciudad, cultura y población, por ejemplo, como en Bogotá manejan un clima muy frio es necesario

tener un buen aire acondicionado que permita mantener un ambiente cálido y frío como el clima de su ciudad, haciendo sentir al huésped como en casa. También se pueden realizar eventos y actividades de interés acorde a los sectores económicos de empresas que más se alojen en dicho periodo. Por ejemplo, para el caso financiero, en los diferentes televisores (Lobby, restaurante) se puede colocar canales que indique constantemente como están los indicadores económicos en el momento, además de noticias financieras, esto con el fin de que el huésped esté a la vanguardia.

Figura 19 – Análisis de gráfica, Reporte de total de alojamientos por sector económico y ciudad en los años 2009 y 2010.



## 4. CONCLUSIONES

Como se buscó desde el comienzo, el objetivo de este proyecto consistió en la creación de un Datawarehouse para una Pyme del sector Hotelero de la ciudad de Medellín, utilizando el concepto de inteligencia de negocio.

Dentro del trabajo práctico y las investigaciones realizadas se logró llegar a las siguientes conclusiones:

- La construcción de un Datawarehouse, en ocasiones depende de cómo tengamos la información fuente u origen. El hotel Casa Victoria, pyme con la que se realizó el trabajo práctico, aunque no contaba con un Datawarehouse, contaba con una base de datos relacional, en la que los datos se encontraban en un porcentaje alto, normalizados, lo cual facilitó el número de transformaciones que hubo que realizar sobre los datos al cargarlos al Datawarehouse.
- Para seleccionar la información de interés para el Datawarehouse se tuvo en cuenta la realización de entrevistas y encuestas; La comunicación con los usuarios que operarán el sistema es fundamental para una buena definición de la información a almacenar. Aunque es recomendable que en las entrevistas a realizar participen gerentes, o personas con alto poder de decisión, en la práctica encontramos que existen otras personas encargadas de la selección de la información, es por ello que en nuestra definición participaron personas del área administrativa y del departamento de sistemas, quienes en el diario vivir del hotel, son los encargados de generar la información, reportes, estadísticas y preparar las decisiones que luego serán aprobadas o rechazadas por quienes tienen

este poder de decisión; Obteniendo de esta forma, información útil para la toma de decisiones y la creación del Datawarehouse.

- En la actualidad existen diversas herramientas para gestionar la información de un datawarehouse, muchas de estas herramientas pueden ser usadas por los usuarios finales de un reporte, mientras otras, deben ser manipuladas solo por aquellos con conocimientos técnicos en sistemas. El hotel casa victoria es un hotel en la ciudad de Medellín, el cuál, aunque posee un buen nivel de reconocimiento en la ciudad, no tiene en sus prioridades actuales la inversión de altos costos en sistemas de inteligencia de negocios, es por ello que dentro del marco teórico de esta tesis, mostramos las herramientas BI en diferentes modalidades de contratación, como son las Saas, muy valoradas en los últimos tiempos, utilizando para nuestra aplicación práctica el software de BI de Microstrategy, reconocida, y que de acuerdo al movimiento del mercado actual, ha sacado una herramienta para uso de empresas como el Hotel Casa Victoria, Microstrategy Cloud Personal.

Esta herramienta permite a un bajo costo, hacer uso de las bondades de algunas funcionalidades de BI en su versión Cloud Personal, las cuales les permite llevar el control de los indicadores y perspectivas por medio del cual desee hacer seguimiento a sus principales procesos de negocio.

- El Hotel casa victoria podría contar con la información justo a tiempo, a la altura de sus competidores, teniendo a mano, el comportamiento de la demanda en el hotel de acuerdo a cada una de las tarifas ofrecidas y a las demandas del mercado del hotel, identificando la procedencia de la mayoría de los huéspedes que suelen frecuentar el hotel, permitiendo mejorar la calidad del servicio, según las expectativas de los huéspedes,

entre otras tomas de decisiones que podrá tomar de acuerdo a la información obtenida.

- Crear Inteligencia de Negocios, no necesariamente implica entonces, contar con tecnología de punta, o altos costos de inversión, se puede crear inteligencia de negocios, partiendo inclusive de información contenida en un Excel, garantizando únicamente la calidad de los datos y con herramientas On demand que son asequible para las Pymes.

## 4. BIBLIOGRAFIA

### REFERENCIAS BIBLIOGRAFICAS PARA LIBROS, FOLLETOS E INFORMES

ADAMSON Cristopher. Mastering datawarehouse aggregates. Wiley 2006

JIAWEY Han, KAMBER Micheline. Data Mining Concepts and Techniques. Morgan Kaufmann 2000

### REFERENCIAS DOCUMENTALES PARA FUENTES DE INFORMACION ELECTRONICA

[1].Sinnexus.[http://www.sinnexus.com/business\\_intelligence/piramide\\_negocio.aspx](http://www.sinnexus.com/business_intelligence/piramide_negocio.aspx)

[2].Sinnexus, [http://www.sinnexus.com/business\\_intelligence/olap\\_avanzado.aspx](http://www.sinnexus.com/business_intelligence/olap_avanzado.aspx)

[3].Consulmatica, [http://www.consulmatica.com.py/casos\\_bancard.html](http://www.consulmatica.com.py/casos_bancard.html)

[4].Dataprix, <http://www.dataprix.com/data-warehousing-y-metodologia-hefest>

### REFERENCIAS PROGRAMAS DE COMPUTACION (SOFTWARE)

Microstrategy,<https://cloud2.microstrategy.com>

<http://cxo-community.com/articulos/blogs/blogs-metodologia-legislacion/2016-como-proteger-informacionfidencial-acuerdos-de-confidencialidad.html>

[http://es.wikipedia.org/wiki/Computaci%C3%B3n\\_Nube](http://es.wikipedia.org/wiki/Computaci%C3%B3n_Nube).

<http://www.gartner.com/it/page.jsp?id=707508>

## 4. ANEXOS

### Encuestas realizadas

Encuesta No. 1. Para conocer las tecnologías de información con la que cuentan los hoteles en Medellín, en especial el Hotel Casa Victoria.

ENCUESTA INFORMACION UTIL EN HOTELES DE MEDELLIN  
DIRIGIDA A GERENTES Y ADMINISTRADORES DE HOTELES  
MEDELLIN – ANTIOQUIA

Abandonar->

Continuaré más tarde

#### 1.- Proyecto de Grado estudiantes Universidad EAFIT

Creación de una Bodega de Datos aplicando el concepto de Business Intelligence en las PYMES, Sector Hotelero Medellín

##### \*1. Nombre del Hotel

##### \*2. El numero de empleados del Hotel esta en el rango de:

- Menos de 5 Empleados
- De 5 - 50 Empleados
- De 50 - 100 Empleados
- Mas de 100 Empleados

##### \*3. El hotel cuenta con sistemas de información automatizados para

- Ventas
- Contabilidad
- Software hotelero ( reservas y servicios del hotel)
- Recursos humanos
- Compras
- Otro (Por favor especifique)

**\*4. Cuando necesita información que incluya resultado de operaciones del negocio que hace?**

- Saca un reporte de un sistema integrado usted mismo
- Lo solicita al área de sistemas para que consulten en el sistema o base de datos
- Pregunta a los empleados por los datos que necesita
- Diseña manualmente un informe que le arroje datos en Excel u otro similar
- Otro (Por favor especifique)

**\*5. Por lo general los informes que solicita o necesita tienen que ver con el área de:**

- Ventas
- Contabilidad
- Atención Al Cliente
- Marketing y publicidad
- Operación diaria del hotel: habitaciones y servicios en el hotel, flujo de caja diario...
- Otro (Por favor especifique)

**\*6. La información que generalmente solicita es para tomar decisiones acerca de:**

- Nuevas estrategias de ventas
- Mejorar el servicio en el hotel
- Tarifas y descuentos promocionales
- Archivar resultados mensuales
- Estimación y reducción de costos
- Otro (Por favor especifique)

**\*7. Conocer información para tomar decisiones es importante. Califique según su criterio de importancia de mayor a menor los siguientes aspectos, siendo 4 el más importante y 1 el menos**

	1. 1	2. 2	3. 3	4. 4
Conocer el total facturado por empresas/clientes	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Conocer el total facturado por cada uno de los servicios que ofrece	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Conocer de que ciudad/país provienen la mayoría de los clientes	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Saber cuales son los meses más rentables dentro de un periodo	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**\*8. Tecnologías, aplicaciones y sistemas utilizados en el hotel**

- Software hotelero (reservas, servicios...)
- Manejador de bases de datos
- ERP( Planificación de Recursos Empresariales) u otros sistemas integrados
- CRM (Administración de la relación con los clientes)
- Reportes dinámicos
- Otro (Por favor especifique)

Fin->

Encuesta No. 2 para conocer la necesidad del Hotel, que se tendrá en cuenta en el Datawarehouse.

## REQUERIMIENTOS HOTEL CASA VICTORIA

Abandonar->

Continuaré más tarde

### 1.- Cambia este título por el de la página

**\*1. Cual es el proceso de negocio que considera más importante para el Hotel?**

- Proceso de Alojamientos
- Inventario de Productos
- Facturación y costos
- Otro (Por favor especifique)

**\*2. Que metrica considera importante para medir el proceso de negocio que considera más importante?**

- Cantidad por periodo de tiempo
- Ingresos por procesos de negocios
- tarifas por numero de huéspedes
- sexualidad por periodo de tiempo
- ventas por cliente
- ingresos por mes

Que preguntas debe responder un informe del proceso seleccionado?

---