

**UNIVERSIDAD GABRIELA MISTRAL**  
**FACULTAD DE INGENIERIA**

**INTEGRACIÓN DE E-COMMERCE Y BUSINESS INTELLIGENCE CON  
TECNOLOGÍA OPEN SOURCE PARA PYMES**

Memoria para optar al título de Ingeniero de Ejecución en Informática

Autor : Daniel Andrés Manque Paineo  
Profesor Guía : Roberto Carú Cisternas

Santiago – Chile  
Diciembre, 2017

## INDICE

I.	INTRODUCCION .....	6
I.1	Que es PETS O MASCOTAS ON LINE.....	8
I.2	Servicio y Ventajas Competitivas.....	8
I.3	Motivación .....	9
I.4	Hipótesis.....	9
I.5	Objetivo General.....	9
I.6	Objetivos Específicos .....	10
I.7	Alcances .....	11
I.8	Actividades .....	11
II.	MARCO TEORICO.....	13
II.1	Comercio electronico.....	13
II.2	Inteligencia de Negocios.....	14
II.3	Tecnología OLTP.....	15
II.4	Tecnología OLAP .....	20
II.5	Tecnología OLAP vs Tecnología OLTP.....	23
II.6	Conceptos Básicos de un ETL .....	25
II.7	Definición de Inteligencia de Negocios.....	27
II.8	Características de una herramienta BI .....	29
II.9	Data Mart.....	30
II.10	Data Warehouse.....	31
III.	CONTEXTO .....	35
III.1	La Importancia de la información.....	35
III.1.1	Cifras en PETS.....	35
III.2	Situación Actual.....	36
III.3	Propuesta .....	38
III.4	Beneficiarios .....	40
III.5	Organización del Proyecto.....	40
IV.	DESARROLLO DEL PROYECTO.....	41

---

IV.1 Planificación Y ADMINISTRACIÓN del Proyecto .....	41
IV.1.1 Metodología utilizada.....	42
IV.1.2 Planificación del proyecto.....	43
IV.1.3 Definición del Proyecto.....	45
IV.1.4 Estudio de Factibilidad .....	46
IV.2 Selección de PRODUCTOS PARA e-commerce y Business Intelligence .....	56
IV.2.1 MAGENTO E-COMMERCE.....	57
IV.2.2 JASPERSOFT Y TALEND.....	59
IV.2.3 pentaho BUSINESS ANALYTICS.....	60
IV.2.4 Elección de los Productos Business Intelligence.....	62
IV.3 Capacitación en la Herramienta.....	65
IV.4 Diseño de la Arquitectura Técnica .....	66
IV.5 Modelamiento Dimensional .....	67
IV.6 Instalacion y Configuracion de Productos.....	70
IV.7 MAGENTO 1.9.....	70
IV.8 JASPERSOFT .....	72
IV.9 Diseño Físico.....	73
IV.10 Extracción, Transformación y carga (ETL).....	75
IV.11 Desarrollando las Aplicaciones .....	77
IV.12 Implementacion y Go-Live.....	82
IV.13 Mantencion y Crecimiento.....	89
V. RESULTADOS.....	90
V.1 Resultados Intangibles .....	90
V.2 Resultados Tangibles .....	90
V.3 Ventajas y Desventajas .....	91
V.3.1 Desventajas.....	91
V.3.2 Ventajas .....	91
VI. Conclusión .....	93
VII. Glosario.....	95
VIII. Bibliografía .....	101

## INDICE DE IMÁGENES

IMAGEN 1, EJEMPLO DE UN SISTEMA TRANSACCIONAL (OLTP).....	15
IMAGEN 2, HOJA DE CÁLCULO EN EXCEL. ....	19
IMAGEN 3, TABLA DE UNA BASE DE DATOS RELACIONAL.....	20
IMAGEN 4, EJEMPLO DE CUBO MULTIDIMENSIONAL (VITT, LUCKEVICH, MISNER. 2002). ....	21
IMAGEN 5, DISTINTAS IMPLEMENTACIONES DE LAS TECNOLOGÍAS OLAP. ....	23
IMAGEN 6, RECUADRO CON LAS PRINCIPALES DIFERENCIAS ENTRE UN OLTP VS OLAP. ....	24
IMAGEN 7, COMPARACIÓN VISUAL ENTRE UN OLTP Y OLAP. ....	24
IMAGEN 8, DIAGRAMA SIMPLE DEL PROCESO EXTRACCIÓN, TRANSFORMACIÓN Y CARGA.....	25
IMAGEN 9 DIAGRAMA PARA DIFERENCIAR DATOS, INFORMACIÓN Y CONOCIMIENTO .....	28
IMAGEN 10 CICLO VIRTUOSO DE BUSINESS INTELLIGENCE .....	28
IMAGEN 11 ARQUITECTURA TÍPICA DE UN SISTEMA DE INTELIGENCIA DE NEGOCIO. ....	29
IMAGEN 12 REPRESENTACIÓN DE LA TECNOLOGÍA DATA WAREHOUSE .....	30
IMAGEN 13 REPRESENTACIÓN DE UN DATA MART.....	31
IMAGEN 14 DIAGRAMA DEL PROCESO DE UN DATA WAREHOUSE .....	34
IMAGEN 15 EJEMPLO DE UNA CONSULTA SQL PARA UN INFORME.....	37
IMAGEN 16 EJEMPLO EN EXCEL DE INFORME GENERADO.....	38
IMAGEN 17 FASES DE LA METODOLOGÍA DE CASCADA.....	42
IMAGEN 18 CARTA GANTT DEL PROYECTO.....	45
IMAGEN 19 PLATAFORMAS E COMMERCE .....	58
IMAGEN 20 ARQUITECTURA DE TALEND Y JASPERSOFT.....	60
IMAGEN 21 ARQUITECTURA PENTAHO ANALYTICS.....	62
IMAGEN 22 GOOGLE ANALYTICS.....	65
IMAGEN 23 ARQUITECTURA PROPUESTA PARA LA SOLUCIÓN .....	66
IMAGEN 24 MODELO LÓGICO MULTIDIMENSIONAL.....	69
IMAGEN 25 PARTE DEL PROCESO DE INSTALACIÓN DE MAGENTO .....	70
IMAGEN 26 PARTE DEL PROCESO DE CONFIGURACIÓN DE MAGENTO 1.9.....	71
IMAGEN 27 PARTE DEL PROCESO DE INSTALACIÓN DE JASPERSOFT .....	72
IMAGEN 28 PARTE DEL PROCESO DE CONFIGURACIÓN JASPERSERVER.....	73

---

IMAGEN 29 MODELO FÍSICO MULTIDIMENSIONAL – DATAMART VISITA Y DATAMART VENTAS.....	74
IMAGEN 30 PROCESO ETL TALEND OPEN STUDIO FOR DATA INTEGRATION.....	76
IMAGEN 31 PORTAL E-COMMERCE EN MAGENTO .....	77
IMAGEN 32 PORTAL MAGENTO LOGIN MÓDULO ADMIN.....	78
IMAGEN 33 PORTAL MAGENTO, MÓDULO DE ADMINISTRACIÓN.....	78
IMAGEN 34 IMPLEMENTACIÓN DE GOOGLE ANALYTICS EN PORTAL MAGENTO 1.9.....	79
IMAGEN 35 ETL DATAMART DE VENTAS.....	79
IMAGEN 36 ETL DATAMART DE VISITAS .....	80
IMAGEN 37 CALENDARIZANDO EL CRON .....	80
IMAGEN 38 DESARROLLANDO EN IREPORT – DATAMART VISITAS .....	81
IMAGEN 39 DESARROLLANDO EN IREPORT - DESIGNER.....	81
IMAGEN 40 PAGINA LOGIN DE USUARIOS JASPERSERVER.....	83
IMAGEN 41 PORTAL DE PROYECTOS DE JASPERSERVER .....	84
IMAGEN 42 PARÁMETROS REPORTE DE TIPO DE VISITAS.....	85
IMAGEN 43 REPORTE DE TIPO DE VISITAS.....	86
IMAGEN 44 PARÁMETROS REPORTE DE CLIENTES .....	87
IMAGEN 45 REPORTE DE USUARIOS .....	88
IMAGEN 46 EL CIRCULO DE DEMING PDCA.....	89

## Dedicatoria

---

La tecnología es algo que me mueve. No todos tenemos la posibilidad de trabajar y entretenerse a la vez. Doy gracias a Dios por darme la posibilidad de trabajar en lo que me apasiona.

---

## Agradecimientos

En primer lugar, agradezco a mi profesor guía señor Roberto Carú, quien tuvo siempre una excelente disponibilidad para ayudarme y apoyarme en lo que necesitaba. Además, agradezco sus consejos y recomendaciones para cumplir con mis objetivos.

A mis profesores de la Universidad por su entrega de conocimiento y apoyo.

¡Gracias Dios, Familia, Amigos y Compañeros!

## **I. INTRODUCCION**

En una sociedad capitalista y poco equitativa como la nuestra, donde las tiendas On Line es una realidad para unos pocos, se hace necesario darle la oportunidad a nuestras pymes para que puedan competir y desarrollarse de igual a igual con las grandes empresas o trasnacionales. Sin embargo, el alto costo de soluciones informáticas de e-commerce y ni hablar del costo de una solución business intelligence, además de licencias y mano de obra tecnológica, hace de esto una “utopía” para la mayoría de las pymes. Esto aumenta aún más la brecha entre las grandes corporaciones y nuestras pymes.

En los tiempos actuales el e-commerce tiene variedad de soluciones en Open Source, Soluciones profesionales que nada tienen que envidiar a las versiones de pago y pueden ser utilizados por cualquier emprendedor que quiera beneficiarse de los canales de pago que ofrece la nube.

Un e-commerce Open Source versión community si bien representa un gran ahorro, se debe poseer un conocimiento especializado para su implementación. Pero que este sea un software gratis representa una gran oportunidad para dar soluciones profesionales de ventas On Line a nuestras pymes.

Si el e-commerce Open Source es una gran oportunidad para las pymes, esta se convierte en una solución de ligas mayores si la integramos con una solución Business Intelligence Open Source versión community. Soluciones reservadas únicamente para las grandes corporaciones como la banca y el retail, nuestra finalidad es traer esta tecnología y dejarlas al alcance de las pymes a un mínimo costo y/o subsidiarlo por el estado.



Se pretende dar un servicio de implementación de e-commerce y Business intelligence que consta de dos datamart, un datamart de visitas y un datamart de ventas, donde se generen reportes periódicos que sean enviados de forma automática a las respectivas pymes, es decir, la solución debe soportar múltiples pymes y deben coexistir entre ellas. De esta manera cada pyme estará en condiciones de visualizar sus tendencias y tomar acciones preventivas o correctivas para incrementar sus ventas.

La presente tesis, está orientada al desarrollo e implementación de un e-commerce para una tienda On Line de mascotas Pets, pero además se implementará una solución Business Intelligence, es decir: Una solución que según muchos expertos posicionan al BI como la inversión e innovación prioritaria para las grandes empresas y la consideran indispensable para sobrevivir en el entorno del mercado actual.

### ***1.1 QUE ES PETS O MASCOTAS ON LINE***

Mascotas On Line es un proyecto de pyme que comercializa principalmente casas para mascotas, además de otros insumos para las mascotas, es decir una pequeña empresa con escasas posibilidades de implementar una solución e-commerce y business intelligence para mejorar sus ventas.

### ***1.2 SERVICIO Y VENTAJAS COMPETITIVAS***

Actualmente las mascotas forman parte de nuestras vidas cotidianas, por lo tanto, la demanda de productos y servicios para el bienestar de nuestras mascotas es cada vez mayor. Si bien existe un gran oferta y demanda de los diversos productos, accesorios y servicios para nuestras mascotas, “La pequeña empresa Pets” no ha tenido dado la posibilidad de tener un canal de venta On Line por ser una pequeña empresa, sin embargo, sabemos que aventurarse en el comercio electrónico es una excelente oportunidad para que Pets aumente sus niveles de ventas en sus productos y servicios.

Si bien el e-commerce es una excelente oportunidad para aumentar nuestras ventas 24 x 7 y darnos a conocer al mundo. La solución Business Intelligence que acompaña al e-commerce tiene por finalidad aventajarnos con respecto a nuestros competidores y mantenernos vigentes en el mercado, estrechando la brecha con las grandes empresas.

### ***1.3 MOTIVACIÓN***

Nuestra motivación es apoyar al desarrollo de las pymes, darle la oportunidad de tener acceso a herramientas tecnológicas exclusivas para las grandes organizaciones. Hacer competir a la pyme de igual a igual con las grandes corporaciones, darle la posibilidad que su negocio opere 24 x 7 todos los días del año. Todo esto respaldado con un sistema de Inteligencia de negocio para la tomar acciones preventivas o correctivas si el negocio lo requiere, ver tendencias, campañas, etc.

Creemos firmemente que bajando de la elite al e-commerce y el Business Intelligence y dar la posibilidad a nuestras pymes de contar con estas tecnologías a un precio razonable, entonces, estaremos aportando nuestro grano de arena a la sociedad Chilena.

### ***1.4 HIPÓTESIS***

Con la implementación de un e-commerce y una solución Business Intelligence, se pretende aumentar las ventas de la tienda “Pets”. Rompiendo con el paradigma del uso exclusivo del business intelligence sólo en grandes organizaciones.

### ***1.5 OBJETIVO GENERAL***

Incursionar con tecnología Open Source versión community en las pymes a un precio razonable. Dar la posibilidad a la pyme de abrirse al comercio electrónico e-commerce y business intelligence.

## ***1.6 OBJETIVOS ESPECÍFICOS***

Demostrar que podemos romper el paradigma de que el e-commerce y el business intelligence están reservados para las grandes empresas.

Implementando soluciones e-commerce y business intelligence Open source en una pequeña empresa para mascotas podemos aumentar las ventas y estar operativa 24 x 7, con una solución BI estándar para incrementar las ventas y tener el dato duro para una buena toma de decisiones.

En resumen, los objetivos son:

- Implementar una tienda On Line con Magento e-commerce para que el negocio funcione 24 x 7 los 365 días del año.
- Implementar botones de pago On Line.
- Implementar 2 Datamart, Datamart de Ventas y Datamart de Visitas utilizando una herramienta ETL y Base de datos Open source.
- Implementar reportes genéricos con herramientas BI Open source community.
- Posicionar a Pets como una pyme con tecnología de punta en ventas On Line a un precio razonable o subsidiado por el estado.

## **I.7 ALCANCES**

El alcance es desarrollar el proyecto de e-commerce e Inteligencia de Negocios, cumpliendo los principales hitos:

- ✓ Implementar e-commerce con configuración básica y productos simples.
- ✓ Diseñar 2 datamart genéricos, uno de ventas y otro de visitas.
- ✓ Implementar el proceso de extracción, transformación y carga de los datos.
- ✓ Implementar la herramienta Business Intelligence para el análisis y explotación de la información.
- ✓ Desarrollar el modelo inicial, demostrando las bondades de tener implementado una solución Business Intelligence.

## **I.8 ACTIVIDADES**

Para desarrollar este proyecto se deben realizar una serie de actividades, las que se detallan a continuación:

### 1) Planificación y Análisis del requerimiento

- ✓ *Definición del proyecto*
- ✓ *Necesidad, alcances y objetivos del cliente.*
- ✓ *Levantamiento de la información.*

### 2) Diseño de la solución

- ✓ *Metodologías de desarrollo.*
- ✓ *Diseño de DataMart, ETL,*
- ✓ *Diseño de Indicadores, Reportes, Interfaces, Dashboard.*
- ✓ *Selección de las herramientas.*

### 3) Construcción y Pruebas

- ✓ *Configuración y customize de e-commerce*
- ✓ *Preparación Ambiente de pruebas*
- ✓ *Construcción de Datamart de visitas y ventas*
- ✓ *Construcción de Procesos ETL*
- ✓ *Orquestación de procesos ETL*
- ✓ *Construcción de Indicadores, Reportes, Interfaces*
- ✓ *Pruebas unitarias e integrales.*

## **II. MARCO TEORICO**

### **II.1 COMERCIO ELECTRONICO**

Definición.

En el mundo del mercado económico, existen varias formas por las cuales se puede hacer transacciones de bienes o servicios entre individuos o entre un grupo de ellos. Con el avanzar del tiempo, estas formas de transacción han ido cambiando de manera progresiva, buscando siempre el dinamismo y la facilidad para un intercambio de mayor eficiencia. En la actualidad, la tecnología es una herramienta que ha construido una nueva puerta en el mercado de los negocios: las redes de comunicación electrónica o también llamado el comercio electrónico.

Según Teresa Moreda (2013), el concepto de comercio electrónico no sólo incluye la compra y venta electrónica de bienes, información o servicios, sino también el uso de la Red para actividades anteriores o posteriores a la venta, como son:

- ✓ La publicidad
- ✓ El establecimiento del contacto inicial entre un cliente potencial y un proveedor potencial.
- ✓ La búsqueda de información sobre productos, proveedores, etc.
- ✓ Intercambio de información
- ✓ La negociación entre comprador y vendedor sobre precio, condiciones de entrega, etc.
- ✓ La venta y distribución
- ✓ La atención al cliente antes y después de la venta.
- ✓ El pago electrónico
- ✓ La cumplimentación de trámites administrativos relacionados con la actividad comercial.

- ✓ La colaboración entre empresas con negocios comunes (a largo plazo o sólo de forma coyuntural)

## **II.2 INTELIGENCIA DE NEGOCIOS**

La Inteligencia de Negocios se ha consolidado en el mundo empresarial con dos niveles básicos; en primer lugar, los altos ejecutivos necesitan obtener información estratégica para la toma de decisiones en tiempo real y en segundo lugar, los responsables del análisis táctico necesitan esta información, como también para creación de campañas comerciales y de marketing.

Desde la década de los 90 las herramientas BI han evolucionado exponencialmente, desde informes operacionales, hasta ambientes OLAP multidimensionales, esto significa que las organizaciones demanden más formas de analizar y realizar reportes de datos.

Generalmente las Medianas y Grandes empresas, invierten en herramientas de planificación de recursos (ERP) y también en herramientas de administración basada en la relación con el cliente (CRM) o en comercio electrónico e-commerce. En consecuencia han almacenado una enorme cantidad de datos la que desean usar para ayudarles a ser más ágiles frente a la competencia, tomar mejores decisiones, y tener mayor comprensión de sus negocios. Pero estos sistemas están almacenados en sistemas con procesamiento para transacciones en línea OLTP, un ejemplo podrían ser los sistemas de retail para compras en tienda, como también a través de ventas on-line vía Web y Telefónica.



### II.3 TECNOLOGÍA OLTP

En muchas ocasiones, un almacén de datos se utiliza como el fundamento de un sistema de ayuda para la toma de decisiones. Los almacenes de datos se han diseñado para superar algunos de los problemas que una organización encuentra, cuando intenta realizar un análisis estratégico, mediante la misma base de datos que utiliza para realizar el proceso de transacciones en línea (OLTP).



Imagen 1, Ejemplo de un sistema transaccional (OLTP).

**Normalmente, los sistemas OLTP:**

- Admiten el acceso simultáneo de muchos usuarios que agregan y modifican datos.
- Representan el estado, en cambio constante, de una organización, pero no guardan su historial.
- Contienen grandes cantidades de datos, incluidos los datos extensivos utilizados para comprobar transacciones.
- Tienen estructuras complejas.
- Se ajustan para dar respuesta a la actividad transaccional.
- Proporcionan la infraestructura tecnológica necesaria para admitir las operaciones diarias de la empresa.

**Algunas dificultades al utilizar bases de datos OLTP para realizar análisis en línea:**

- Los analistas carecen de la experiencia técnica necesaria para crear consultas "ad-hoc" contra la compleja estructura de datos.
- Las consultas analíticas que resumen grandes volúmenes de datos afectan negativamente a la capacidad del sistema para responder a las transacciones en línea.
- El rendimiento del sistema cuando está respondiendo a consultas analíticas complejas puede ser lento o impredecible, lo que causa un servicio poco eficiente a los usuarios del proceso analítico en línea.
- Los datos que se modifican con frecuencia interfieren en la coherencia de la información analítica.
- La seguridad se complica cuando se combina el análisis en línea con el proceso de transacciones en línea.

**Ejemplo concreto de un sistema transaccional OLTP**

En una empresa de retail o financiera, el sistema transaccional, debe poseer la característica de controlar las transacciones en línea, para mantener consistencia en los datos y brindar seguridad a los datos que serán almacenados. Por ejemplo una compra en la web con tarjeta de la misma tienda, debe ser capaz que al momento de confirmar la compra, el disponible de su tarjeta (almacenado en una Base de datos transaccional), debe ser disminuido al valor del producto adquirido, en caso de que la transacción no se realiza de manera completa, la “transacción” debe ser reversada.

El sistema de procesamiento transaccional puede ser capaz de realizar cálculos y procesos, administrando múltiples transacciones, eso sí, tiene la obligación en determinar la prioridad de estas. Tal como pasa por ejemplo en la compra online de entradas para un concierto con asientos numerados, esa entrada debe quedar bloqueada de manera temporal, hasta que la transacción se complete, ya que otro cliente, de manera paralela, podría reservar el mismo asiento del concierto.

**Los sistemas OLTP tienen principalmente las siguientes propiedades:**

- Los procesos y sus cálculos son generalmente simples.
- Es uno de los principales sistemas que se implementan en una organización
- Están enfocados a las áreas de venta, marketing, finanzas o Administración y Recursos Humanos.
- Automatizan las tareas operativas en una organización, de esta manera permite ahorrar tiempo y en mano de obra.
- Son optimizados para almacenar grandes volúmenes de datos, pero no para analizar los mismos.
- Los beneficios de este tipo de sistemas en una organización son rápidamente visibles.

### **Las principales características de los sistemas transaccionales**

- ✓ Seguridad: Los sistemas OLTP deben ser altamente fiables, ya que de lo contrario podría afectar al negocio o clientes y por ende la reputación de la empresa. Para eso se deben mantener mecanismos de respaldo y recuperación de datos.
- ✓ Velocidad: Deben responder de manera rápida, dependiendo de la cantidad de datos almacenados, estas respuestas puede ser desde milésimas de segundos, a un par de segundos.
- ✓ Consistencia: No debe permitir agregar información distinta a la establecida, por ejemplo, no ingresar valores de texto, donde solo deben ser números, o ingresar un valor distinto a los permitidos de acuerdo a ciertas referencias.

Actualmente este modelo de base de datos relacionales son los más usados, este modelo está basado en la teoría de conjuntos. Es decir se almacena en relaciones o con modelo EAV (Magento) y cada relación es un conjunto de datos. De esta manera es más fácil de entender para usuarios que no son expertos.

El modelo relacional o EAV es como una colección de relaciones, la relación representa una tabla, que es un conjunto de filas y cada fila contiene campos y cada campo contiene un valor. En algunos casos a la fila se le denomina tupla, y a la columna se le denomina campo.

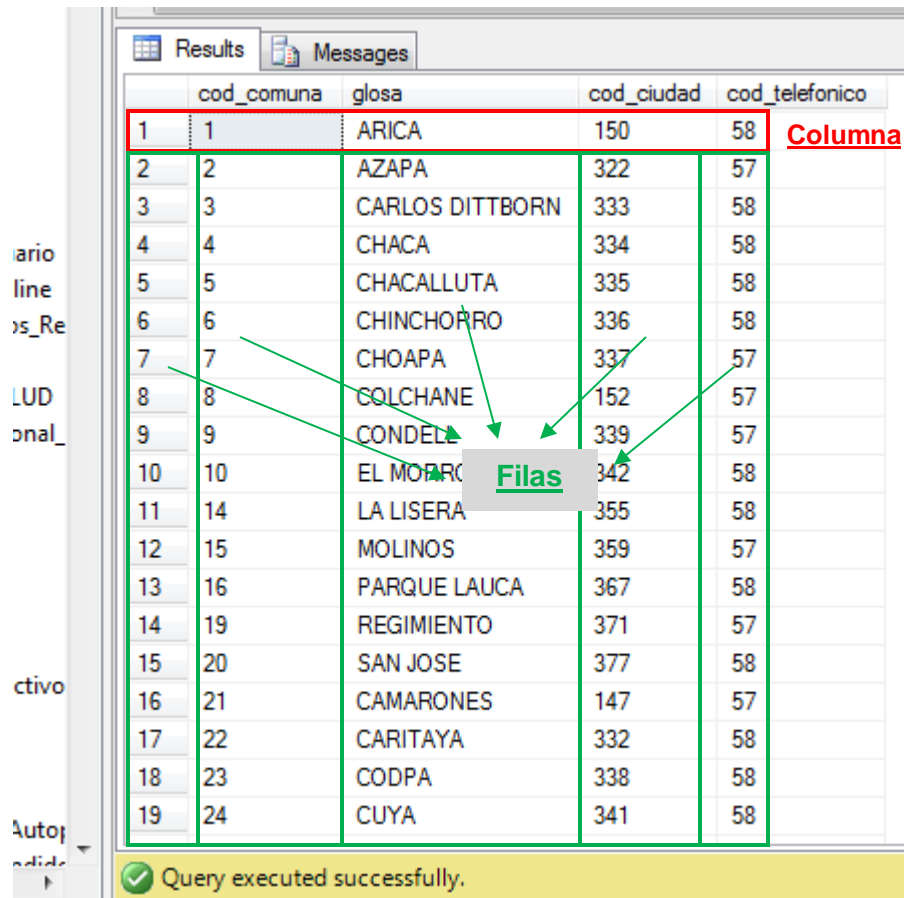
Las tablas en Base de Datos son como una hoja de Excel. Las filas de la tabla son el equivalente a las filas de las planillas Excel, las columnas de la tabla equivalen a las

columnas de una planilla Excel y el campo en una tabla, representa a una celda en la planilla.

The image shows the Microsoft Excel interface with a data table. The table has four columns: 'cod\_comuna', 'glosa', 'cod\_ciudad', and 'cod\_telefonico'. The first row of data is highlighted with a red border. A green box labeled 'FILAS' is positioned over the first column of data, with green arrows pointing to the first four rows of data. The word 'Columnas' is written in red in the right margin of the table area.

	A	B	C	D	E
1	cod_comuna	glosa	cod_ciudad	cod_telefonico	
2	1	ARICA	150	58	<b>Columnas</b>
3	2	AZAPA	322	57	
4	3	CARLOS DITTBORN	333	58	
5	4	CHACA	334	58	
6	5	CHACALLUTA	335	58	
7	6	CHINCHORRO	336	58	
8	7	CHOAPA	337	57	
9	8	COLCHANE	152	57	
10	9	CONDELL	339	57	
11	10	EL MORRO	342	58	
12	14	LA LISERA	355	58	
13	15	MOLINOS	359	57	
14	16	PARQUE LAUCA	367	58	
15	19	REGIMIENTO	371	57	
16	20	SAN JOSE	377	58	
17	21	CAMARONES	147	57	
18	22	CARITAYA	332	58	
19	23	CODPA	338	58	
20	24	CUYA	341	58	
21	26	CAMINA	145	57	
22	27	AICO	320	57	
23	28	CARIQUIMA	331	58	

Imagen 2, Hoja de cálculo en Excel.



cod_comuna	glosa	cod_ciudad	cod_telefonico
1	ARICA	150	58
2	AZAPA	322	57
3	CARLOS DITTBORN	333	58
4	CHACA	334	58
5	CHACALLUTA	335	58
6	CHINCHORRO	336	58
7	CHOAPA	337	57
8	COLCHANE	152	57
9	CONDELLA	339	57
10	EL MORO	342	58
11	LA LISERA	355	58
12	MOLINOS	359	57
13	PARQUE LAUCA	367	58
14	REGIMIENTO	371	57
15	SAN JOSE	377	58
16	CAMARONES	147	57
17	CARITAYA	332	58
18	CODPA	338	58
19	CUYA	341	58

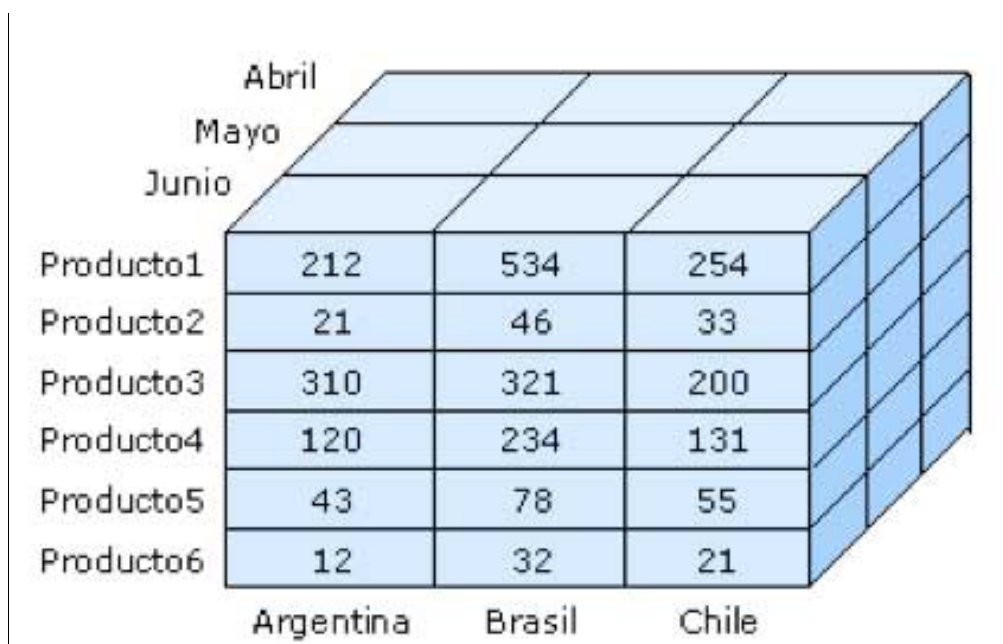
Query executed successfully.

Imagen 3, Tabla de una Base de Datos Relacional.

## II.4 TECNOLOGÍA OLAP

El procesamiento analítico en línea OLAP, (*acrónimo en inglés: On-Line Analytical Processing*) define a una tecnología que se basa en el análisis multidimensional de los datos y que le permite al usuario tener una visión más rápida e interactiva de los mismos. Básicamente consiste en la realización de consultas a estructuras multidimensionales, también conocidas como cubos OLAP que contienen datos resumidos de grandes bases de datos o sistemas transaccionales en línea conocidos como OLTP.

OLAP nace en contraposición al procesamiento transaccional en línea OLTP (*acrónimo en inglés: On-Line Transactional Processing*). El modelo OLAP es llamado análisis multidimensional, el cual permite responder a preguntas como las siguientes: *¿Cuáles son las ventas actuales comparadas con las ventas en presupuesto, por zona, por vendedor y por producto?, ¿Cuál es nuestra rentabilidad por cliente, por producto?* Los sistemas OLAP organizan los datos directamente como estructuras multidimensionales, incluyendo herramientas para conseguir la información en múltiples y simultáneas vistas dimensionales. Otra de las características es que son rápidos en tiempos de respuesta a la hora de consultar los datos a diferencia de OLTP que es mejor para inserciones, modificaciones y eliminaciones en la base de datos.



	Abril	Mayo	Junio
Producto1	212	534	254
Producto2	21	46	33
Producto3	310	321	200
Producto4	120	234	131
Producto5	43	78	55
Producto6	12	32	21
	Argentina	Brasil	Chile

Imagen 4, Ejemplo de Cubo multidimensional (Vitt, Luckevich, Misner. 2002).

#### Algunas ventajas de OLAP:

- Es un modelo de datos intuitivo y multidimensional que facilita la selección, recorrido y exploración de los datos.

- Posee un lenguaje analítico de consulta que proporciona la capacidad de explorar las complejas relaciones existentes entre los datos empresariales.
- Genera cálculos previos en los datos a consultar esto permite una respuesta más eficiente a las consultas ad hoc.

Existen diversos tipos de implementaciones de la tecnología OLAP y pueden variar según el tipo de motor de base de datos, estos pueden ser:

**MOLAP:** Procesamiento Analítico Multidimensional en línea (*Multidimensional On-line Analytical Processing*), la diferencia más significativa es que almacena los datos en una matriz de almacenamiento multidimensional optimizada, en vez de una base de datos relacional.

**ROLAP:** Procesamiento Analítico en línea Relacional (Relational On-line Analytical Processing), en síntesis las herramientas ROLAP accede a datos desde una base de datos relacional y generan consultas en SQL para calcular la información al nivel apropiado. Con ROLAP, es posible crear tablas de base de datos adicionales (tablas resumen o agregaciones) las cuales resumen los datos en cualquier combinación deseada de dimensiones.

**HOLAP:** Procesamiento Analítico en línea híbrido (Hybrid On-line Analytical Process), es una combinación entre MOLAP y ROLAP. En resumen es almacenar los datos agregados y pre-calculados en forma multidimensional, mientras que los datos que tienen un menor nivel de detalle se almacenan en bases de datos relacionales.



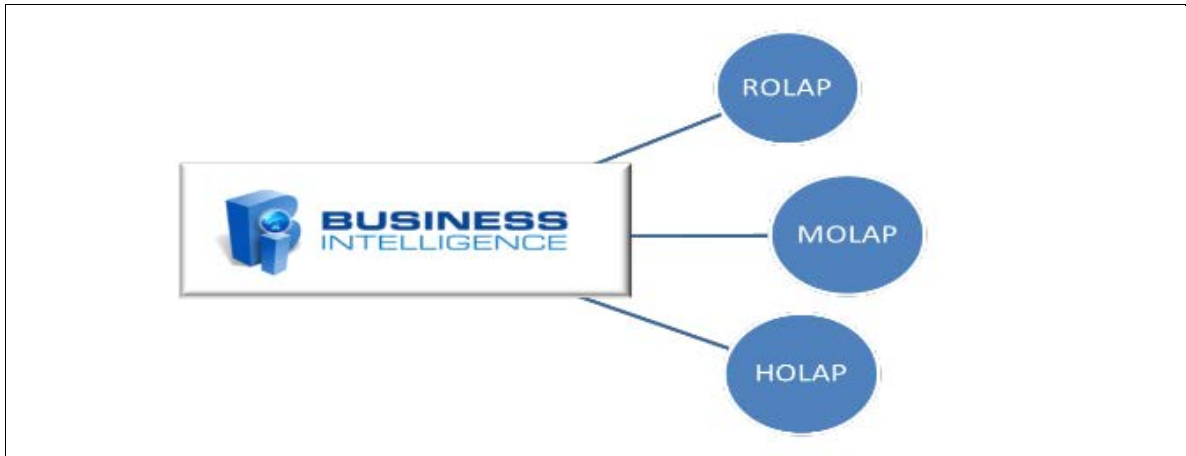


Imagen 5, Distintas Implementaciones de las Tecnologías OLAP.

## II.5 TECNOLOGÍA OLAP VS TECNOLOGÍA OLTP

Los sistemas transaccionales (OLTP) y las aplicaciones con tecnología OLAP son muy distintos en cuanto a sus requerimientos de diseño y sus características de operación. Mientras que las OLTP están organizadas para realizar transacciones como transferencias de fondos bancarias, un OLAP está orientado a los conceptos, como ventas, productos, inventario, clientes, etc.

Generalmente los sistemas OLAP posee un número reducido de usuarios, en cambio un sistema OLTP es mucho mayor la cantidad de usuarios conectados.

Los sistemas OLTP son muy fijos en cuanto a su estructura, mientras que los OLAP sufren cambios mucho más seguido, debido a que los tipos de consultas son muy variados y difíciles de prever.

El siguiente recuadro muestra las principales diferencias que existen entre un OLTP y OLAP



<b>Objetivo</b>	<b>Procesos Operacionales</b>	<b>Toma de Decisiones</b>
<b>Cliente</b>	Personal Operacional	Gestores del Negocio
<b>Datos</b>	Atómicos, Actualizados y Dinámicos	Consolidados, Históricos y Estables
<b>Estructura</b>	Normalizada	Dimensional
<b>Tiempo Respuesta</b>	Segundos	De segundos a minutos
<b>Orientación</b>	A las aplicaciones	A la información
<b>Acceso</b>	Alto	Moderado a bajo
<b>Actualización</b>	Continuamente	Periódicamente
<b>Aplicación</b>	Estructuras y Procesos Repetitivos	No estructurados y Procesos Analíticos

Imagen 6, Recuadro con las principales diferencias entre un OLTP vs OLAP.

Esta imagen muestra una comparación entre la tecnología OLTP y OLAP

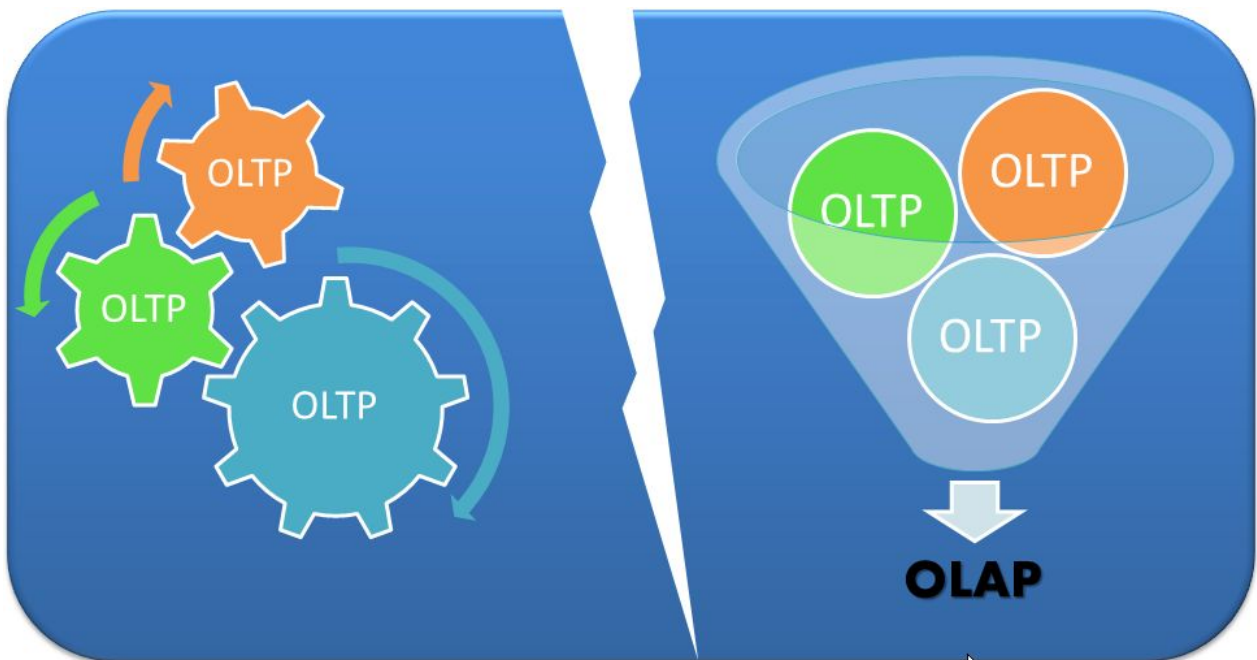


Imagen 7, Comparación visual entre un OLTP y OLAP.

## II.6 CONCEPTOS BÁSICOS DE UN ETL

### Definición de ETL

ETL viene del origen inglés Extract, Transform and Load, en español, Extraer - Transformar y Cargar, el cual corresponde al proceso que permite a las organizaciones, mover datos desde múltiples fuentes, ya sea Base de datos, planillas de cálculo, archivos de texto, etc., reformateándolos, limpiándolos y cargándolos en otra base de datos, ya sea DataMart o DataWarehouse para análisis. También puede ser utilizado para otro sistema operacional que apoya al proceso del negocio.

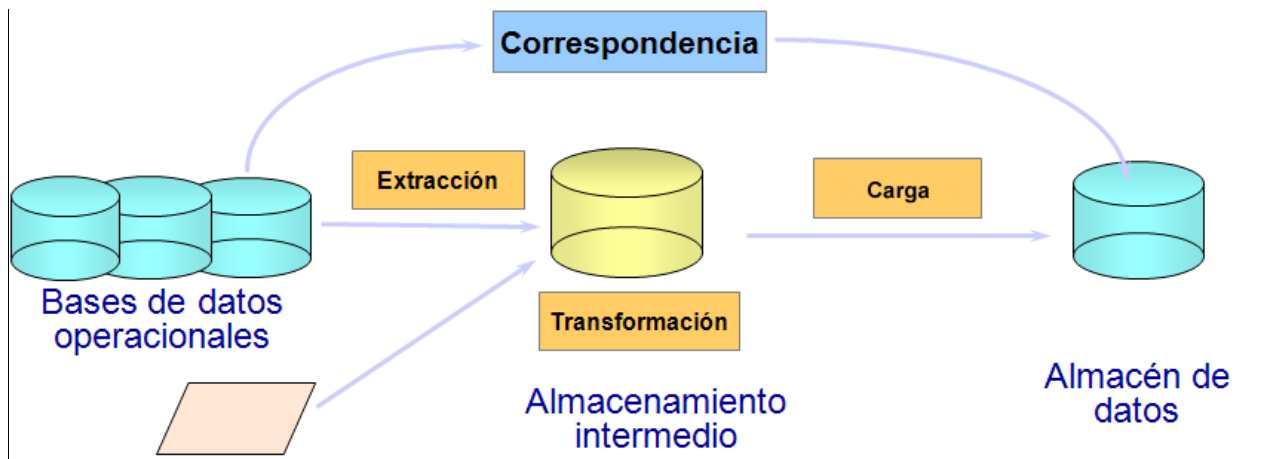


Imagen 8, Diagrama simple del proceso extracción, transformación y carga.

Hay que considerar que los procesos ETL generalmente son muy complejos. Un sistema ETL mal diseñado puede provocar importantes problemas en la fiabilidad de la información.

### Extracción

Es la primera parte del proceso ETL, y consiste en extraer los datos desde los sistemas de origen. Generalmente el origen proviene de distintos sistemas y formatos, por lo tanto la extracción convierte los datos a un formato preparado para iniciar el proceso de transformación.

## Transformación

La fase de transformación aplica una serie de reglas de negocio o funciones sobre los datos extraídos para convertirlos en datos que serán cargados. Algunas veces se requerirán pequeñas manipulaciones en los datos, sin embargo en otros casos pueden ser necesarias aplicar transformaciones como:

- Seleccionar sólo ciertas columnas para su carga (por ejemplo, que las columnas con valores nulos no se carguen).
- Traducir códigos (ejemplo, si la fuente almacena una "H" para Hombre y "M" para Mujer pero el destino tiene que guardar "1" para Hombre y "2" para Mujer).
- Codificar valores libres (ejemplo, convertir "Hombre" en "H" o "Sr" en "1").
- Obtener nuevos valores calculados (ejemplo,  $total\_venta = cantidad * precio$ ).
- Calcular totales de múltiples filas de datos (ejemplo, ventas totales de cada región).
- Generación de campos clave.
- Dividir una columna en varias (ejemplo, "Nombre: García, Miguel"; a "Nombre: Miguel" y "Apellido: García").

## Carga

La fase de carga es el momento en el cual los datos de la fase anterior (transformación) son cargados en el sistema de destino, este proceso puede abarcar una amplia variedad de acciones diferentes. En algunas bases de datos se sobrescribe la información antigua con nuevos datos.

La fase de carga interactúa directamente con la base de datos de destino. Al realizar esta operación se aplicarán todas las restricciones que se hayan definido, (por ejemplo, valores únicos, integridad referencial, campos obligatorios, rangos de valores). Estas restricciones y disparadores, si están bien definidos, contribuyen a que se garantice la calidad de los datos en el proceso ETL, y deben ser tomados en cuenta.

Algunas herramientas y aplicaciones ETL del mercado son:

- Talend Open Studio for Data Integration (Open Source)
- IBM Websphere DataStage (anteriormente Ascential DataStage y Ardent DataStage)
- Pentaho Data Integration (Open Source)
- SAS ETL Studio
- Oracle Warehouse Builder
- Informática PowerCenter
- Cognos Decisionstream
- Ab Initio
- BusinessObjects Data Integrator (BODI)
- Microsoft SQL Server Integration Services (SSIS)

## **II.7 DEFINICIÓN DE INTELIGENCIA DE NEGOCIOS**

La Inteligencia de Negocio (*del Ingles Business Intelligence o BI*) engloba el conjunto de estrategias y soluciones tecnológicas que tienen como objetivo la administración y creación de conocimiento mediante el análisis de datos existentes en una organización o empresa.

La Inteligencia de Negocio se basa principalmente en la diferenciación de los conceptos de dato, información y conocimiento: un dato es algo vago (por ejemplo "300.000"). La información se refiere a algo más preciso (por ejemplo "*Las ventas del mes de octubre fueron de \$300.000*"). En cambio el conocimiento se obtiene mediante el análisis de la información (por ejemplo "*Las ventas del mes de octubre fueron de \$300.000.-, por lo tanto, comparado con otros meses, se puede constatar que octubre es el mes más bajo en ventas*").



**Imagen 9 Diagrama para diferenciar datos, información y conocimiento**

Al igual que en el punto anterior existe el concepto de “ciclo de vida de la información” aunque en este caso enfatiza que las acciones ejecutadas a partir del conocimiento obtenido, proporciona nuevos datos que son susceptibles de volver a analizar para obtener nueva información y nuevo conocimiento, esto se llama “ciclo virtuoso” del Business Intelligence.



**Imagen 10 Ciclo Virtuoso de Business Intelligence**

Para un proyecto de inteligencia de negocio se toman como fuente los sistemas de información de la organización. Después de extraer los datos relevantes, es necesario transformarlos y cargarlos en un nuevo sistema especialmente diseñado para soportar un acceso rápido a la información, que posteriormente será analizada multidimensionalmente mediante análisis OLAP. En la siguiente figura se pueden observar gráficamente las fases y el proceso para la transformación de dato en bruto a datos analizables por los usuarios.

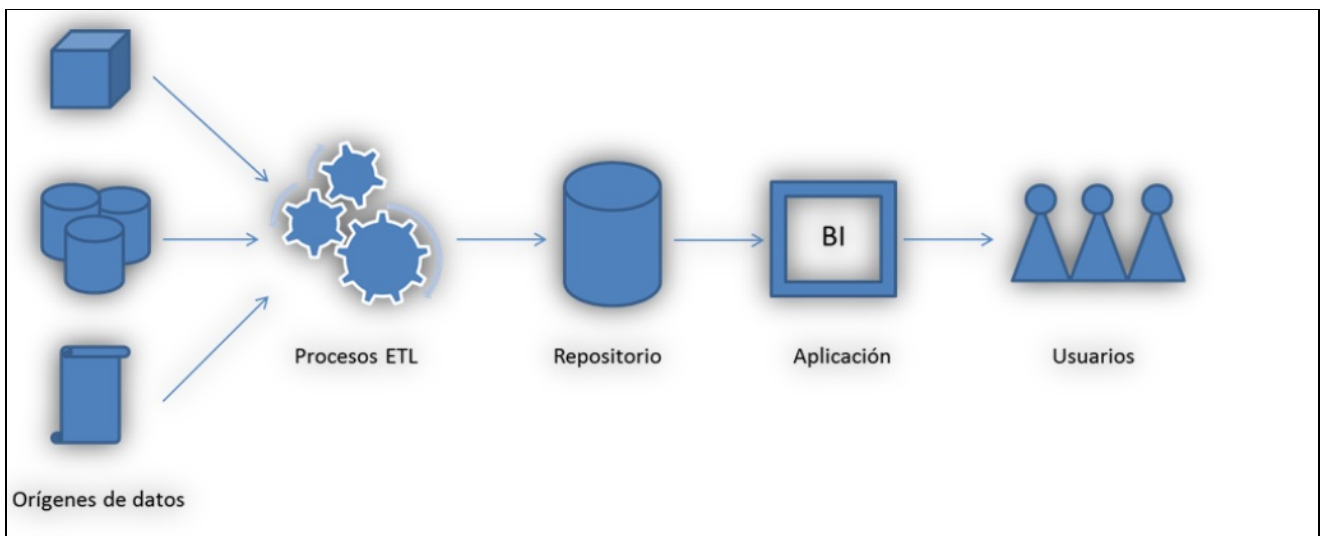


Imagen 11 Arquitectura típica de un sistema de Inteligencia de Negocio.

## II.8 CARACTERÍSTICAS DE UNA HERRAMIENTA BI

Para que una herramienta de Inteligencia de Negocios pueda ser considerada como tal, debe garantizar:

- ✓ **Accesibilidad:** garantizar el acceso de los usuarios a los datos.
- ✓ **Orientada al usuario:** buscar la independencia entre los conocimientos técnicos que puedan poseer los distintos usuarios.

- ✓ **Apoyo a la toma de decisiones:** debe poseer herramientas de análisis que permita seleccionar y manipular solamente aquellos datos que sean interesantes para el usuario.

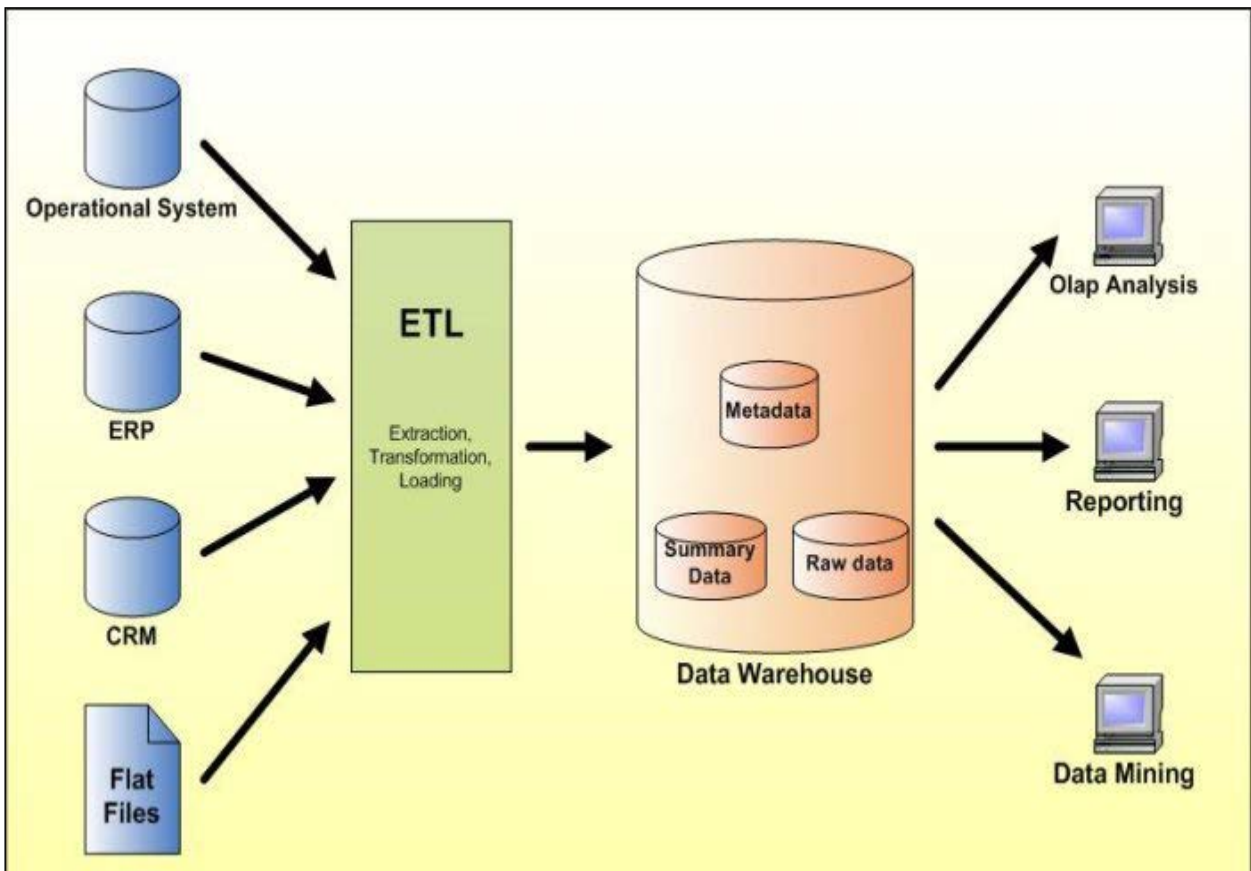


Imagen 12 Representación de la Tecnología DataWarehouse

## II.9 DATA MART

Un DataMart es una Base de Datos orientada y estructurada para el análisis, pero especializada en el almacenamiento de los datos de un área de negocio específica. 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. La información que se maneja es histórica.



Los DataMart generalmente no mantienen datos operacionales detallados, por lo que son más fáciles de entender y navegar.

Provee de los datos en una forma que concuerda con vista colectiva por un grupo de usuarios del mismo departamento o función de negocio.

A diferencia de un DataWarehouse un DataMart puede mejorar el tiempo de respuesta al usuario final debido a la reducción en el volumen de información a ser accedido.

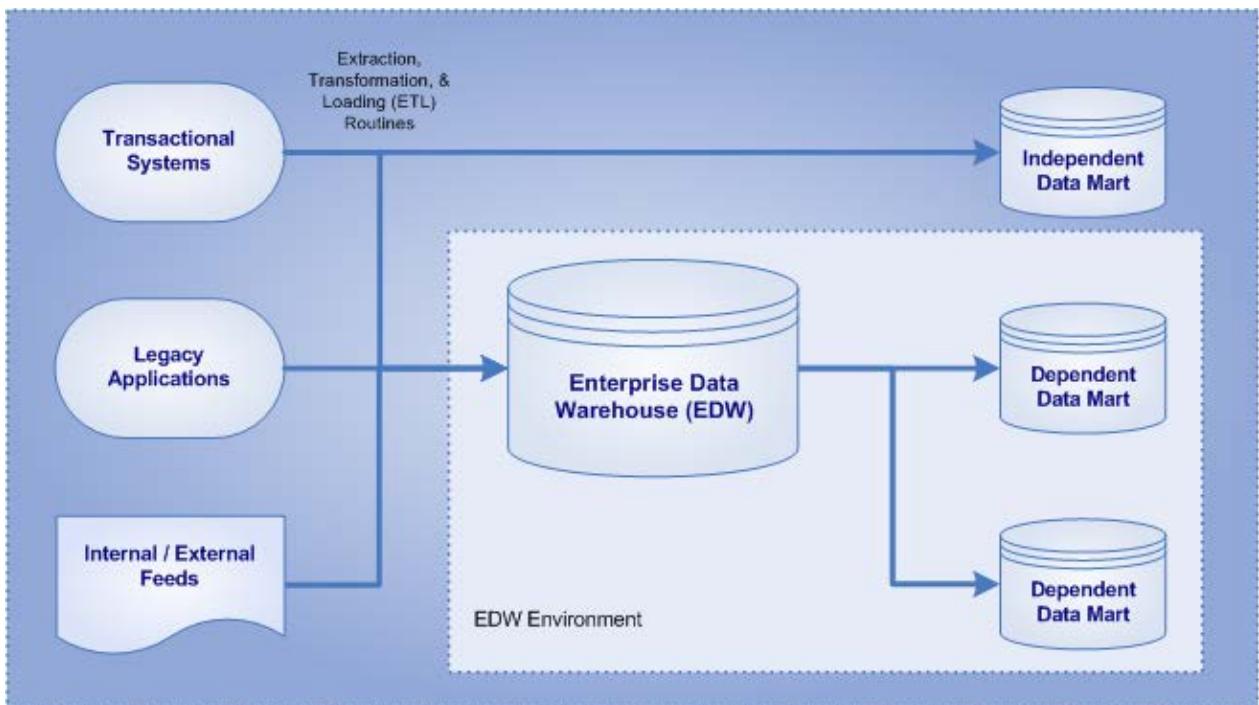


Imagen 13 Representación de un Data Mart

## II.10 DATA WAREHOUSE

La definición de un DataWarehouse según Ralph Kimball: *"es una copia de las transacciones de datos específicamente estructurada para la consulta y el análisis"*, también dijo que es *"la unión de todos los Datamart de una entidad"*.

En cambio Bill Inmon ve la necesidad de transferir la información de los OLTP a un lugar centralizado donde los datos puedan ser usados para el análisis. Escribe además con respecto a las características que posee un DataWarehouse:

Como solución a las necesidades de obtener información fiable y homogénea en términos globales dentro de la organización, surge el concepto de DataWarehouse y se basa en su concepto fundamental, la estructura de la información.

- ✓ **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.

Otra característica del DataWarehouse es que contiene “datos relativos a los datos”, concepto que se ha venido asociando al término de “Metadatos”. Los metadatos permiten mantener información de su procedencia, la periodicidad de refresco, su fiabilidad, forma de cálculo, etc.

Estos metadatos serán los que permitan simplificar y automatizar la obtención de la información desde los sistemas operacionales a los sistemas informacionales.

Los objetivos que deben cumplir los metadatos, según el colectivo al que va dirigido, serían:

- **Soportar al usuario final:** ayudándole a acceder al DataWarehouse con su propio lenguaje de negocio, indicando qué información hay y qué significado tiene. Ayudar a construir consultas, informes y análisis, mediante herramientas de navegación.
- **Soportar a los responsables técnicos del DataWarehouse:** en aspectos de auditoría, gestión de la información histórica, administración del DataWarehouse, elaboración de programas de extracción de la información, especificación de las interfaces para la realimentación a los sistemas operacionales de los resultados obtenidos, etc.

Para comprender el concepto de DataWarehouse, es importante considerar los procesos que lo conforman, estos procesos son los siguientes:

- **Extracción:** obtención de información de las distintas fuentes tanto internas como externas.
- **Elaboración:** filtrado, limpieza, depuración, homogeneización y agrupación de la información.
- **Carga:** organización y actualización de los datos y los metadatos en la base de datos relacional OLAP.
- **Explotación:** extracción y análisis de la información en los distintos niveles de agrupación a través de herramientas enfocadas a la inteligencia de negocios.

En la siguiente imagen se puede observar de manera más sencilla el proceso de un DataWarehouse y su posterior explotación.

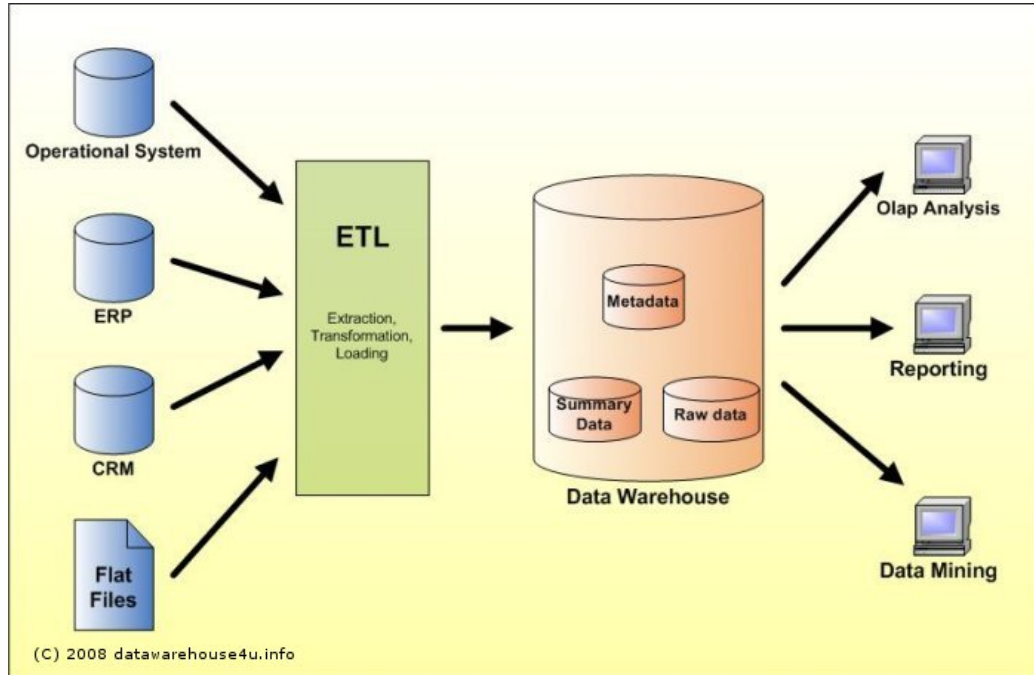


Imagen 14 Diagrama del proceso de un Data Warehouse

### En Resumen un DataWarehouse:

- ✓ **Proporciona** una herramienta para la toma de decisiones en cualquier área funcional, basándose en información integrada y global del negocio.
- ✓ **Facilita** la aplicación de técnicas estadísticas de análisis y modelización para encontrar relaciones ocultas entre los datos del almacén; obteniendo un valor añadido para el negocio de dicha información.
- ✓ **Eficiente** la capacidad de aprender de los datos del pasado y podría predecir situaciones futuras en diversos escenarios.
- ✓ **Simplifica** dentro de la empresa la implantación de sistemas de gestión integral de la relación con el cliente.
- ✓ **Supone** una optimización tecnológica y económica en entornos de Centro de Información, estadística o de generación de informes con retornos de la inversión espectaculares.

### **III. CONTEXTO**

#### **III.1 LA IMPORTANCIA DE LA INFORMACIÓN**

En la actualidad, las empresas saben que la información es uno de los activos más importantes de la empresa (Bitam, 2002), pero se comienza a tratar de manera especial aquella que tiene relación con los datos para la toma de decisiones.

Si la información es un activo, debemos poder asignarle un valor en pesos, la pregunta que surge inmediatamente es, ¿cómo podemos asignarle un valor en pesos a la información, dado un mercado libre?, la primera respuesta es, que el valor de la información es lo que el mercado pague por ella. Este recurso simple, basado en el valor percibido, muchas veces es suficiente para asignarle un valor a la información, sin embargo, no es suficiente en otros casos, por ejemplo, en el caso de una pieza de información que no vende y que es utilizada únicamente en procesos internos de toma de decisiones.

La importancia de una buena información puede ser vista como la diferencia en valor entre una decisión correcta y una decisión equivocada, en donde la decisión está basada en esa información. Mientras más grande sea la diferencia, entre decisión correcta y errónea, mayor será la importancia de contar con una buena información [Bitam, 2002].

##### **III.1.1 CIFRAS EN PETS<sup>1</sup>**

La pyme Pets no tiene ningún sistema informático a excepción de Excel y correo electrónico. Sus mayores ingresos son la venta de casas de mascotas y correderas de

---

mascotas, valores que fluctúan entre los 35.000 pesos hasta una corredera de 500.000 pesos aproximado la unidad. Si bien tienen registrados a sus clientes en planillas Excel hay perdida de información por lo tanto el dato no es confiable.

### ***III.2 SITUACIÓN ACTUAL***

De manera periódica nuevos clientes se contactan con Pets para comprar casas, correderas o insumos para sus mascotas y es común que se hagan órdenes de trabajo con medidas específicas de casas para mascotas o correderas. Estos últimos se han convertido en los principales ingresos para Pets.

Pets no tiene sistemas informáticos y tener un sistema e-commerce con soluciones de Business Intelligence es una utopía para ellos y para nosotros un gran oportunidad al tener la posibilidad de implementar tecnología de elite a una simple pyme.

```

1 select distinct T7.CODE, T5.REGION_ID, T7.REGION_ID, T2.increment_id, T8.increment_id FACTURA
2 , nombres, method modalidad_pago, round(T3.PRICE) monto,
3 IF(T3.WEIGHT=0,0,
4 IF((T2.BASE_GRAND_TOTAL - T2.SUBTOTAL_INCL_TAX)>0,
5 CASE
6 WHEN T7.CODE = 'RM' THEN (SELECT RM FROM tarifas_envio WHERE peso = round(T3.WEIGHT))
7 WHEN T7.CODE = 'I' THEN (SELECT I FROM tarifas_envio WHERE peso = round(T3.WEIGHT))
8 WHEN T7.CODE = 'II' THEN (SELECT II FROM tarifas_envio WHERE peso = round(T3.WEIGHT))
9 WHEN T7.CODE = 'III' THEN (SELECT III FROM tarifas_envio WHERE peso = round(T3.WEIGHT))
10 WHEN T7.CODE = 'IV' THEN (SELECT IV FROM tarifas_envio WHERE peso = round(T3.WEIGHT))
11 WHEN T7.CODE = 'V' THEN (SELECT V FROM tarifas_envio WHERE peso = round(T3.WEIGHT))
12 WHEN T7.CODE = 'VI' THEN (SELECT VI FROM tarifas_envio WHERE peso = round(T3.WEIGHT))
13 WHEN T7.CODE = 'VII' THEN (SELECT VII FROM tarifas_envio WHERE peso = round(T3.WEIGHT))
14 WHEN T7.CODE = 'VIII' THEN (SELECT VIII FROM tarifas_envio WHERE peso = round(T3.WEIGHT))
15 WHEN T7.CODE = 'IX' THEN (SELECT IX FROM tarifas_envio WHERE peso = round(T3.WEIGHT))
16 WHEN T7.CODE = 'X' THEN (SELECT X FROM tarifas_envio WHERE peso = round(T3.WEIGHT))
17 WHEN T7.CODE = 'XI' THEN (SELECT XI FROM tarifas_envio WHERE peso = round(T3.WEIGHT))
18 WHEN T7.CODE = 'XII' THEN (SELECT XII FROM tarifas_envio WHERE peso = round(T3.WEIGHT))
19 WHEN T7.CODE = 'XIV' THEN (SELECT XIV FROM tarifas_envio WHERE peso = round(T3.WEIGHT))
20 WHEN T7.CODE = 'XV' THEN (SELECT XV FROM tarifas_envio WHERE peso = round(T3.WEIGHT))
21 END,0)) AS costo_envio_p
22 , round(qty_invoiced) as Cantidad, T2.created_at fecha_trx
23 FROM opbe_sales_flat_order as T2
24 LEFT JOIN opbe_sales_flat_order_item T3 ON T3.ORDER_ID = T2.ENTITY_ID
25 LEFT JOIN opbe_sales_flat_order_payment T4 ON T3.ORDER_ID = T4.PARENT_ID
26 LEFT JOIN opbe_sales_flat_order_address T5 ON T2.ENTITY_ID = T5.PARENT_ID
27 LEFT JOIN opbe_directory_country_region T7 ON T7.REGION_ID = T5.REGION_ID
28 LEFT JOIN Listado_Usuarios vista ON upper(trim(T2.customer_email)) = upper(trim(vista.mail))
29 left join opbe_sales_flat_invoice T8 on T8.order_id = T2.entity_id
30 Left join marca_carro T9 on T9.INCREMENT_ID = T2.INCREMENT_ID
31 where T2.created_at between '2016-12-01 00:00:00' and '2016-12-31 23:59:59'
32 and T2.status = 'processing'
33 and T2.base grand total > 1000
    
```

CODE	REGION_ID	REGION_ID	increment_id	FACTURA	nombres	modalidad_pago	monto	costo_envio_p	Cantidad	fecha_trx
NULL	NULL	NULL	1400237475	1400003665	Eleodoro Franci...	transbank	31500	0	2	2016-12-15 11:27:37
RM	497	497	1400237475	1400003665	Eleodoro Franci...	transbank	31500	0	2	2016-12-15 11:27:37
NULL	NULL	NULL	1400237531	1400003666	yesica	transbank	13990	0	1	2016-12-15 12:47:16
RM	497	497	1400237531	1400003666	yesica	transbank	13990	0	1	2016-12-15 12:47:16
V	489	489	1400237557	1400003667	claudio	transbank	47600	3990	1	2016-12-15 13:29:22
RM	497	497	1400237589	1400003668	cristian	botonpago	25200	2990	1	2016-12-15 14:14:38

Result 14 ×

Output

Action Output

Time	Action	Message
40 19:08:25	select distinct T7.CODE, T5.REGION_ID, T7.REGION_ID, T2.increment_id, T8.increment_id FACTURA, replace(upper(trim(rut)), '.', '') rut, nombres, apellid...	412 row
41 19:11:52	select distinct T7.CODE, T5.REGION_ID, T7.REGION_ID, T2.increment_id, T8.increment_id FACTURA, nombres, apellidos, mail email, sku, method moda...	412 row
42 19:14:05	select distinct T7.CODE, T5.REGION_ID, T7.REGION_ID, T2.increment_id, T8.increment_id FACTURA, nombres, apellidos, method modalidad_pago, ro...	399 row
43 19:15:47	select distinct T7.CODE, T5.REGION_ID, T7.REGION_ID, T2.increment_id, T8.increment_id FACTURA, nombres, method modalidad_pago, round(T3.PR...	399 row

Imagen 15 Ejemplo de una consulta SQL para un informe

	A	B	C	D	E	F	G	H	I	J	K
	CODE	REGION	REGION	increment	FACTURA	nombres	modalidad_pa	mon	costo_envio	Cantid	fecha_trx
5	RM	497	497	1400226526	1400003496	Fernado	transbank	245000	4490	1	01-12-2016 16:05
6	XIV	498	498	1400226539	1400003497	DIEGO ANDRES	transbank	245000	7990	1	01-12-2016 16:24
7	RM	497	497	1400226540	1400003498	Andres	transbank	245000	0	1	01-12-2016 16:25
8	NULL	NULL	NULL	1400226544	1400003499	ivan	transbank	245000	0	1	01-12-2016 16:39
9	RM	497	497	1400226544	1400003499	ivan	transbank	245000	0	1	01-12-2016 16:39
11	XI	495	495	1400226588	1400003501	claudio	transbank	245000	9990	1	01-12-2016 17:31
12	NULL	NULL	NULL	1400226596	1400003502	Emanuel Eduardo	transbank	245000	0	1	01-12-2016 17:42
13	RM	497	497	1400226596	1400003502	Emanuel Eduardo	transbank	245000	0	1	01-12-2016 17:42
14	VII	491	491	1400226633	1400003503	paula	transbank	245000	6990	1	01-12-2016 18:34
18	XV	499	499	1400226677	1400003506	jeannette amalia	transbank	245000	9990	1	01-12-2016 19:45
19	IX	493	493	1400226690	1400003507	waldo arturo	transbank	245000	7990	1	01-12-2016 20:04
20	RM	497	497	1400226691	1400003508	gabriela	transbank	245000	4490	1	01-12-2016 20:07
21	RM	497	497	1400226698	1400003509	alejandro	transbank	245000	0	1	01-12-2016 20:15
22	V	489	489	1400226859	1400003510	Fabiola	transbank	11990	3990	1	02-12-2016 1:55
23	RM	497	497	1400227235	1400003511	leonor	transbank	21700	0	1	02-12-2016 13:35
26	RM	497	497	1400227314	1400003514	monica	transbank	21700	2990	1	02-12-2016 15:30
27	VIII	492	492	1400227317	1400003515	CLAUDIA	transbank	11990	4990	1	02-12-2016 15:31
28	NULL	NULL	NULL	1400227353	1400003516	Eduardo	transbank	21700	0	2	02-12-2016 16:27
29	RM	497	497	1400227353	1400003516	Eduardo	transbank	21700	0	2	02-12-2016 16:27
30	V	489	489	1400227370	1400003517	MARIA SOLEDAD	transbank	21700	3990	2	02-12-2016 16:47
37	IX	493	493	1400227607	1400003522	freddys	transbank	239990	7990	1	02-12-2016 22:44
39	RM	497	497	1400228359	1400003524	Kareen	transbank	28700	0	1	03-12-2016 22:26
40	RM	497	497	1400228359	1400003524	Kareen	transbank	11200	0	1	03-12-2016 22:26
41	RM	497	497	1400228359	1400003524	Kareen	transbank	38500	0	1	03-12-2016 22:26

Imagen 16 Ejemplo en Excel de Informe generado

### III.3 PROPUESTA

Hacer de Pets una tienda On Line 24 x 7 que este operativo los 365 días del año, por medio de este canal queremos incrementar las ventas y cambiar la forma de pago de esta misma, es decir: Al hacer una venta On Line Pets tendrá en su caja el dinero antes de tener el producto fabricado, de esta manera los productos vendidos por Pets por el canal On Line recibirán el pago antes de fabricar el producto final, cambiando así la forma de venta que venían operando “sin sistema e-commerce”. Además se propone implementar un servicio Business Intelligence que reporte vía email, reportes automáticos de periodicidad diario con las ventas realizadas y un reporte de consolidado mensual de todo el sitio e-commerce, es decir, las visitas al sitio completo, visitas al home, histórico de clientes registrados, etc.



De esta manera se pretende:

- Implementar e-commerce Open Source community.
- Implementar Google Analytics para nuestro sitio e-commerce.
- Generar 2 Datamart genéricos:
  - ✓ Datamart Visitas.
  - ✓ Datamart Ventas.
- Generar reporte diario de Ventas.
- Generar reporte Mensual de Consolidado de Pets.
  - ✓ Visitas al Sitio completo de Pets
  - ✓ Visitas sólo al Home de Pets
  - ✓ Gráficos de visitas Desktop vs Mobile
  - ✓ Distribución de Visitas por Categorías de mascotas
  - ✓ Categoría de mascotas que genera mayor cantidad de Ventas
  - ✓ Histórico de clientes promedios Mensual
  - ✓ Gráfico de torta de ventas por canal
  - ✓ Porcentajes de Visitas vs ventas por canal
  - ✓ Ranking de productos más vendidos por categorías
  - ✓ Tiempo de permanencia en la página de registro por canal
  - ✓ Tasa de fuga en la página de registro por canal
- Por último definir un modelo de los datamart estándar, debido a que en el futuro queremos integrar a más empresas que requieran nuestras soluciones e-commerce y Business Intelligence para pymes.

### ***III.4 BENEFICIARIOS***

El principal beneficiario de esta iniciativa es principalmente el emprendedor dueño de la pyme Pets, no obstante, nuestro objetivo es de carácter social y queremos que todas las pymes de Chile tengan la posibilidad de implementar tecnología e-commerce y Business Intelligence Open Source community a un precio razonable.

### ***III.5 ORGANIZACIÓN DEL PROYECTO***

Las áreas que intervienen en la ejecución del proyecto son:

- 1) Administración de la pyme: Principalmente su rol es recopilar imágenes de los productos que desean vender en el portal, además de capacitación en la administración de Magento y herramientas de BI.

## **IV. *DESARROLLO DEL PROYECTO***

### ***IV.1 PLANIFICACIÓN Y ADMINISTRACIÓN DEL PROYECTO***

La planificación se centró en la máxima estandarización de las soluciones e-commerce y Business Intelligence, algunos puntos altamente importantes en el ciclo de vida son:

En esta primera etapa sólo configurar productos simples en el e-commerce

Generar Datamart estándares para los Datamart para en un futuro cercano implementar más empresas pymes en estos mismos datamart haciendo de esta solución una solución aplicable a más y más pymes

### IV.1.1 METODOLOGÍA UTILIZADA

La metodología utilizada es cascada.

Al crear productos simples con Magento nos aseguramos que en esta primera fase los ETL no serán tan complejos y podrán ser reutilizados para que otras pymes se integren en el futuro.

Por otro lado, los datamart son implementados y estandarizados para el modelo EAV de Magento y el procedimiento de implementación para nuevas pymes será repetitivo.

Creemos que la metodología cascada es la metodología que mejor nos representa para este tipo de proyectos.

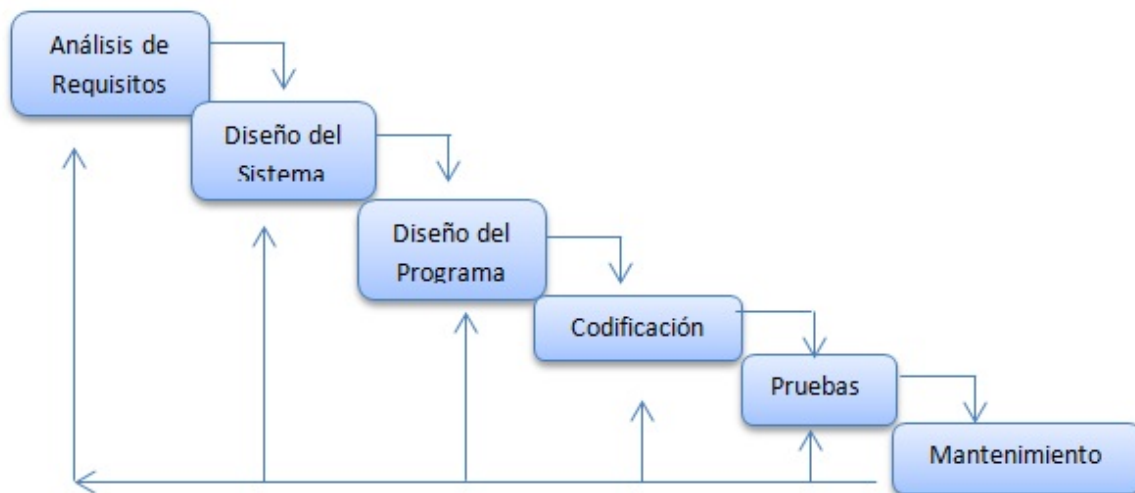


Fig. 1 fases del modelo en cascada

Imagen 17 Fases de la metodología de cascada

### IV.1.2 PLANIFICACIÓN DEL PROYECTO

Las tareas que conforman el proyecto, tienen establecidas un tiempo estimado para cada fase, planeamos dos fases, la primera fase corresponde a la implementación del e-commerce y la segunda fase corresponde a la implementación del Business Intelligence. La realización de dichas tareas tiene un carácter secuencial y en otras de forma paralela. La siguiente tabla muestra las etapas del proyecto identificando los hitos más importantes.

Nombre de tarea	Duración
<b>E-commerce y BI para Pymes</b>	<b>30 días</b>
<b>Planificación y Análisis del requerimiento</b>	<b>2 días</b>
Definición del proyecto	1 día
Definición de los requerimientos	1 día
<b>Diseño de la solución e-commerce</b>	<b>1 día</b>
Diseño de la Arquitectura Técnica	1 día
<b>Construcción y Pruebas e-commerce</b>	<b>10 días</b>
Instalación de Magento 1.9	0,5 días
Personalización de Magento (Customización)	2 días
Creación de Categorías y productos	1 día
Implementar Google Analytics	1 día
Implementar botón de pago	1 día
Implementar Slider en Magento	1,5 días
Pruebas Funcionales y de desempeño	3 días
<b>Implementación e-commerce</b>	<b>6,5 días</b>
Paso a Producción (Go-Live)	2 días
Entrenamiento a Usuarios Finales	1,5 días
Pruebas de Aceptación	2 días
Cierre	1 día
<b>Diseño de la solución BI</b>	<b>5 días</b>
Análisis de las Fuentes de Datos	2 días
Modelado Dimensional	3 días
<b>Construcción y Pruebas BI</b>	<b>12,5 días</b>
Instalación de Productos	1 día
Diseño Físico	0,5 días

---

Extracción, Transformación y carga (ETL)	5 días
Desarrollo de Reportes	3 días
Pruebas Funcionales y de desempeño	3 días
<b>Implementación BI</b>	<b>5 días</b>
Paso a Producción (Go-Live)	1 día
Entrenamiento a Usuarios Finales	1 día
Pruebas de Aceptación	2 días
Cierre	1 día

El total de días se estableció utilizando como base un día hábil que consta de ocho horas diarias de trabajo de lunes a viernes, la semana hábil comprende un total de cuarenta y cinco horas hombre.

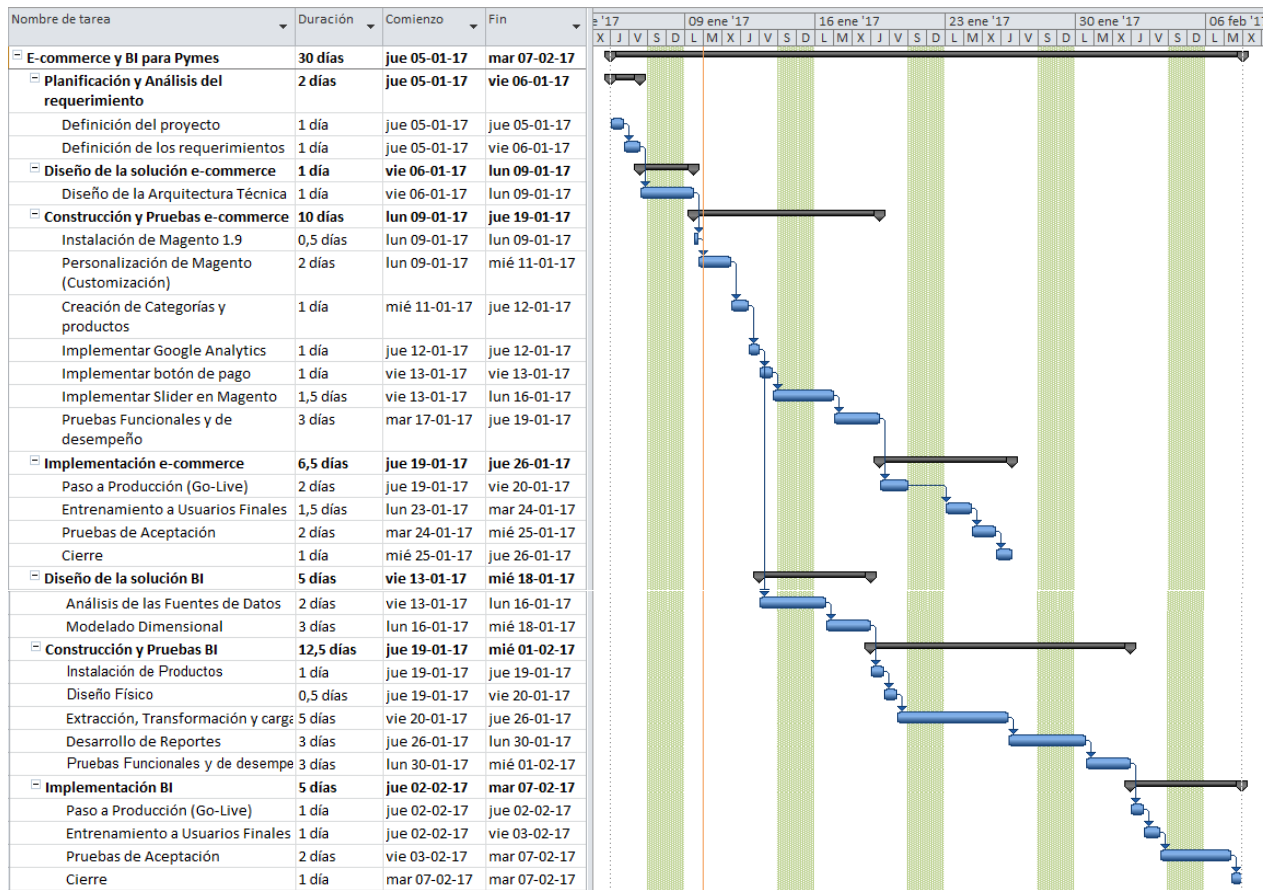


Imagen 18 Carta Gantt del Proyecto

### IV.1.3 DEFINICIÓN DEL PROYECTO

La definición del proyecto, corresponde al arranque de este mismo, se incluye básicamente la identificación de los interesados del negocio. Los usuarios finales y los expertos del negocio ayudaron y son aportes significativos en la definición conceptual de la plataforma, aportando con información técnica de los productos a publicar en el e-commerce y fotografías de estos mismos.

### **IV.1.3.1 DEFINICION DE LOS REQUERIMIENTOS**

El requerimiento debería redefinirse como definición de la solución y son básicamente un portal e-commerce y una solución Business Intelligence.

Para esto se propone reuniones con el emprendedor y/o los stakeholders de la pyme para alinearse con ellos e incluirlos en la solución del portal como actores relevantes para el buen término del proyecto.

Algunas cualidades de la solución son:

- Portal e-commerce debe ser responsivo para cualquier dispositivo
- Portal debe tener botón de pago On Line
- Portal debe ser entendible fácilmente para los usuarios
- Portal debe ser administrable para los usuarios calificados de la Pyme
- Portal debe tener la capacidad de crear nuevas categorías, subcategorías y productos
- La solución Business Intelligence debe ser capaz de generar reportes automáticos y enviarlos vía e-mail a los administradores de la plataforma
- Los reportes deben ser claros y precisos.

### **IV.1.4 ESTUDIO DE FACTIBILIDAD**

El estudio de Factibilidad, fue analizada solo con sus cuatro apartados más comunes: **Técnica, Operativa, Económica-Financiera y Legal**, los cuales permitirán concluir y tomar la decisión de poner en práctica un proyecto, en este caso: El e-commerce y DataWarehouse.



#### IV.1.4.1 **FACTIBILIDAD TÉCNICA**

Es el estudio a la función, rendimiento y restricciones que puedan afectar la consecución de un sistema aceptable para la implementación y uso del e-commerce, esto implica contratar un web hosting para alojar el portal e-commerce y la solución Business Intelligence. Dejando en manos de terceros la seguridad del Portal.

En primer lugar se procedió a calcular la capacidad y establecer el perfil técnico de los elementos, en función de los datos y aplicaciones que serán procesados por la futura instalación. También se necesitó determinar que aplicaciones son necesarias para el desarrollo del proyecto de e-commerce y BI.

En la siguiente tabla, se identificaron los elementos considerados necesarios para el proyecto e-commerce e Inteligencia de negocios.

Categoría	Elemento	Cant.	Consideraciones	Requerimientos
Hardware	Servidor BD	1	- Escalable en Procesador, memoria y disco duro - Backup diario - Con seguridad UPS.	16 GB Memoria Procesador 1,8Ghz, 8 Core. 350GB Disco Duro. S.O. RHEL 6.1
	Servidor BI	1	- Escalable en Procesador, memoria y disco duro - Backup diario - Con seguridad UPS.	- 16 GB Memoria - Procesador 1,8Ghz, 8 Core. - 1TB Disco Duro. - S.O. RHEL 6.1
	Servidor Web	1	Copias de seguridad 3 veces al día	4GB RAID SSD, transferencia mensual ilimitado
Software	Base de datos Analítica	1	Postgres	N/A
	ETL	1	Talend	N/A

	Herramienta BI	1	Jaspersoft	N/A
	Herramienta Modelado BD	1	Enterprise Architect	N/A

#### **IV.1.4.2 LA FACTIBILIDAD OPERATIVA**

Permite verificar si la organización está en condiciones de absorber el cambio que significa la implementación y uso de un e-commerce y BI para la cultura organizacional, entre estos podemos mencionar:

- ✓ El personal no tiene experiencia en el manejo de este tipo de sistemas.
- ✓ Los gerentes o dueño de la pyme deberán estar a la cabeza del cambio.
- ✓ Se debe agregar el factor de la resistencia natural al cambio en todos los niveles.
- ✓ Se requerirá la capacitación del personal.

Además de todo esto se deberá seguir cumpliendo con las actividades diarias para que el Negocio siga funcionando.

Como vemos no se puede trivializar el estudio de factibilidad operativa, pues este permite poder reducir el margen de riesgo del proyecto y asegurar su éxito dentro de la organización.

El e-commerce y DataWarehouse tendrá un profundo impacto en el funcionamiento de la organización, debido a que una vez implementado, la pyme contara con un canal nuevo de ventas, y un BI elemento estratégico de la empresa pues disminuye el grado de incertidumbre con la que se toman las decisiones.

Algunas de las ventajas obtenidas del uso de un e-commerce e BI y que modifica de manera positiva el desempeño de la organización son:

- ✓ Contar con un nuevo Canal de venta On Line acorde a los tiempos modernos y reservado sólo para grandes organizaciones.
- ✓ Contar con una fuente confiable de información que presenta una visión consistente de los datos e información de la empresa y su entorno.
- ✓ Los reportes generados son enviados automáticamente vía correo electrónico.
- ✓ Permite monitorear oportunamente los indicadores del negocio.
- ✓ Habilita el análisis de tendencias, lo que permite descubrir oportunidades de utilidades para la empresa y debilidades que se deben corregir.

Todo esto implica cambios organizacionales importantes en el esquema de trabajo, estos cambios impactarán desde alta gerencia, hasta los trabajadores de cada de cada área, sin embargo, a pesar de las ventajas que presenta la introducción de tecnología de elite, la curva de aprendizaje y la resistencia al cambio son factores que se deben considerar.

#### ***IV.1.4.3 DOTACIÓN***

El personal mínimo para construir e implementar un e-commerce y una solución Business Intelligence dentro de los plazos establecidos anteriormente consta de dos consultores expertos. Al tener conocimiento y experiencia en la arquitectura de Magento, las actividades para implementar es un proceso aprendido, por lo tanto ya sabemos con anterioridad que podemos designar a un consultor para el portal e-commerce y otro consultor para la solución Business Intelligence.

Los datos proporcionados a continuación son una descripción de las funciones básicas que desempeñaremos, para construir nuestro e-commerce y solución Business Intelligence.

Función	Cargo	Tarea
Implementador de Magento	Consultor especialista en e-commerce (Magento)	1- Descarga de Magento
		2- Instalación de Magento
		3- Personalización de Magento
		4- Creación de Categorías y Productos
		5- Implementar Google Analytics
		6- Implementar botón de pago
		7- Implementar Slider Magento
		8- Pruebas funcionales y desempeño
		9- Paso a producción
		10- Entrenamiento usuarios finales
		1- Análisis fuentes de datos
Implementador de solución Business Intelligence	Consultor Especialista en Business Intelligence (Talend y Jaspersoft)	2- Modelamiento Dimensional (Datamart)
		3- Descarga e Instalación de productos BI
		4- Diseño físico
		5- Desarrollador ETL
		6- Orquestación de procesos
		7- Desarrollo de Reportes
		8- Automatización de reportes
		9- Pruebas funcionales y desempeño
		10- Paso a producción
		11- Entrenamiento usuarios finales

Según nuestra experiencia, estos cargos son los mínimos para implementar un proyecto e-commerce e Inteligencia de negocios, sin embargo, una sola persona puede realizar todo el proyecto, tal como se está realizando en esta tesis, donde un consultor senior con experiencia en e-commerce y business Intelligence tiene las competencias necesarias para el desafío.

Se pueden resumir que el personal necesario para construir el e-commerce y la solución Business Intelligence es un grupo base reducido mínimo de dos consultores expertos, lo cual implica un costo menor en el presupuesto.

#### **IV.1.4.4 FACTIBILIDAD ECONÓMICO-FINANCIERA**

En la evaluación del costo de desarrollo del proyecto, sopesado con los ingresos netos o beneficios obtenidos del producto a desarrollar, en este caso un e-commerce y una solución business Intelligence, la justificación económica incluye una amplia gama de aspectos a tener en cuenta como son el análisis de costo-beneficios, las estrategias de ingresos de la empresa a largo plazo, costo de recursos necesarios para el desarrollo y crecimiento potencial del mercado.

Para efectuar el estudio de Factibilidad económico-financiera resultó imprescindible conocer la totalidad de los costos en que debe incurrir la empresa, para incorporar el nuevo sistema e-commerce y Business Intelligence.

#### **IV.1.4.5 COSTOS DEL PERSONAL**

<b>Función</b>	<b>Personal</b>	<b>Tiempo (mes)</b>	<b>Sueldo (mes) Expresado en M\$</b>	<b>Costo Total Expresado en M\$</b>
<b>Consultor e-commerce</b>	<b>Jefe de Proyecto</b>	<b>1</b>	<b>1.500</b>	<b>1.500</b>
Implementador de Magento	Incluido	-	0.000	0.000
<b>Consultor Business Intelligence</b>	<b>Consultor Senior</b>	<b>0,8</b>	<b>1.500</b>	<b>1.500</b>
Desarrollador solución BI	Incluido	-	0.000	0.000
Desarrollador de Reportes	Incluido	-	0.000	0.000
			<b>Total</b>	<b>3.000</b>

Tabla realizada de acuerdo a lo óptimo según estudios realizados, sin embargo una vez implementada la solución Business Intelligence esta será capaz de integrar a más y más pymes, por lo tanto quedaran disponibles las hh del Consultor Business Intelligence y este tendrá que prestar apoyo a la implementación de nuevos e-commerce.

#### **IV.1.4.6 COSTOS DE SOFTWARE Y HARDWARE**

Estos requerimientos son esenciales para el desarrollo del proyecto. Estos requerimientos son de carácter físico (hardware) y lógico (software) y sirven para dar soporte al e-commerce y la solución Business Intelligence.

<b>Elemento</b>	<b>Cantidad</b>	<b>Valor Unitario Miles de pesos</b>	<b>Valor Total Miles de pesos\$</b>
Hosting e-commerce y BI	1	\$65.000	\$65.000
Servidor BD	-	\$0	\$0
Sistema Operativo	-	\$0	\$0
Licencia Herramienta BI	-	\$0	\$0
Soporte Base de Datos Analítica	-	\$0	\$0
Licencia Herramienta de modelamiento (Enterprise Architect)	-	\$0	\$0
Curso Herramienta BI	-	\$0	\$0
Implementación, consultoría y pruebas de concepto	-	\$0	\$0
<b>Total</b>			<b>\$65.000</b>

#### **IV.1.4.7 COSTOS INDIRECTOS**

Los costos que afectan de manera indirecta al proceso productivo de la empresa, pero que no se puede asignar directamente de acuerdo a algún criterio en específico.

- ✓ Aumento de recursos (H.H.) para mantención/soporte de aplicaciones y desarrollo.
- ✓ Aumento de recurso (H.H.) para la administración de la infraestructura tecnológica.
- ✓ Upgrade de hardware para uso de aplicaciones a un número determinado de usuarios.
- ✓ Aumento en capacitaciones para personal del negocio y también de tecnología.

#### **IV.1.4.8 ANÁLISIS COSTO V/S BENEFICIO**

¿Por qué una pyme necesita presencia online?

De acuerdo a un estudio de red.es, cada vez son más las personas que buscan información y compran a través de la web diversos productos online. Si bien podríamos distinguir entre tres principales grupos donde la penetración del comercio electrónico es bien distinta, siendo el sector turismo el principal grupo, seguido de otros servicios como venta de entradas, servicios financieros o electrónica, y un último grupo conformado por ventas de artículos como ropa, electrodomésticos o casas para mascotas el crecimiento en todos ellos es similar a medida que madura el sector del e-commerce.

¿Es rentable el comercio electrónico para pymes?

Eso depende de cómo se aborde, un alto porcentaje (en torno al 80%) de las pymes tradicionales (que carecen de presencia online) fracasan con menos de 3 años de vida, en Internet la cifra de tiendas online que fracasan sea probablemente más elevada, debido a la alta competencia.

Podríamos decir que montar una tienda online es algo sencillo, ya que cualquier webmaster con experiencia podría crear una web para vender tus productos o servicios, pero montar una tienda online rentable es algo mucho más difícil, que requiere de un asesoramiento en temas tan dispares como optimización de posicionamiento en buscadores, presencia en redes sociales, marketing en internet o logística del producto, que en conjunto determinarán el margen de venta por producto o servicio que vendas y por tanto, la viabilidad de tu proyecto e-commerce.

Si no conoces los factores relevantes para montar un comercio electrónico rentable, estás condenado al fracaso.

¿Cómo puedo determinar la rentabilidad de un comercio electrónico?

De manera parecida a como se haría para un negocio tradicional, con una proyección de ingresos y gastos, que se ajuste a tu estrategia online y a los principales factores que la afectan.





Supuesto: En siguiente embudo podemos ver que existe un ingreso de 4,9 millones de pesos aproximado, son 97 ventas de casas para mascotas para un periodo mensual. Son 97 ventas de nuevos clientes web a los que ya se le tienen sus datos para futuras campañas u marketing para potenciar las ventas.

Este dato es capturado y almacenado por la solución Business Intelligence y acá podemos apreciar todo el potencial que nos entrega una solución Business Intelligence para Pymes.

#### **IV.1.4.9 LA FACTIBILIDAD LEGAL**

Para este proyecto se utilizan herramientas Open Source en sus versiones community, y no existe ningún impedimento legal para realizar este proyecto.

#### **IV.2 SELECCIÓN DE PRODUCTOS PARA E-COMMERCE Y BUSINESS INTELLIGENCE**

Para la selección de los productos se consideran sólo herramientas Open Source versión community, evaluando cada solución por su potencialidad, escalabilidad, nativos, multi idioma, multi moneda y facilidad en la implementación.

La selección de una u otra herramienta está en función de múltiples aspectos a considerar:

**¿Qué e-commerce necesita?:** Que sea escalable internacionalmente, responsivo, múltiples idiomas, múltiples almacenes y tiendas.

**¿Qué información se necesita?:** Se necesita un Datamart de Visitas y un Datamart de Ventas.

**¿Para qué se necesita la información?:** Bajo el concepto general “soporte a la toma de decisiones” se esconden múltiples necesidades como por ejemplo, contrastar que todo esté bien, analizar diferentes aspectos de la evolución de la empresa, presentar información de forma más intuitiva, comparar información en diferentes períodos de tiempo, comparar resultados, identificar comportamientos y evoluciones excepcionales, confirmar o descubrir tendencias e interrelaciones, identificar nuestros productos top del e-commerce, las url dónde se producen mayor cantidad de rebotes, etc.

**A quién va dirigida:** Organización en general, gestión, dirección, dirección estratégica, etc.

**Aspectos técnicos:** (tiempos de respuesta, integración, seguridad, etc.) y funcionales (navegación, entorno gráfico, etc.).

#### ***IV.2.1 MAGENTO E-COMMERCE***

Magento es una plataforma de código abierto para comercio electrónico escrita en PHP. Fue desarrollada con apoyo de voluntarios por Varien Inc (ahora Magento Inc), una compañía privada con sede en Culver City, California.

Varien publicó la primera versión del software el 31 de marzo de 2008. Roy Rubin, ex director general de Varien, vendió más tarde una parte sustancial de la compañía a eBay, que eventualmente la adquirió por completo y luego se escindió de la empresa.

Magento emplea el sistema de base de datos relacional MySQL/MariaDB, el lenguaje de programación PHP, y elementos de Zend Framework. Aplica las prácticas de la programación orientada a objetos y la arquitectura modelo–vista–controlador. También utiliza el modelo entidad–atributo–valor para almacenar los datos.

Magento es el proveedor de dos plataformas distintas, Magento Community Edition y Magento Enterprise Edition. También existieron las plataformas Magento Professional Edition and Magento Go.

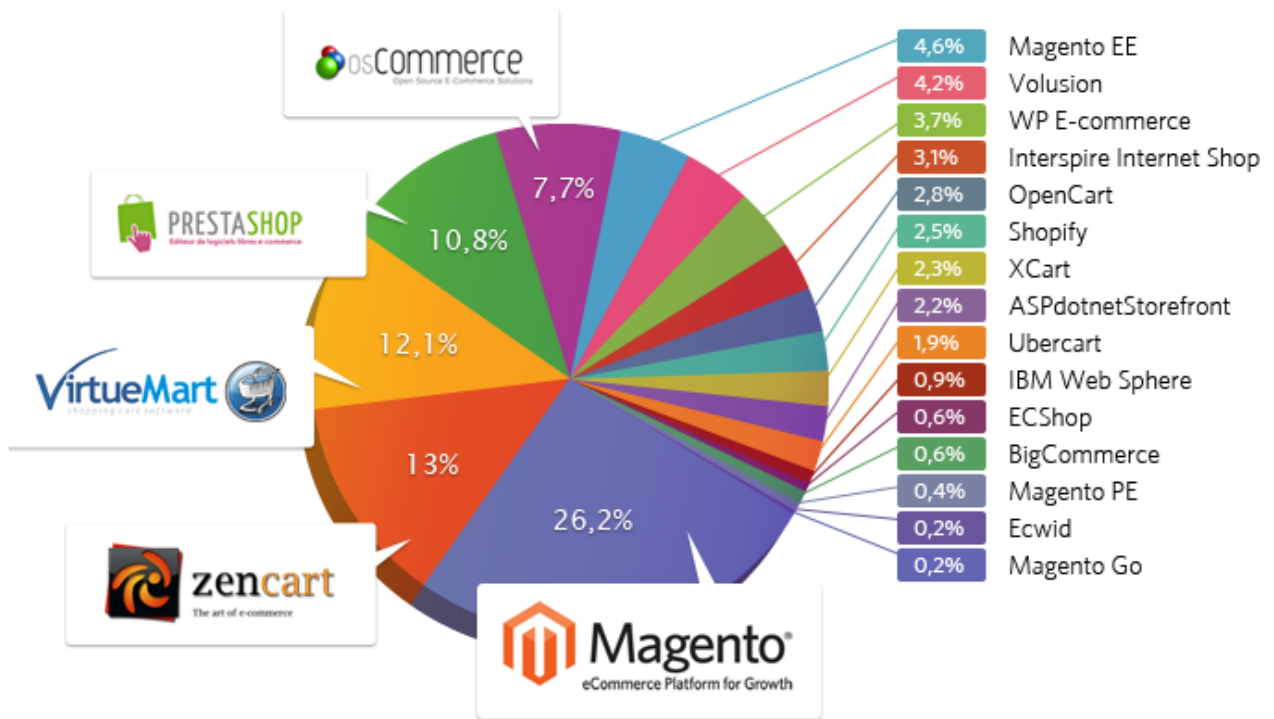


Imagen 19 Plataformas E commerce

Según el grafico el líder indiscutido es Magento y curiosamente responde con todos los requisitos de una tienda On Line para grandes ligas:

- ✓ Aplicaciones básicas para que funcione.
- ✓ Aplicaciones para que funcione.
- ✓ Formas de pago.
- ✓ Transportistas.
- ✓ Volumen de catálogo.
- ✓ Cómo va a crecer.
- ✓ ¿Queremos hacer multi tienda o no?
- ✓ ¿Vamos a tener 1.000 productos o 50.000?
- ✓ ¿Sólo vamos a vender nosotros o vamos a tener afiliados?
- ✓ ¿Cuántos idiomas queremos manejar?
- ✓ ¿Cómo manejaremos el stock (desde la tienda, desde un ERP...)?

## **IV.2.2 JASPERSOFT Y TALEND**

Jasper Reports son bibliotecas de creación de informes que tiene la habilidad de entregar contenido enriquecido al monitor, a la impresora o a ficheros PDF, HTML, XLS, CSV y XML.

Talend Studio es una solución ETL Open Source community, muy fácil de usar.

Está escrito completamente en Java y puede ser usado en gran variedad de aplicaciones de Java, incluyendo J2EE o aplicaciones web, para generar contenido dinámico. Se ha desarrollado un sub-proyecto que es un servidor integrado para informes: JasperReports Server.

Su propósito principal es ayudar a crear documentos de tipo páginas, preparados para imprimir en una forma simple y flexible.

JasperReports se usa comúnmente con iReport, un front-end gráfico de código abierto para la edición de informes, si bien a partir de la versión 5.5.0 iReport ha sido sustituido por Jaspersoft Studio, un front-end gráfico de código abierto basado en Eclipse.

Se encuentra bajo licencia libre GNU, por lo que es Software libre. Forma parte de la iniciativa apilada open source Lisog.

### **Características**

JasperReports es una biblioteca que puede ser embebida (incrustada) en cualquier aplicación Java. Sus funciones incluyen:

- ✓ Scriptlets, que pueden acompañar a la definición del informe,<sup>1</sup> y pueden ser invocados en cualquier momento por la definición para realizar un procesamiento adicional. El scriptlet se basa en Java, y tiene muchos ganchos (hooks) que se

pueden invocar antes o después de las etapas de la generación de informes, como el Informe, Página, Columna o Grupo.

- ✓ Sub-informes

Para usuarios con requisitos más sofisticados de gestión, los informes diseñados para JasperReports pueden ser fácilmente importados a JasperServer- el servidor de informes interactivos.

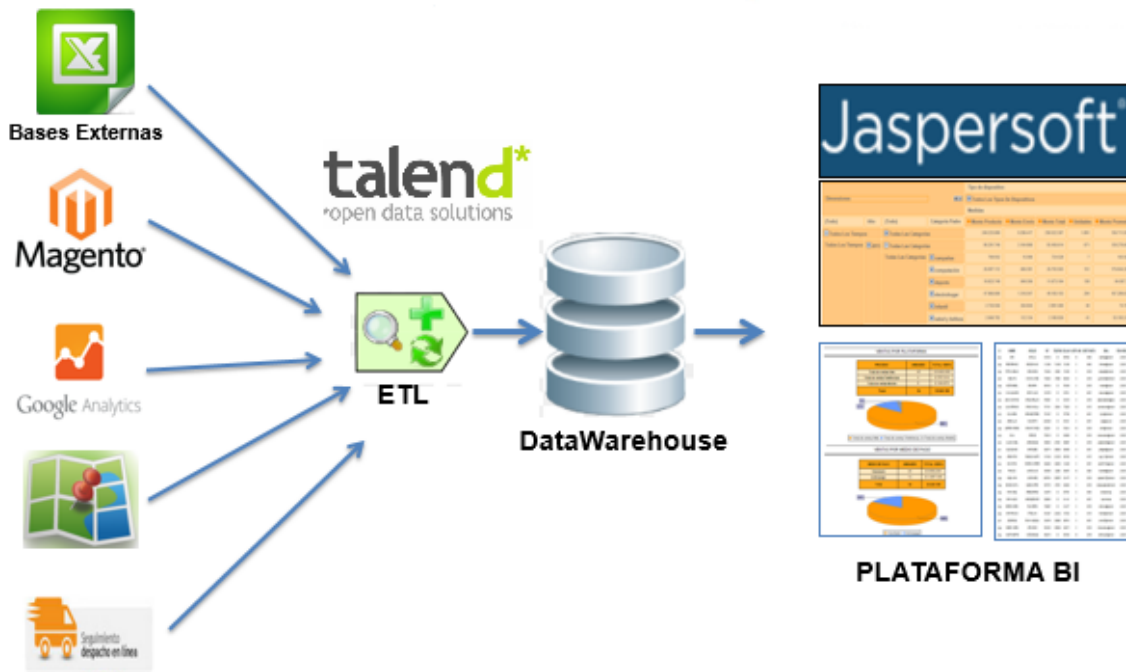


Imagen 20 Arquitectura de Talend y JasperSoft

### IV.2.3 PENTAHO BUSINESS ANALYTICS

Pentaho BI Suite es un conjunto de programas libres para generar inteligencia empresarial (Business Intelligence). Incluye herramientas integradas para generar informes, minería de datos, ETL, etc, “Nosotros nos centraremos en la Reportería”

Pentaho se define a sí mismo como una plataforma de BI “orientada a la solución” y “centrada en procesos” que incluye todos los principales componentes requeridos para implementar soluciones basadas en procesos tal como ha sido concebido desde el principio.

Las soluciones que Pentaho pretende ofrecer se componen fundamentalmente de una infraestructura de herramientas de análisis e informes integrados con un motor de workflow de procesos de negocio. La plataforma será capaz de ejecutar las reglas de negocio necesarias, expresadas en forma de procesos, actividades, además es capaz de presentar y entregar la información adecuada en el momento adecuado.

Su modelo de ingresos parece estar orientado a los servicios (soporte, formación, consultoría y soporte a ISVs y distribuciones OEM) aunque en alguno de los documentos y páginas que hemos examinado se mencionan algunas funcionalidades “Premium” que hacen pensar en ingresos por futuras versiones o funcionalidades de pago.

En su web presenta una organización por productos: Reporting, Analysis, Dashboard y DataMining, acompañado por dos introducciones: a la plataforma y a los productos. En dichas introducciones se hace mención específica al workflow como una de las capacidades BI claves de la plataforma.

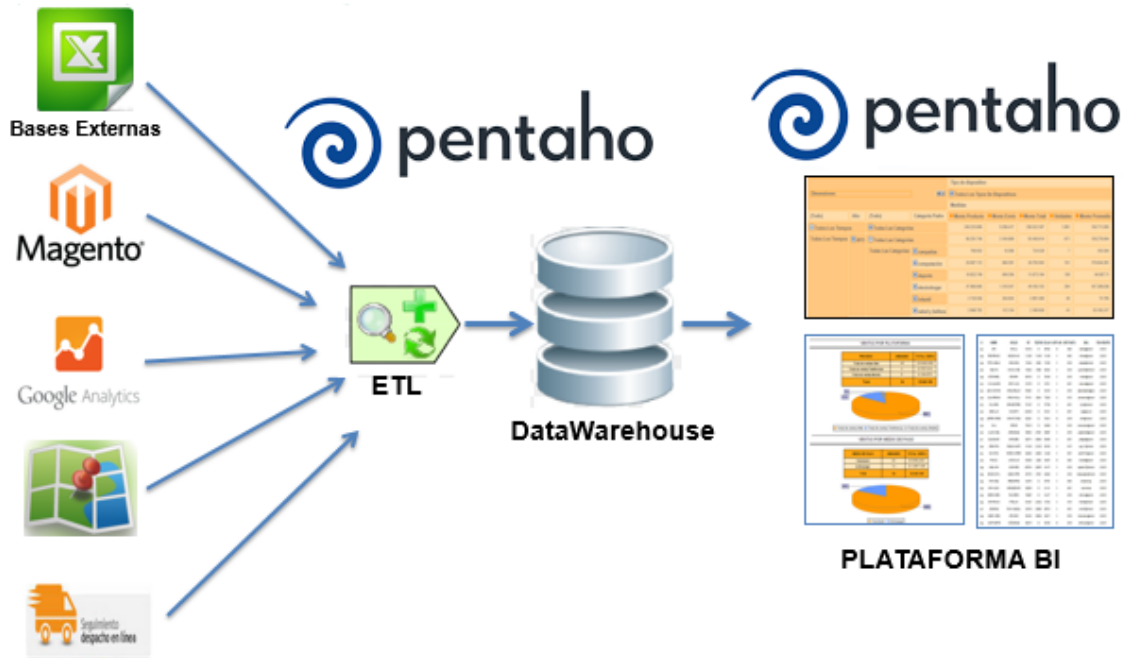


Imagen 21 Arquitectura Pentaho Analytics

## IV.2.4 ELECCIÓN DE LOS PRODUCTOS BUSINESS INTELLIGENCE

### IV.2.4.1 JASPERSOFT Y TALEND

Los productos seleccionados para el desarrollo y ejecución del proyecto de Inteligencia de Negocios es “Jaspersoft community” y Talend community, ya que al contrario de muchas herramientas, algunas por su alta complejidad de instalación y otras por restricciones propias de la empresa desarrolladora, Jaspersoft y Talend proporciona gratuitamente desde su web una versión para descargar con todas las opciones de desarrollo disponibles, debido principalmente a que es un producto Open Source. Esta versión gratuita no necesita de licencias para usuarios que necesiten implementar ETL, extraer reportes web o administrar la plataforma.



Los componentes de servidor son de hasta 1 CPU como límite, aunque se instale en un servidor con múltiples procesadores y múltiples core's el software está limitado para que trabaje con uno solo, excepto en el caso de adquirir el soporte pagado.

De esta manera tenemos la posibilidad de que en el futuro podamos potenciar nuestro DataWarehouse una vez comprobada su eficiencia sin arriesgar un gran capital destinado a la herramienta BI, siendo una de las variables más importante a la hora de implementar un Sistema de Inteligencia de negocios.

#### **IV.2.4.2    *POSTGRES SQL***

Para llegar a una decisión final, no fue muy difícil, basándonos en nuestra experiencia se decidió que Postgres cumple con todos los requisitos para nuestros Datamart, La principal razón por la cual se tomó esta decisión, pasó por un factor económico "Open Source a costo cero", ya que otros motores realizan prácticamente las mismas funciones, sin mostrar una diferencia sustancial entre los sistemas en las pruebas de carga y consulta de datos (prueba de concepto).

Además, Postgres tiene incorporado técnicas de tuning SQL de particionado para que las consultas se comporten como una función lineal independiente que la base de datos tenga volúmenes de datos considerables.

### **IV.2.4.3 GOOGLE ANALYTICS**

Google Analytics es una herramienta de analítica web de la empresa Google. Ofrece información agrupada del tráfico que llega a los sitios web según la audiencia, la adquisición, el comportamiento y las conversiones que se llevan a cabo en el sitio web.

Se pueden obtener informes como el seguimiento de usuarios exclusivos, el rendimiento del segmento de usuarios, los resultados de las diferentes campañas de marketing online, las sesiones por fuentes de tráfico, tasas de rebote, duración de las sesiones, contenidos visitados, conversiones (para ecommerce), etc. Este producto se desarrolló basándose en la compra de Urchin (hasta entonces la mayor compañía de análisis estadístico de páginas web) por parte de Google.

Google anunció el 2 de abril del 2014 el lanzamiento de la última y más completa actualización de esta herramienta: Universal Analytics.

Para la tecnología web, se comienza añadiendo un código JavaScript a cada una de las páginas que se desea analizar, al que se denomina GATC (Google Analytics Tracking Code). Éste carga algunos archivos desde los servidores Google y monitoriza para luego enviar toda esta información al servidor Google y almacenarla en la cuenta de cada usuario.

Para funcionar, el GATC carga un archivo más grande desde el servidor web de Google, y luego asigna a las variables con el número de cuenta del usuario. El archivo más grande (actualmente conocido como ga.js) es típicamente de 18 KB en tamaño y solo se descarga una vez al comienzo de la visita, ya que se almacenará en la caché durante el resto de la sesión. Como todos los sitios web que implementan Google Analytics con el código de ga.js usan el mismo archivo maestro de Google, un visitante que anteriormente hubiera visitado cualquier otro sitio con este código implementado

también tendrá el archivo en el caché de su máquina. El resultado es que el aumento del tiempo de carga de la página al incluir el código es mínimo.

Tiene una interfaz muy completa de informes con gráficos desarrollados en Adobe Flash. Recientemente se actualizó la interfaz, pero aún está disponible la interfaz antigua habilitada.

También es posible enlazar esta herramienta con Google Webmaster Tools para obtener resultados referidos al estado del sitemap, robots.txt y estado de indexación del sitio general.

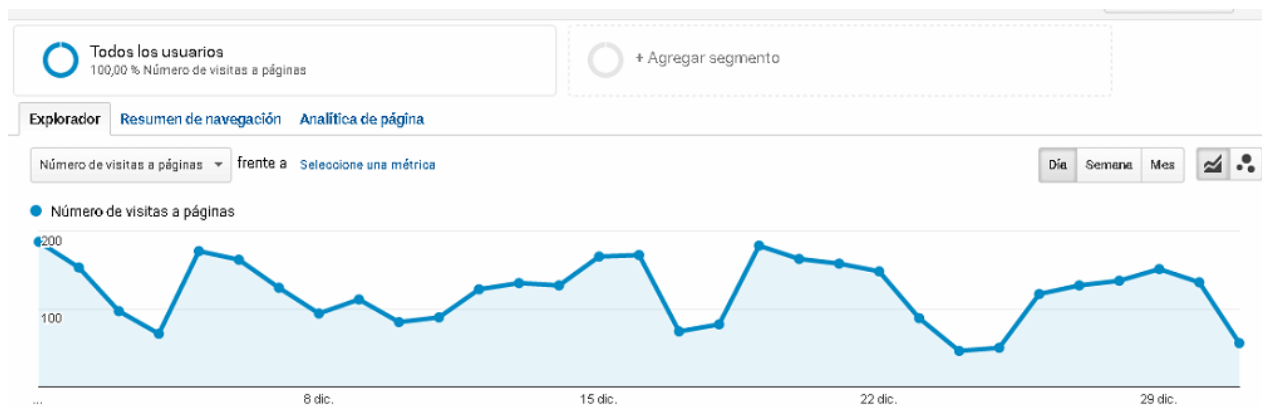


Imagen 22 Google Analytics

### IV.3 CAPACITACIÓN EN LA HERRAMIENTA

Una vez que se ha tomado la decisión de los recursos que serán utilizados tanto de hardware como de software, se comienza con la capacitación del producto.

Debido al tiempo que se debe considerar en la capacitación de la herramienta JasperServer para los administradores y usuarios avanzados, se comienza de manera temprana, casi en paralelo con la definición del proyecto.

Para realizar los ejemplos y trabajos, se utilizaron datos reales desde las bases de datos de Magento e-commerce (transaccionales), que tienen una estructura EAV, pero que se encuentran en una base de datos MySQL.

#### IV.4 DISEÑO DE LA ARQUITECTURA TÉCNICA

Para poblar el DataWarehouse, es decir los datamart de Ventas y de Visitas es necesario realizar integración con modelo EAV de Magento y Google Analytics, para esto se deben considerar algunos factores como, por ejemplo: las directrices técnicas.

La arquitectura propuesta se encuentra representada en la siguiente figura:

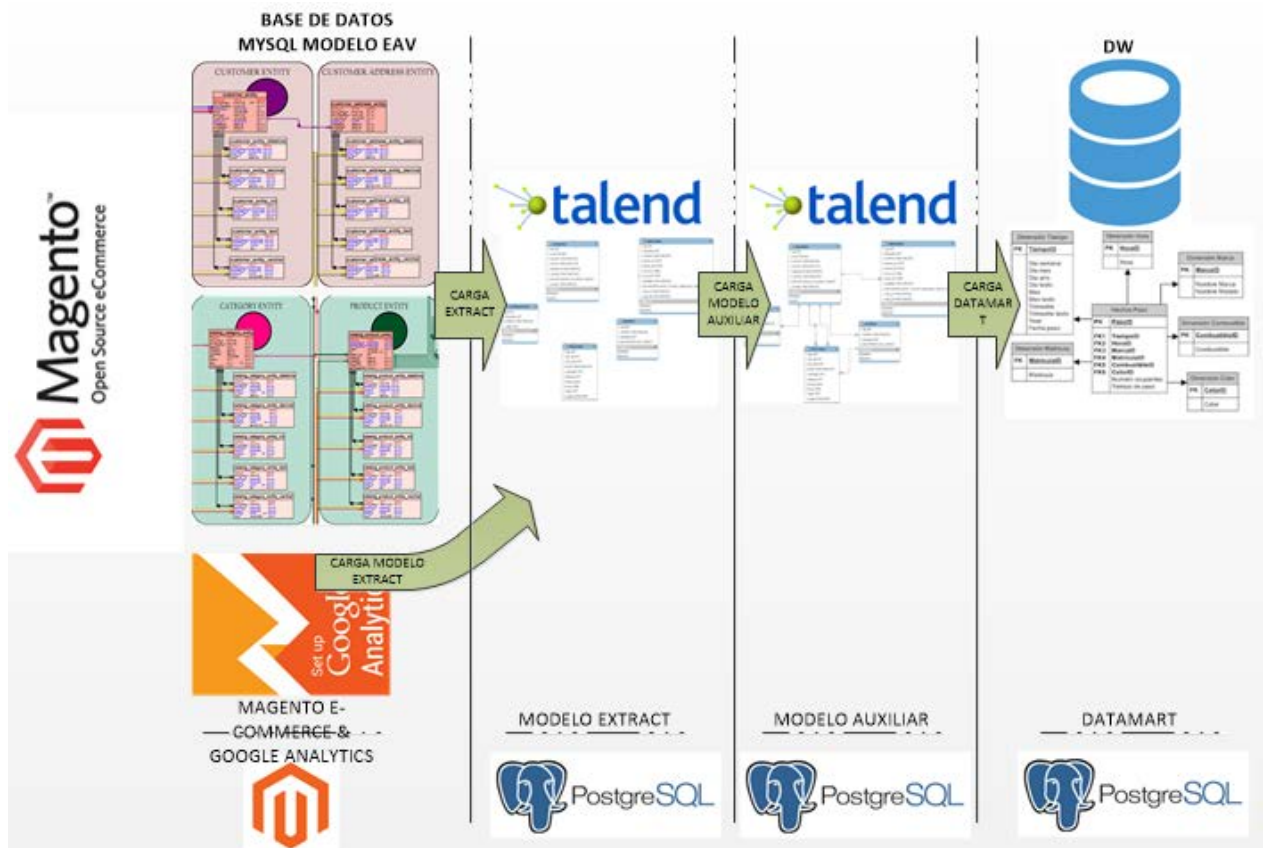


Imagen 23 Arquitectura propuesta para la solución

#### **IV.5 MODELAMIENTO DIMENSIONAL**

Para el modelado dimensional, se decidió en realizar un modelo de Esquema Estrella, sin embargo, se debe considerar que, con este modelo, la cantidad requerida de almacenamiento crece exponencialmente, en nuestro caso utilizaremos técnicas de particionamiento en las tablas fact para que los datamart crezcan linealmente y no exponencialmente. El DataWarehouse se concibe como una composición de DataMarts unidos por dimensiones conformadas por lo que no introduce una nueva plataforma, pero requiere de distintos procesos de población de Datos para distintas estrellas, lo que conlleva a dificultad en mantener concordantes los datos cuando estos aparecen en DataMarts que cumplen distintas funciones.

La decisión de usar un modelado de estrella, fue considerando que un e-commerce pequeño o medio “no genera grandes cantidades de datos diaria”, es decir son de tamaño mediano, además sus sistemas operacionales fueron implementados hace poco más de una década. Esto hace que los volúmenes de datos que albergue un DataWarehouse construido en este ambiente, sean medianos en escala comparativa a grandes empresas como lo puede ser un retail o una compañía telefónica. Lo anterior también es responsable de una disminución en la consideración de la duración del proyecto y de las etapas de cada uno. Por lo tanto, es adecuado aprovechar las ventajas del modelado dimensional y es está la opción que se aplica en el desarrollo del proyecto.

En síntesis, en lo correspondiente al tipo de modelado del DataWarehouse la solución apropiada es modelar los datos utilizando el esquema estrella siempre y cuando las cantidades de datos a manipular no impliquen un volumen de datos tan grande que vuelva más complicada la solución que el problema.

Una vez elegido como método el modelado dimensional se decidió construir el DataMart inicial, basándose en el contexto del negocio y la necesidad actual de la organización.

Ya elegido el DataMart, se comenzó con una abstracción del modelo para visualizar las dimensiones y los hechos y la estructura tipo estrella, hasta llegar a construir las definiciones formales del modelo lógico y físico que tendrá de manera definitiva el DataWarehouse.

El siguiente paso era transformar el modelo lógico de Base de datos, al diseño físico, el diagrama lógico además muestra las características de los campos en dependencia del tipo de dato que debe albergar (tipos de datos y longitud, por ejemplo).

En la ejecución de este nuevo procedimiento, se utilizó la herramienta Enterprise Architect, debido a las bondades de la aplicación de permitir pasar de una definición de alta abstracción como un Diagrama de Entidades, hasta el script de definición para la generación de la Base de Datos.

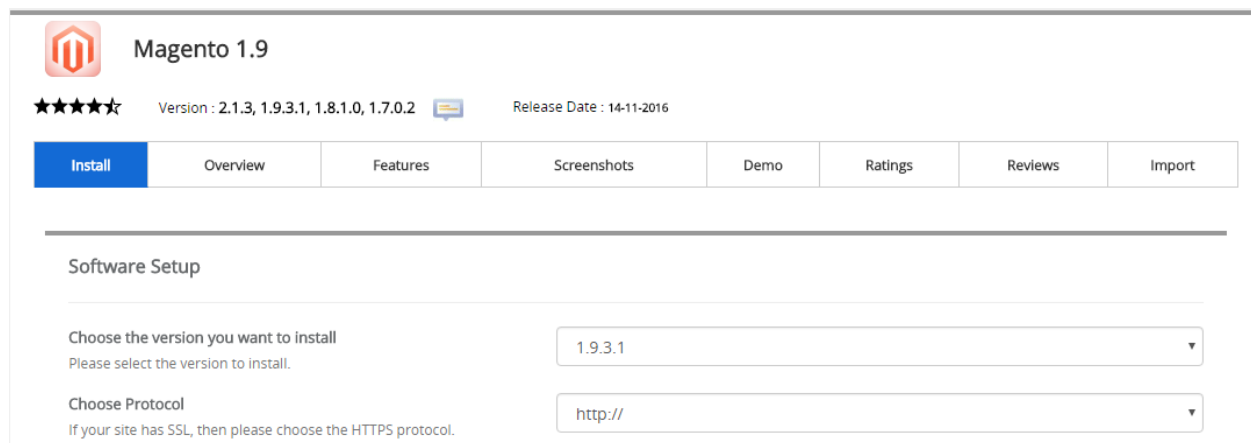


## IV.6 INSTALACION Y CONFIGURACION DE PRODUCTOS

Una vez contratado nuestros servicios de web hosting, se procede con la instalación de los diferentes productos necesarios para implementar el proyecto de e-commerce e Inteligencia de Negocios para pymes. El procedimiento de instalación para las distintas soluciones es el siguiente:

- Verificar prerequisites para instalación.
- Instalación del Software.
- Validación y pruebas de Hardware sobre el Software.
- Aceptación del producto instalado ok.

## IV.7 MAGENTO 1.9



The screenshot shows the Magento 1.9 installation interface. At the top, there is a navigation bar with the following tabs: **Install**, Overview, Features, Screenshots, Demo, Ratings, Reviews, and Import. Below the navigation bar, the 'Software Setup' section is visible. It contains two dropdown menus: 'Choose the version you want to install' (with '1.9.3.1' selected) and 'Choose Protocol' (with 'http://' selected). The text below the second dropdown reads: 'If your site has SSL, then please choose the HTTPS protocol.'

**Imagen 25 Parte del proceso de instalación de Magento**



The screenshot displays the Magento Admin Panel interface. At the top, there is a navigation bar with tabs for Dashboard, Sales, Catalog, Customers, Promotions, CMS, Reports, and System. A search bar and user information (Logged in as admin, Sunday, January 22, 2017) are also present. Below the navigation bar, a message banner indicates a latest message about a new Magento DevBox Beta Speeds Magento 2 Setup. The main content area is titled 'General' and contains several configuration sections: Countries Options, States Options, Locale Options, and Store Information. The Store Information section includes fields for Store Name, Store Contact Telephone, Store Hours of Operation, Country, VAT Number, and Store Contact Address. A 'Validate VAT Number' button is located below the VAT Number field. A 'Save Config' button is located in the top right corner of the configuration area. On the left side, there is a 'Configuration' menu with options for GENERAL, CATALOG, and INVENTORY.

Imagen 26 Parte del proceso de configuración de Magento 1.9

## IV.8 JASPERSOFT

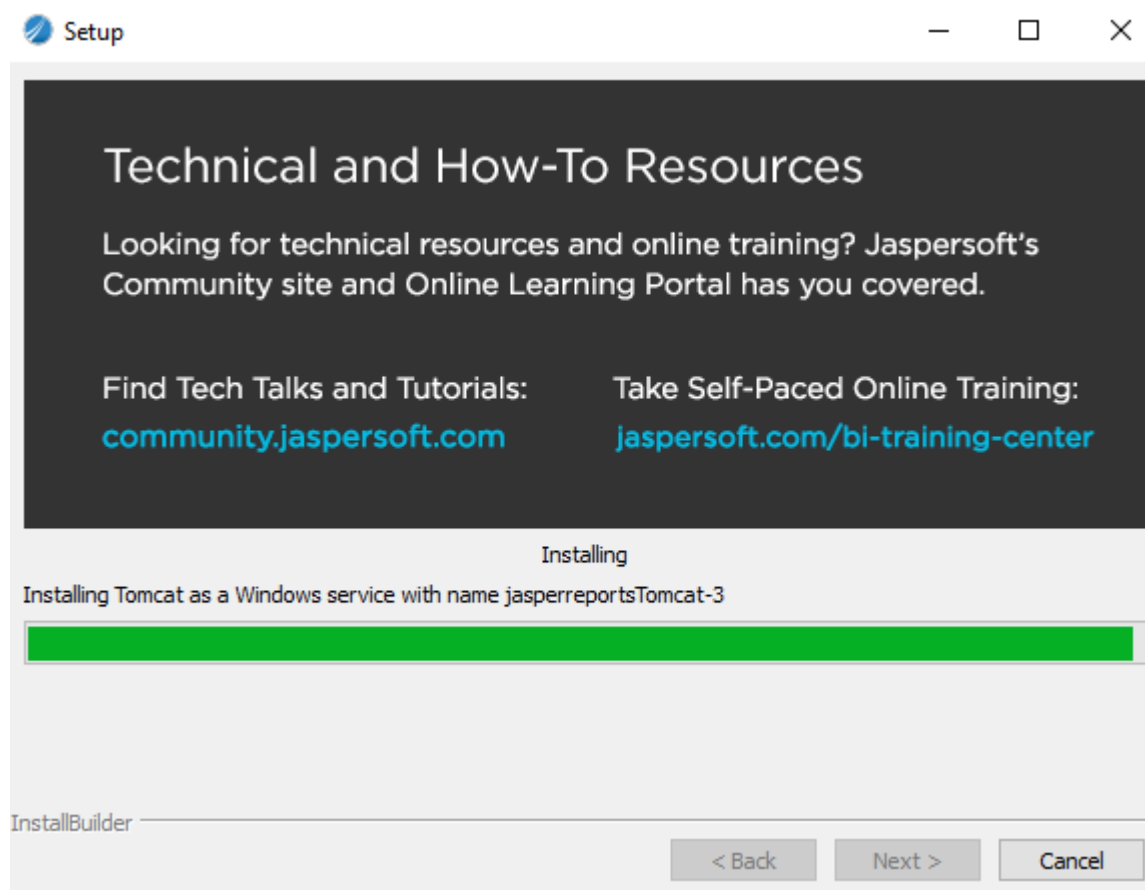
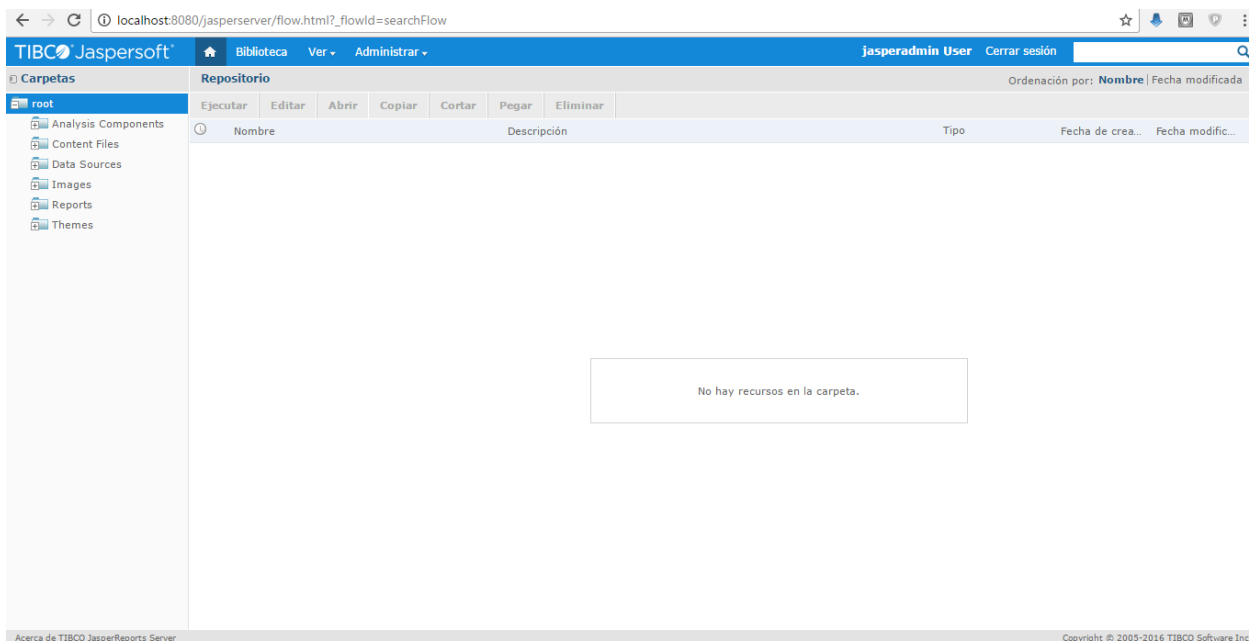


Imagen 27 Parte del proceso de instalación de Jaspersoft



**Imagen 28 Parte del proceso de configuración Jasperserver**

## ***IV.9 DISEÑO FÍSICO***

El diseño físico se focalizó en base a la selección del diseño lógico. En esta etapa se definieron los estándares en los prefijos de los nombres de cada tabla, también conversiones, seteos, además de determinar puntos importantes como la indexación y estrategias de particionamiento.

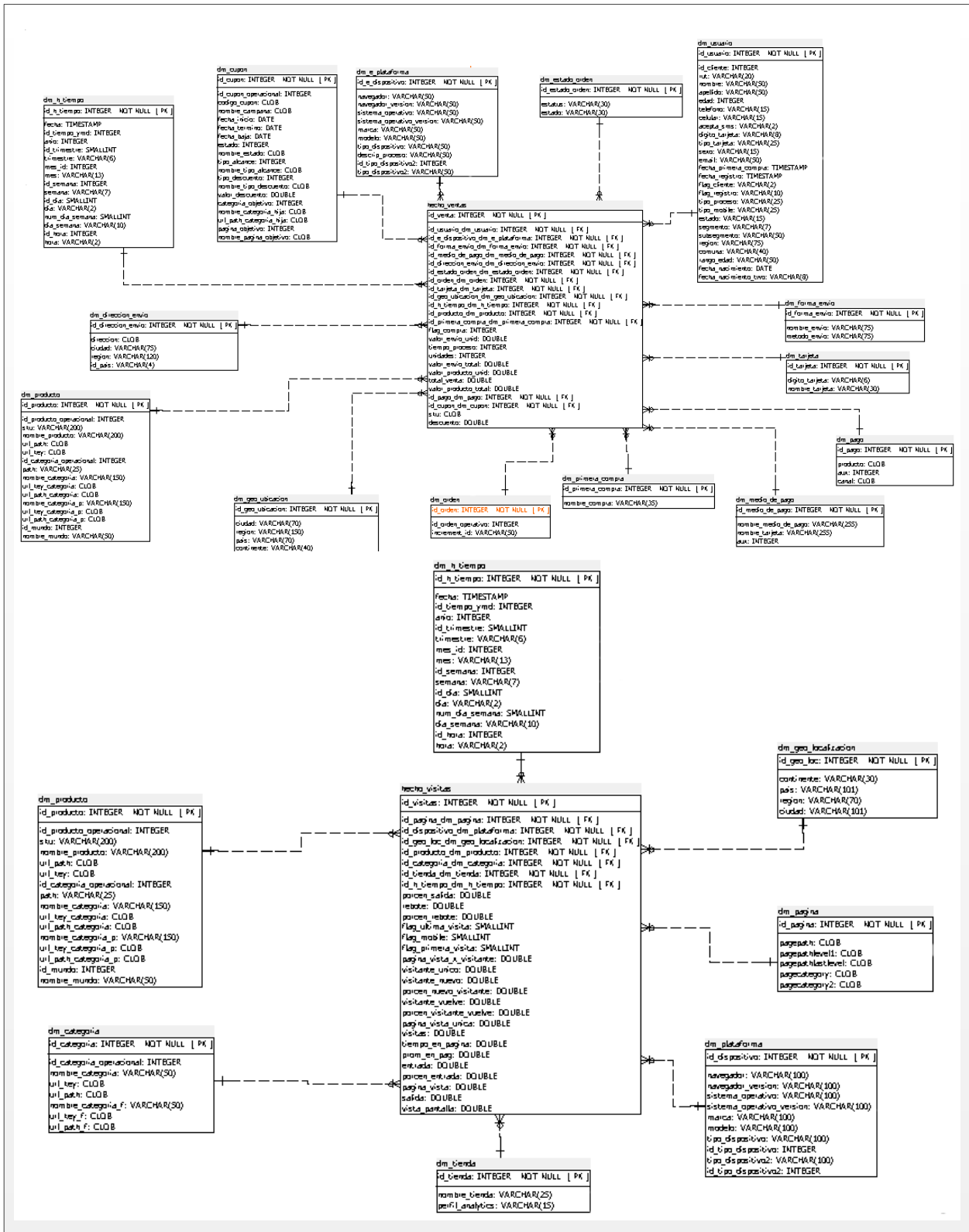


Imagen 299 Modelo físico multidimensional – Datamart visita y Datamart ventas

#### ***IV.10 EXTRACCIÓN, TRANSFORMACIÓN Y CARGA (ETL)***

Este es un proceso muchas veces se menosprecia, pero es bastante importante, ya que es el proceso es delicado y que lleva mucho tiempo. La extracción, transformación y carga en este caso se realizará con Talend Open Studio for Data Integration 6.1.

Postgres es nuestro motor de bases de datos para nuestro Datamart, pero las extracciones se hacen desde el modelo EAV de Magento 9.1 y este se encuentra en MySQL. La solución ETL consta de 2 modelos de paso, el primer modelo es el modelo “extract” este modelo no tiene relación, es decir: No es un modelo entidad relación, sino que es el modelo dónde se guarda la data extraída desde el modelo EAV de Magento “El modelo extract no hace ningún tipo de transformación a la data”. Una vez finalizada la carga del modelo extract, se carga el modelo auxiliar, este modelo es un modelo entidad relación con la data convertida y con las reglas de negocios implementadas, es decir: la data está preparada, limpia y lista para ser insertada en los Datamart, en nuestro caso, principalmente se utilizó en las siguientes funciones:

- Control de la extracción de los datos y su automatización.
- Manejo de excepciones
- Validación en la integridad y consistencia
- Control de redundancias
- Codificar valores (ej. Hombre=1, Mujer=2, etc.)
- Generación de campos clave

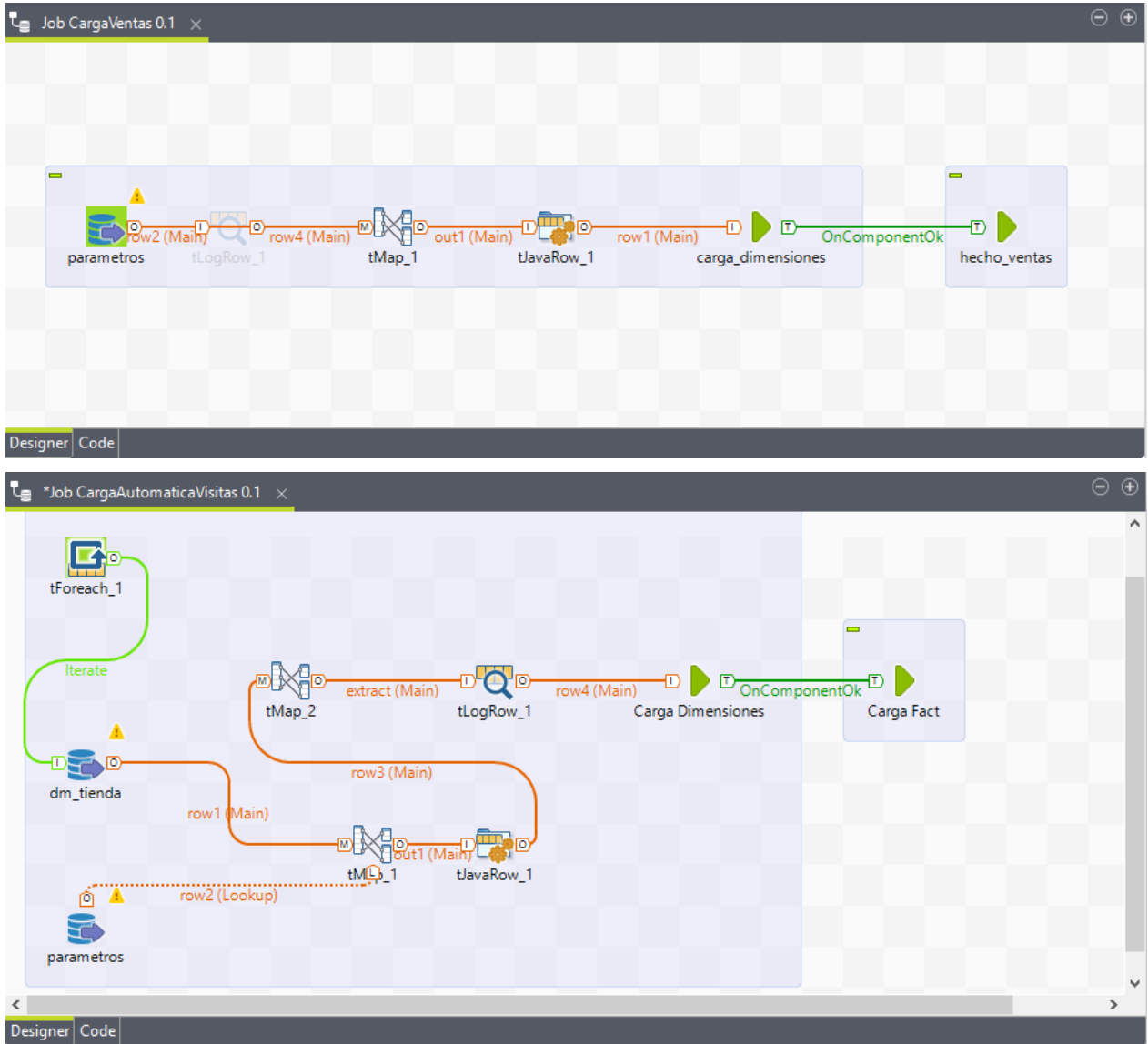


Imagen 30 Proceso ETL Talend Open Studio for Data Integration

## IV.11 DESARROLLANDO LAS APLICACIONES

Los productos a desarrollar: Según la carta gantt se comienza por el portal e-commerce en Magento community 1.9, luego al portal e-commerce en Magento se le implementara el Google Analytics “esto se logra agregando el código UA a Magento” Una vez implementado Magento con Google Analytics se procede con la Implementación de la solución Business Intelligence. Específicamente cargando los datamart de ventas y visitas con Talend. Finalmente se generan los reportes al portal en Jaspersoft “específicamente IReport”

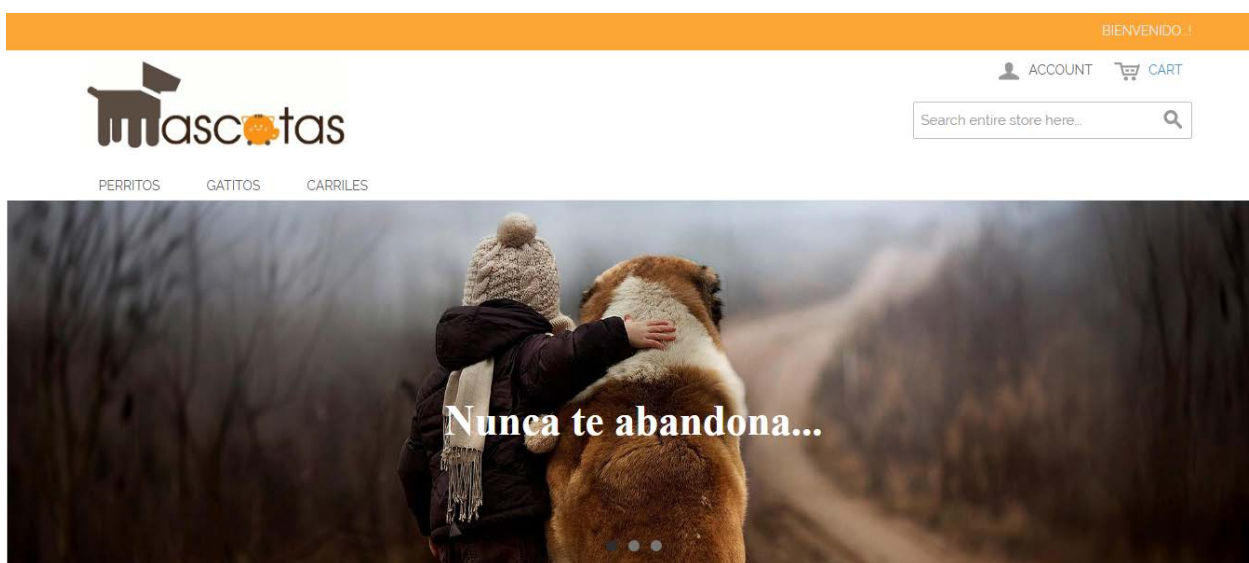


Imagen 31 Portal e-commerce en Magento

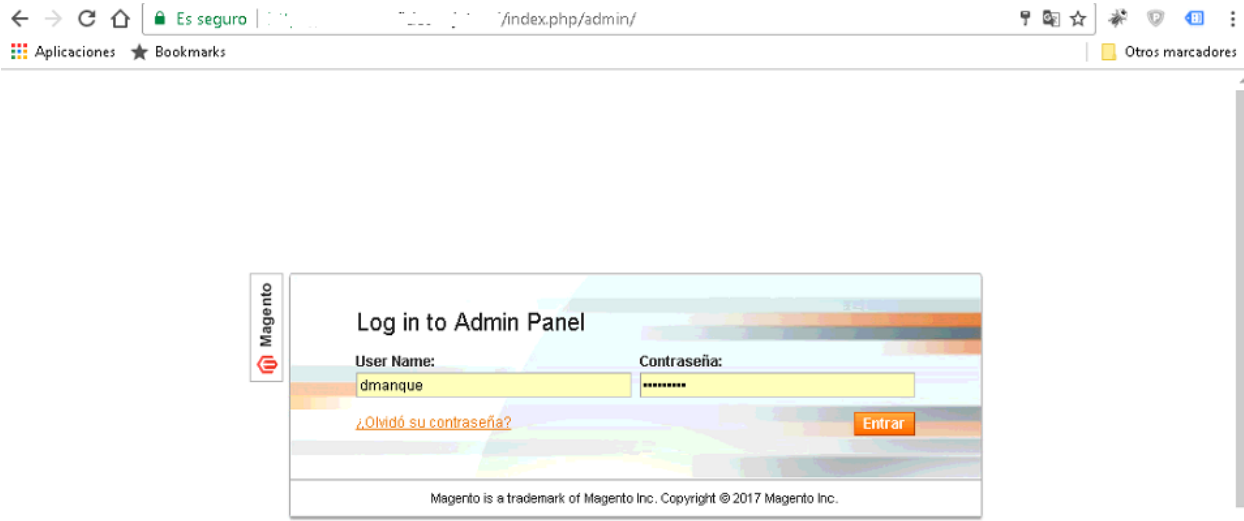


Imagen 32 Portal Magento login módulo Admin

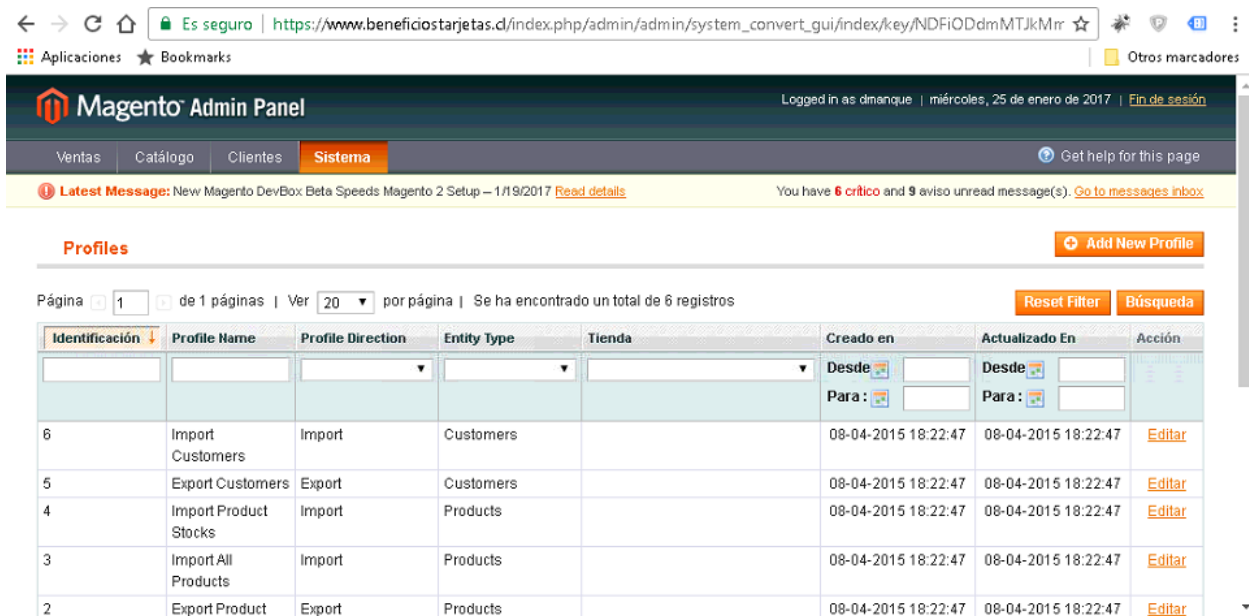


Imagen 33 Portal Magento, módulo de administración



```

</script>
<!-- BEGIN GOOGLE UNIVERSAL ANALYTICS CODE -->
<script type="text/javascript">
  <![CDATA[
    (function(i,s,o,g,r,a,m){i['GoogleAnalyticsObject']=r;i[r]=i[r]||function(){
      (i[r].q=i[r].q||[]).push(arguments)},i[r].l=1*new Date();a=s.createElement(o),
      m=s.getElementsByTagName(o)[0];a.async=1;a.src=g;m.parentNode.insertBefore(a,m)
    })(window,document,'script','//www.google-analytics.com/analytics.js','ga');

    ga('create', 'UA-37890-1', 'auto');
    ga('send', 'pageview');

  ]]>
</script>
<!-- END GOOGLE UNIVERSAL ANALYTICS CODE -->

```

Código UA

Imagen 34 Implementación de Google Analytics en Portal Magento 1.9

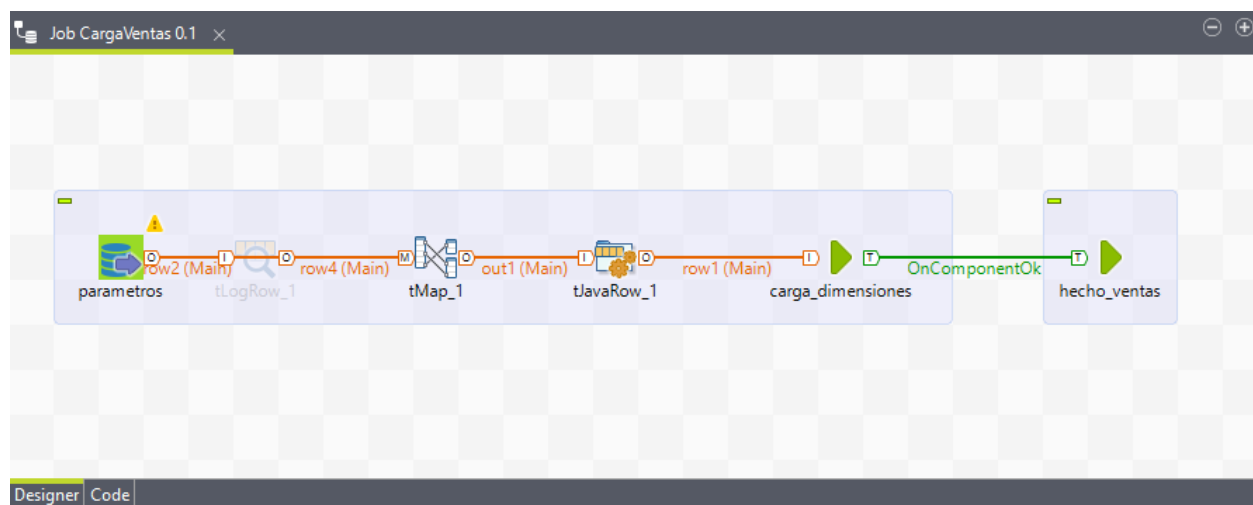


Imagen 35 ETL Datamart de Ventas

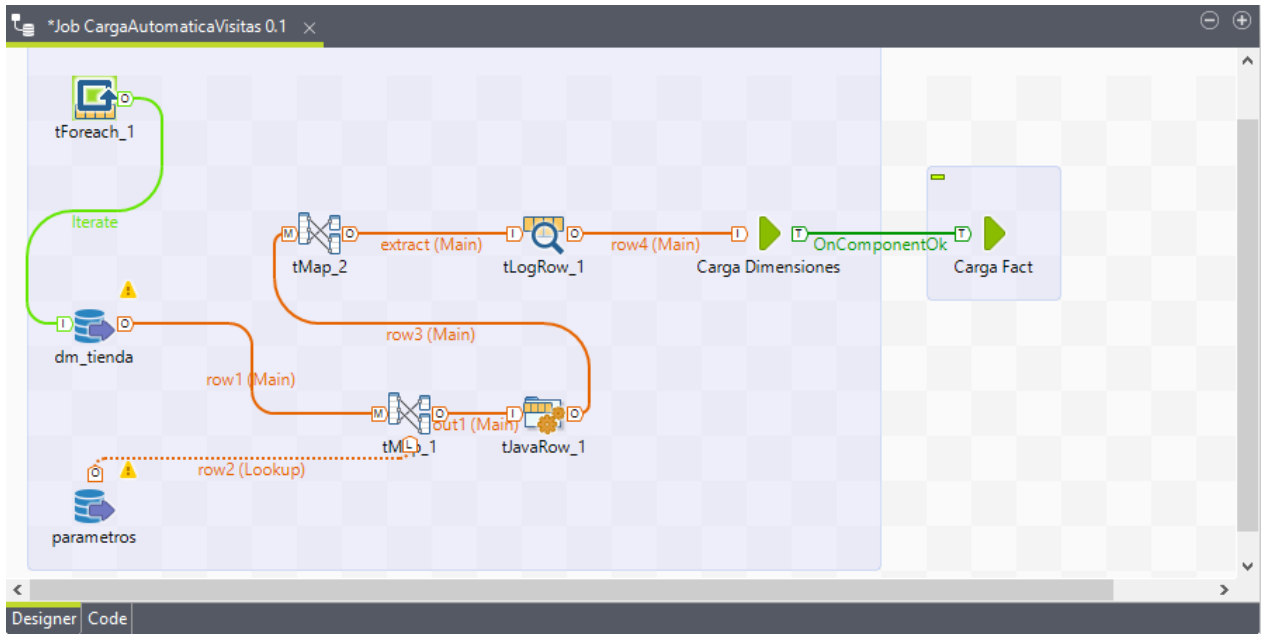


Imagen 36 ETL Datamart de Visitas

```

dmanque@bestadobi:~
00 01 * * * /opt/jasperreports-server-cp-5.5.0/Cron/Portal/Magento/ecommerce.sh

```

Imagen 37 Calendarizando el CRON

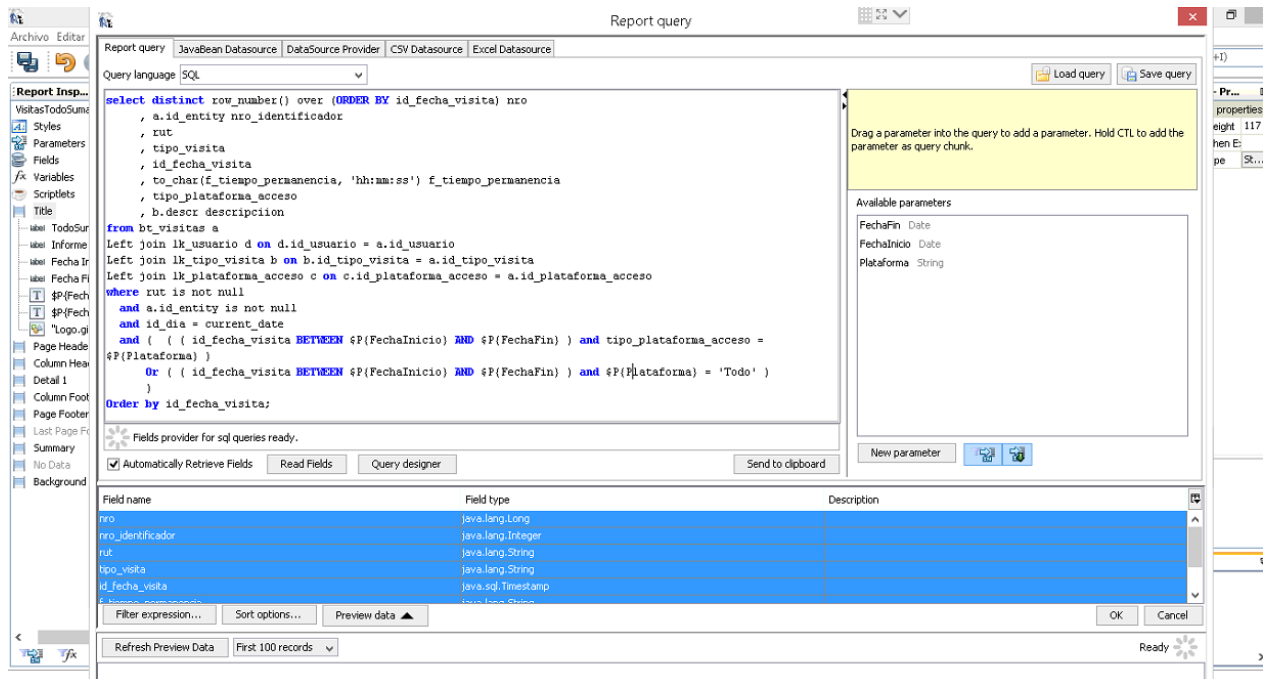


Imagen 38 Desarrollando en IReport – Datamart Visitas

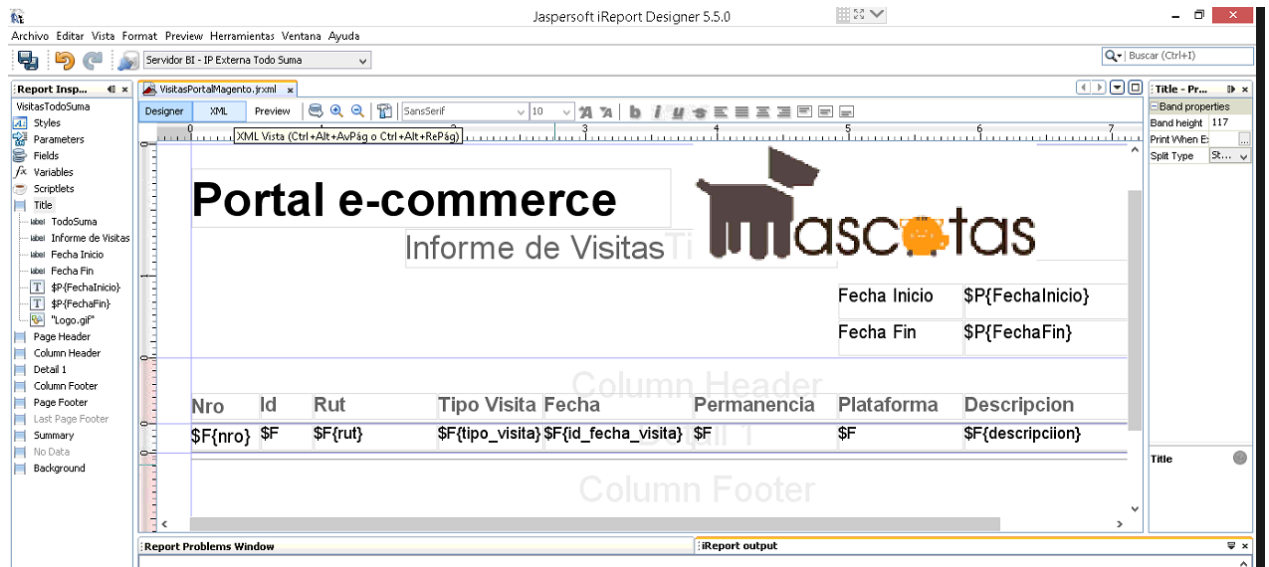


Imagen 39 Desarrollando en IReport - Designer

El desarrollo de estos reportes debe tener la capacidad de hacer filtros por fechas y plataforma de acceso “mobile o desktop”. Estos filtros son implementados vía parámetros y pasados a la Query del reporte “se puede observar en la Imagen 35”

#### ***IV.12 IMPLEMENTACION Y GO-LIVE***

Antes de la liberación al usuario final, se deben analizar varios factores técnicos para su correcto acceso y adecuado funcionamiento, además ha sido considerada desde fases anteriores como en la planificación del proyecto, la correcta difusión y acertada comunicación. Durante el ciclo de vida del proyecto, se deben realizar capacitaciones en el contexto de uso y explotación de la herramienta por cada tipo de usuario.

El implementar este proyecto, significa dejar establecida una solución e-commerce e Inteligencia de Negocio de clase empresarial de un muy buen nivel, ya que lo realizado deja abierta la posibilidad de ir agregando nuevas categorías y productos en la tienda On Line, además de ser flexible para múltiples tipos de monedas, multilinguaje para expandirse al mundo, etc. Es por esta razón que implementar Magento nos parece la solución acertada. Por otro lado, la solución Business Intelligence implementado en Talend y Jaspersoft puede ir escalando a medida que vamos encontrando nuevos patrones e indicadores que nos puedan ir potenciando cualquier tipo de negocio, todo en un ambiente Open Source community, es decir: “libre de pago o licencias”

Los ejemplos presentados solo son los primeros implementados de muchos otros que ya se encuentran en planificación o incluso en desarrollo.

En esta etapa del proyecto se presentará la forma de como es el ingreso a la administración de Magento y el ingreso a herramienta de Inteligencia de Negocios, mostrando ejemplos concretos de administración de productos en el e-commerce ,reportes y funcionalidades generales, los cuales fueron creados en este proyecto de

tesis, dejando en evidencia la riqueza que posee en información. Estos ejemplos ya cuentan con los datos transformados a través del proceso ETL y almacenados en la Base de Postgres.

En la siguiente figuras se muestran ejemplos del ingreso vía web al portal de la herramienta JasperServer, una vez validadas las credenciales aparece la pantalla inicial donde se visualizan los reportes a los cuales se pueden ingresar.



Imagen 40 Pagina Login de Usuarios Jasperserver

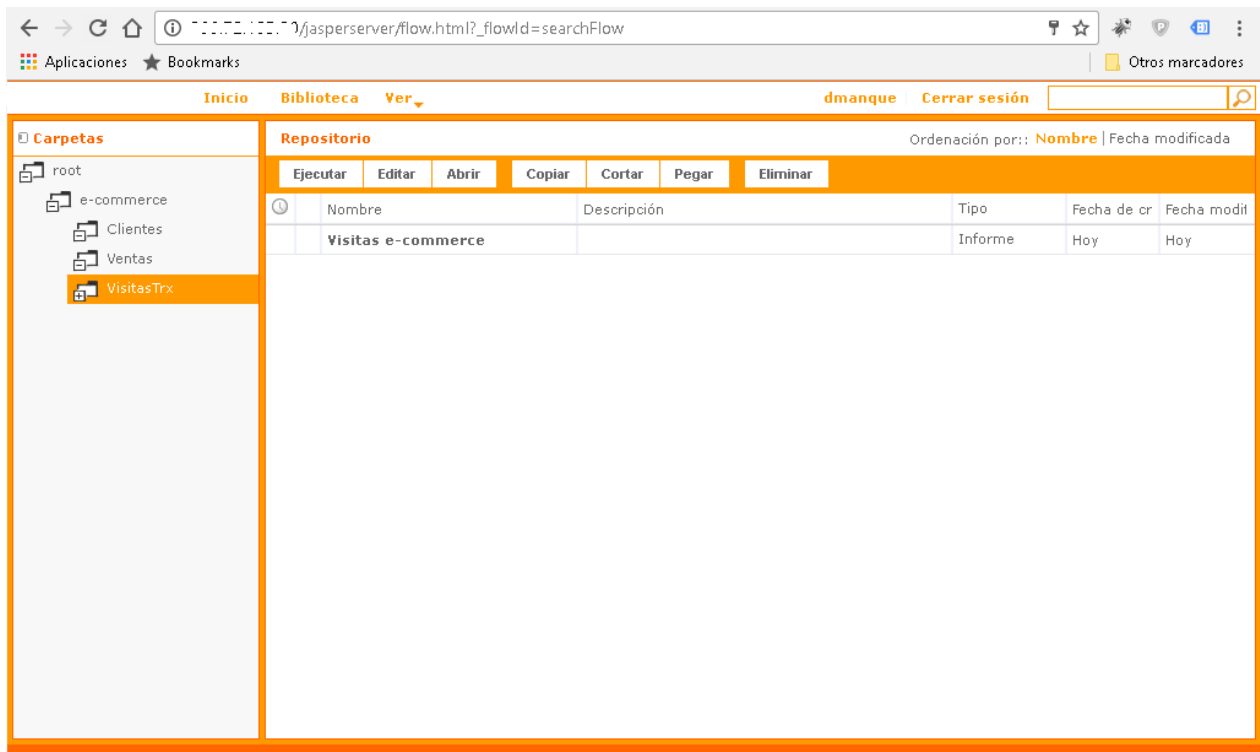
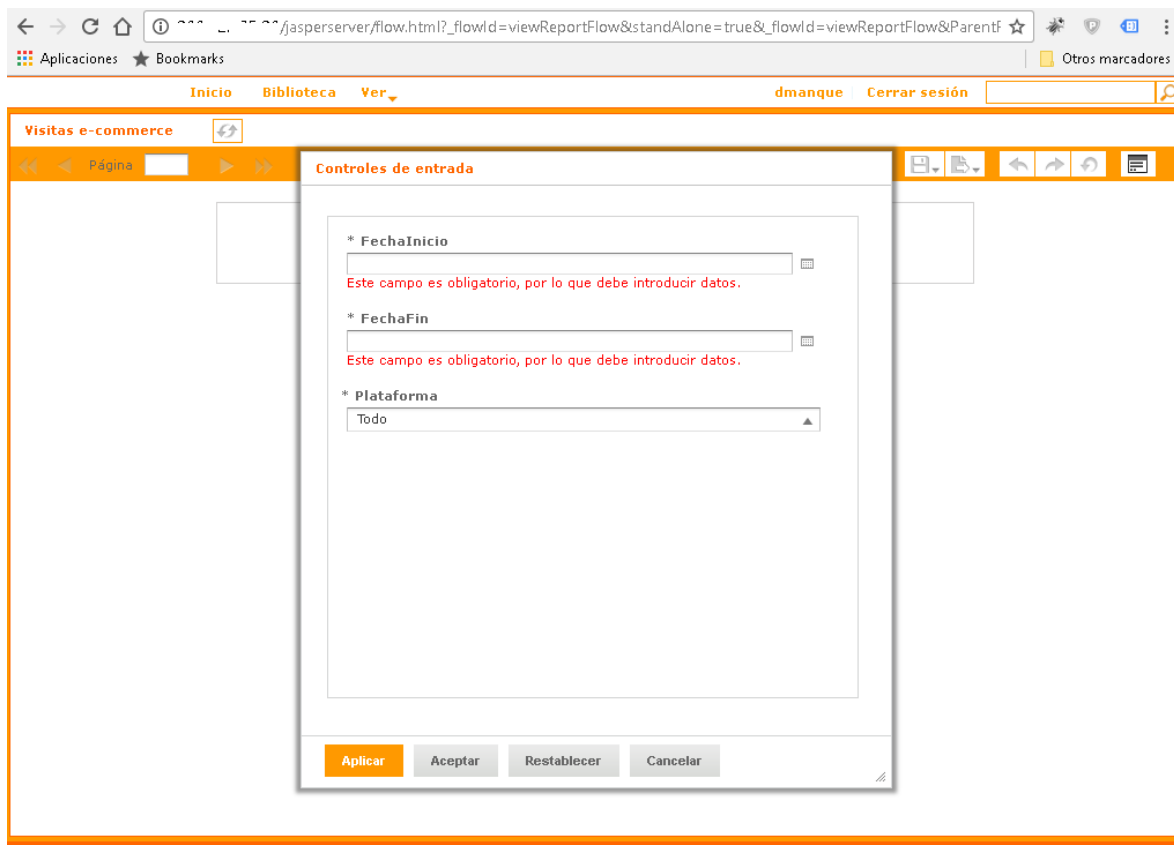


Imagen 41 Portal de Proyectos de JasperServer

- 1) Parámetros de entrada tipos de Reporte tipos de visitas en JasperServer.
- 2) El reporte se puede exportar para múltiples formatos.



The image shows a screenshot of a web browser displaying a report configuration interface. The browser's address bar shows a URL starting with 'jasperserver/flow.html?\_flowId=viewReportFlow&standAlone=true&\_flowId=viewReportFlow&Parentf'. The page title is 'Visitas e-commerce'. A modal dialog titled 'Controles de entrada' is open, containing three input fields:

- \* FechaInicio**: A text input field with a calendar icon. Below it, a red error message reads: 'Este campo es obligatorio, por lo que debe introducir datos.'
- \* FechaFin**: A text input field with a calendar icon. Below it, a red error message reads: 'Este campo es obligatorio, por lo que debe introducir datos.'
- \* Plataforma**: A dropdown menu with 'Todo' selected.

At the bottom of the dialog, there are four buttons: 'Aplicar' (highlighted in orange), 'Aceptar', 'Restablecer', and 'Cancelar'.

Imagen 42 Parámetros Reporte de Tipo de Visitas

The screenshot shows a web browser window displaying a report from 'Portal e-commerce'. The report title is 'Informe de Visitas' and it includes the 'mascotas' logo. The data is for the period from 1/01/17 0:00 to 24/01/17 0:00. The report is presented as a table with 8 columns: Nro, Id, Rut, Tipo, Fecha, Permanencia, Plataforma, and Descripción. A download menu is open on the right side of the page, listing various export options such as PDF, Excel, CSV, DOCX, RTF, ODT, ODS, and XLSX (paginado).

Nro	Id	Rut	Tipo	Fecha	Permanencia	Plataforma	Descripción
1	16364	6796567-1	IS	1/01/17 0:27	11:00:11	WEB	Inicio de Sesión
2	16364	6796567-1	CS	1/01/17 0:28	11:00:11	WEB	Cierre de Sesión
3	17412	10303603-1	ES	1/01/17 0:46	11:00:11	WEB	Enrolamiento en el Sitio
4	17412	10303603-1	IS	1/01/17 0:46	11:00:11	WEB	Inicio de Sesión
5	17412	10303603-1	CS	1/01/17 0:47	11:00:11	WEB	Cierre de Sesión
6	17413	9050397-9	ES	1/01/17 1:32	11:00:11	WEB	Enrolamiento en el Sitio
7	10539	16861540-K	IS	1/01/17 1:34	11:00:11	MOBILE	Inicio de Sesión
8	10539	16861540-K	CS	1/01/17 1:35	11:00:11	MOBILE	Cierre de Sesión
9	17413	9050397-9	IS	1/01/17 1:39	11:00:11	WEB	Inicio de Sesión
10	15346	13110022-1	RC	1/01/17 1:42	11:00:11	MOBILE	Recuperación de Clave
11	9316	6286959-3	IS	1/01/17 1:46	11:00:11	WEB	Inicio de Sesión
12	9316	6286959-3	CS	1/01/17 1:50	11:00:11	WEB	Cierre de Sesión
13	8192	17582507-K	RC	1/01/17 2:08	11:00:11	MOBILE	Recuperación de Clave
14	17414	10199470-8	ES	1/01/17 2:14	11:00:11	WEB	Enrolamiento en el Sitio
15	17415	8810795-0	ES	1/01/17 2:15	11:00:11	WEB	Enrolamiento en el Sitio
16	17415	8810795-0	IS	1/01/17 2:19	11:00:11	WEB	Inicio de Sesión
17	4801	10082186-9	IS	1/01/17 2:33	11:00:11	WEB	Inicio de Sesión
18	4801	10082186-9	CS	1/01/17 2:33	11:00:11	WEB	Cierre de Sesión
19	14246	13058356-2	IS	1/01/17 2:47	11:00:11	WEB	Inicio de Sesión
20	17416	6893570-9	ES	1/01/17 2:58	11:00:11	MOBILE	Enrolamiento en el Sitio

Imagen 43 Reporte de Tipo de Visitas



## 1) Parámetros de entrada, Reporte tipos de usuarios en JasperServer.

The screenshot shows a web browser window with the URL `jasperserver/flow.html?_flowId=viewReportFlow&standAlone=true&_flowId=viewReportFlow&Parentf`. The browser's address bar and navigation icons are visible. The page content includes a header with 'Inicio', 'Biblioteca', and 'Ver' menus, and a user profile 'dmanque' with a 'Cerrar sesión' link. The main content area is titled 'Clientes e-commerce' and features a 'Controles de entrada' dialog box. This dialog box is used for configuring report parameters and includes the following fields:

- \* FechaInicio**: A date input field containing '01-01-2017'. Below it, a red error message reads: 'Este campo es obligatorio, por lo que debe introducir datos.'
- \* FechaFin**: A date input field containing '24-01-2017'. Below it, a red error message reads: 'Este campo es obligatorio, por lo que debe introducir datos.'
- \* Plataforma**: A dropdown menu with 'Todo' selected. The dropdown is open, showing a search bar and a list of options: 'Todo', 'MOBILE', and 'WEB'.

At the bottom of the dialog, there are four buttons: 'Aplicar' (highlighted in orange), 'Aceptar', 'Restablecer', and 'Cancelar'.

Imagen 44 Parámetros Reporte de Clientes

## 2) Reporte de Usuarios en JasperServer.

The screenshot shows a web browser displaying a report titled "Portal e-commerce Datos Personales". The report includes a table with 18 rows of user data. A context menu is open over the table, offering various export options like PDF, Excel, CSV, etc. The browser's address bar shows a URL with parameters for report flow and parent frame.

**Portal e-commerce** Datos actualizados 24-ene-2017 a las 21:52:39

**Portal e-commerce** Datos Personales

Fecha Inicio: 1/01/17 0:00 Fecha Fin: 24/01/17 0:00

Nro	Rut	Nombre	Apellido	email	Fecha Primer	Plataforma
1	11000000-0	Nancy	Bustamante	nancy.bustamante@gmail.com	1/01/17 0:46	WEB
2	11000000-0	carlos	pineda	carlos.pineda@gmail.com	1/01/17 1:39	WEB
3	11000000-0	monica	sobarzo	monica.sobarzo@hotmail.com	1/01/17 2:19	WEB
4	11000000-0	Mima	Jena	mima.jena@gmail.com	1/01/17 2:59	MOBILE
5	11000000-0	Rosa	Canto	rosa.canto@gmail.com	1/01/17 4:25	MOBILE
6	11000000-0	Viviana	Valdivia	viviana.valdivia@hotmail.com	1/01/17 8:46	MOBILE
7	11000000-0	Edison	Sanmartin	edison.sanmartin@hotmail.com	1/01/17 8:59	MOBILE
8	11000000-0	Ivan	Homazabal	ivan.homazabal@live.cl	1/01/17 9:51	MOBILE
9	11000000-0	Ema	Caneo	ema.caneo@hotmail.com	1/01/17 9:54	MOBILE
10	11000000-0	Maria	Gallardo	maria.gallardo@gmail.com	1/01/17 9:58	WEB
11	11000000-0	isabia	neira	isabia.neira@gmail.com	1/01/17 10:00	WEB
12	11000000-0	alonso	Florian	alonso.florian@hotmail.com	1/01/17 10:01	MOBILE
13	11000000-0	Valeria	Jara	valeria.jara@hotmail.com	1/01/17 10:12	WEB
14	11000000-0	Digna	Aldana	digna.aldana@live.cl	1/01/17 10:34	MOBILE
15	11000000-0	Juan	Peralta	juan.peralta@gmail.com	1/01/17 10:36	MOBILE
16	11000000-0	Richard	Vivanco	richard.vivanco@yahoo.es	1/01/17 10:36	MOBILE
17	11000000-0	Paula	Toledo	paula.toledo@gmail.com	1/01/17 11:11	MOBILE
18	11000000-0	Miquel	Santander	miquel.santander@gmail.com	1/01/17 11:20	MOBILE

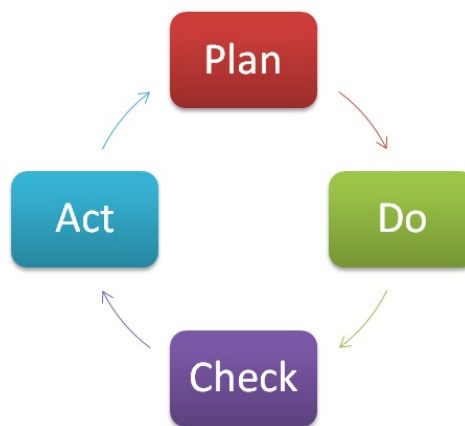
Export options: Como PDF, Como Excel (paginado), Como Excel, Como CSV, Como DOCX, Como RTF, Como ODT, Como ODS, Como XLSX (paginado), **Como XLSX**

Clientes\_e\_comme...xlsx | Mostrar todo

Imagen 45 Reporte de Usuarios

### **IV.13 MANTENCION Y CRECIMIENTO**

Una solución e-commerce y de Inteligencia de negocios se encuentra en constante cambio, ya sea por nuevos productos, nuevos idiomas o por modificaciones, dentro de este marco la Mejora Continua se enfoca en generar una evolución a través de cambios que permitan mejorar los procesos y lo que está en operación, los cuales pueden ser a través de pequeños cambios que no necesiten grandes inversiones y permitan mejorar algún indicador de rendimiento. Esta mejora basada principalmente en el ciclo de vida de Deming compuesto por cuatro fases (PDCA en inglés) o (PHVA en español): Planificar, Hacer, Verificar y Actuar.



**Imagen 46 El círculo de Deming PDCA**

Frente a este escenario, en las soluciones de e-commerce e Inteligencia de negocios, debemos estar preparados para realizar una eficiente mantención y buen manejo en el crecimiento, enfocado siempre en contribuir, desarrollar y expandir a las pymes.

## **V. RESULTADOS**

Este proyecto de tesis tiene como consecuencia distintos tipos de resultados, los que se pueden clasificar como tangibles e intangibles.

### **V.1 RESULTADOS INTANGIBLES**

Como resultados intangibles, se puede rescatar el conocimiento adquirido de tecnologías de punta, como son las tiendas e-commerce y el Business Intelligence en ambientes Open Source, dejando esta tecnología al alcance de cualquier pyme, enfocados siempre en una solución de carácter social y transversal para que cualquier pyme se integre a nuestro BI generando soluciones estándares de acuerdo a los tiempos contemporáneos. Enfocados como una inversión para el negocio de a corto, mediano y largo plazo.

### **V.2 RESULTADOS TANGIBLES**

Los resultados tangibles más importantes en este proyecto son varios:

- Se implementa un nuevo canal de ventas On Line en Magento responsivo para mobiles, tables y desktop.
- Se tiene un modelo o diseño estándar capaz de integrar otras pymes a la solución e-commerce y Business Intelligence, permitiendo así su explotación sin gran esfuerzo, de esta forma las pymes pueden basar sus decisiones comerciales y operativas en la misma información.

- Se logró la implementación del diseño lógico de los Datamart y posteriormente la creación al esquema físico en tablas.
- Se desarrollaron los procesos automáticos a través de Talend Data Integrator y procedimientos almacenados, para la extracción, transformación, normalización y carga de los datos.
- Se lograron terminar informes para el apoyo en la toma de decisiones.

### **V.3 VENTAJAS Y DESVENTAJAS**

De acuerdo al desarrollo del proyecto de tesis, lecciones aprendidas y al trabajo realizado, se rescatan ventajas y desventajas al implementar e invertir en un e-commerce con Business Intelligence Open Source community.

#### **V.3.1 DESVENTAJAS**

- Requiere un fuerte enfoque a la gestión del cambio, donde siempre existe un porcentaje de personas resistentes a la innovación.
- Requiere de una inversión de capacitación a los usuarios finales.
- Incremento continuo en los requerimientos, ya que al ser cambiante las necesidades del usuario y del negocio irán variando durante el tiempo.

#### **V.3.2 VENTAJAS**

- No hay que hacer colas para comprar

- Acceso a tiendas y productos en lugares remotos
- No es necesario tener una tienda física para vender
- Es posible ofrecer y encontrar una gran cantidad de opciones
- Las tiendas online están disponibles todos los días a todas las horas
- Facilidad para crecer y ofrecer más y mejores productos y servicios
- No hay limitaciones ni condicionamientos de espacio, lo que permite tener más productos disponibles
- Personalización de la compra y de la experiencia de cliente
- No hay necesidad de manejar dinero en efectivo
- Transacciones y contrataciones rápidas y eficientes
- Reducción de gastos de personal
- Posibilidad de comprar y vender productos más raros o menos comerciales, pero que tienen su cuota de mercado
- La Inteligencia de Negocios tiene la ventaja de reunir, normalizar y centralizar toda la información de la empresa, mediante un almacén de datos.
- Mejorar la toma de decisiones, realizándola de forma más rápida, informada y basada en hechos
- Eliminar ineficiencias y reducir los costes operativos al generar información de una única fuente fidedigna.
- Acortar los tiempos de respuesta a las peticiones de informes.
- Proveer el autoservicio de información a trabajadores, colaboradores, clientes y proveedores.
- Medición sobre el progreso de la empresa, ya que reúne la información tanto del negocio como la operacional, permitiendo a los ejecutivos rastrear con precisión y eficiencia los KPIs y comparar pasado con presente.
- Planificar el futuro, teniendo la información pasada y presente de la organización a través del modelo de Inteligencia de Negocio, es tener la habilidad para planificar el futuro, reconocer oportunidades de negocio potenciales y adelantarse a la curva.

- Liberar las bases de datos transaccionales a consultas del negocio, permitiendo mayor eficiencia en su desempeño.
- Elevar la capacidad de análisis en el contexto estratégico, analítico y operacional.
- Permite tener la información de manera oportuna y de calidad.
- Con la información recolectada centralizadamente es posible identificar tendencias y anticipar riesgos.

## **VI. CONCLUSIÓN**

En este proyecto de título, se demuestran los elementos esenciales para el desarrollo de una solución e-commerce e Inteligencia de Negocios. También es importante indicar que no existe una metodología universal para tener éxito frente a un proyecto e-commerce o BI, ya que independiente a la metodología, esta debe ser adaptada al contexto de la organización, ya sea en recursos humanos, tiempos, necesidades, etc.

La generación de un proyecto e-commerce y BI introduce un nuevo entorno en el cual es posible realizar ventas On Line, obtener información vital para la empresa, de manera rápida y de manera autónoma, lo que habilita a la rápida toma de decisiones que mejoren el desempeño de la compañía y a los créditos que podría obtener.

La principal enseñanza que se establece con este trabajo es la enorme gama de posibilidades que ofrece implementar un e-commerce y BI, sus herramientas, aquí se mostraron algunos casos, en los cuales se puede aplicar la Inteligencia de Negocios, con diferentes formas de trabajar, soportadas por sistemas de información particulares a cada una de ellas y con distintos contenidos en sus bases de datos.

Adicionalmente gracias a este proyecto se ha tenido la posibilidad de tomar contacto con herramientas muy utilizadas en el ámbito de e-commerce y Business Intelligence.

Este proyecto es la base para futuras implementaciones de Magento e-commerce para otras pymes que se deseen incorporarse al DataWarehouse estándar para pymes, ampliando su enfoque de acuerdo a las necesidades de la pyme.

Se ha podido apreciar que una solución de e-commerce e Inteligencia de Negocios no reemplaza a las tecnologías de información convencionales, sino que convive con ellas e integra un marco conceptual interdisciplinario en donde conviven disciplinas administrativas con las ingenierías del software y el conocimiento del negocio.



## VII. GLOSARIO

### **Almacén de datos o DataWarehouse**

Son repositorios diseñados para facilitar la confección de informes y realización de análisis; son separados del sistema de información principal; lo cual significa una ganancia enorme en el rendimiento de los sistemas cuando se ejecuten las consultas.

### **Atributo**

Corresponde a un elemento de información, que describe una característica de una entidad.

### **Base de datos**

Conjunto de datos no redundantes, almacenados en un soporte informático, organizado de forma independiente de su utilización y accesible simultáneamente por usuarios y aplicaciones.

### **BI**

Forma abreviada para referirse a Business Intelligence.

### **Business Intelligence**

Inteligencia de Negocios en inglés.

### **Cuadrante Mágico de Gartner**

Es una representación gráfica del mercado durante un periodo específico de tiempo. Muestra el análisis desarrollado por Gartner Inc., sobre cómo se comportan ciertos actores del mercado en función de diferentes criterios de evaluación. El informe de Gartner no apoya a ninguno de los productos o proveedores mencionados ni tampoco aconseja el uso o la selección de los proveedores colocados en el cuadrante de líderes.

**DataMart**

Subconjunto de datos de un DataWarehouse.

**Datos**

Información concreta sobre hechos, elementos, etc., que permite estudiarlos, analizarlos o conocerlos.

**Dimensión**

Entidad independiente dentro del modelo multidimensional de una organización, son las variables sobre las que se hace la consulta

**Diseño físico**

Es el proceso de producir la descripción de la implementación y establece el detalle de los componentes y configuraciones de la base de datos

ETL (Extract, Transform and Load)

Extraer, transformar y cargar, es el proceso que permite a las organizaciones mover datos desde múltiples fuentes, reformatearlos y limpiarlos, y cargarlos en otra base de datos, Datamart, o DataWarehouse

**Diseño lógico**

Es el proceso de construir un esquema de la información que utiliza la empresa, basándose en un modelo de base de datos específico, independiente del SGBD concreto que se vaya a utilizar y de cualquier otra consideración física.

**E-commerce**

Es un método de compraventa de bienes, productos o servicios valiéndose de internet como medio, es decir, comerciar de manera online.

**Esquema tipo estrella**

Un Esquema completamente des normalizado, donde todos los niveles están en una misma tabla para toda la dimensión

**Granularidad**

Término que se usa en los DataWarehouse para expresar el nivel de detalle. A más alto nivel de granularidad más bajo nivel de detalle (mayor nivel de abstracción).

**Hecho o Fact (en inglés)**

Es un dato numérico, que puede ser agregable para responder preguntas de negocio en un modelo multidimensional.

**IBM (International Business Machines)**

Empresa multinacional estadounidense de tecnología y consultoría.

**Integridad referencial**

Es una propiedad deseable en las bases de datos, se garantiza que una entidad (fila o registro) siempre se relacione con otras entidades válidas, es decir, que existen en la base de datos.

**Inteligencia de Negocios**

Es básicamente la habilidad para transformar los datos en información, y la información en conocimiento, de forma que se pueda optimizar el proceso de toma de decisiones en los negocios.

**Jaspersoft**

Proveedor líder mundial de plataformas de software empresarial.

**Jerarquía**

Se trata de la gradación de personas, animales u objetos según criterios de clase, tipología, categoría u otro tópico que permita desarrollar un sistema de clasificación

**Microsoft** (Microsoft Corporation)

Empresa multinacional desarrolladoras de software de origen estadounidense

**Modelo Auxiliar**

Modelo de paso donde el dato es extraído y convertido desde el modelo extract, para finalmente cargarlo al modelo auxiliar. Este modelo es cíclico y almacena hasta 30 días.

**Modelo Extract**

Modelo de paso donde el dato es extraído desde el modelo EAV MySQL y cargado a Postgres. El modelo extract no tiene ninguna referencia, es decir: No tiene relaciones.

**Motor de Base de Datos**

Servicio principal para almacenar, procesar y proteger los datos.

**MOLAP**

Un sistema MOLAP usa una base de datos propietaria multidimensional, en la que la información se almacena multidimensionalmente, para ser visualizada multidimensionalmente.

**OLAP**

Es el acrónimo en inglés de procesamiento analítico en línea (On-Line Analytical Processing). Es una solución utilizada en el campo de la llamada Inteligencia de negocios (o Business Intelligence) cuyo objetivo es agilizar la consulta de grandes cantidades de datos.

**OLTP**

Procesamiento de Transacciones En Línea o en inglés OnLine Transaction Processing, es una clase de programa que facilita y gestiona aplicaciones orientadas a transacciones

### **Open Source**

Open Source o Código Abierto es un término que se aplica al Software distribuido bajo una licencia que le permita al usuario acceso al código fuente del Software

### **Periodo**

En el contexto de Pets un periodo corresponde al mes y año en que se realiza el pago de las cotizaciones, ejemplo 201701 corresponde al año 2017, mes Enero.

### **Pets**

Es una pyme que vende inmobiliario y accesorios para mascotas.

### **POTGRES** (Oracle Corporation).

Motor Open Source de Sistemas de Gestión de Bases de Datos.

### **Pyme**

Pequeña y medianas empresas.

### **Qlikview**

Es un software fabricado por la empresa QlikTech, para el análisis visual que aporta conocimiento, datos y claridad en la toma de decisiones

### **ROLAP**

Procesamiento Analítico OnLine Relacional, es decir, se trata de sistemas y herramientas OLAP (Procesamiento Analítico OnLine) construidos sobre una base de datos relacional. Es una alternativa a la tecnología MOLAP

**SAP** (System, Applications and Products)

Sistemas, Aplicaciones y Productos

**SQL**

Structured Query Language (SQL), en español, lenguaje de consulta estructurado, para bases de datos relacionales que permiten realizar diversos tipos de operaciones.

**Stakeholders**

Termino (en inglés “los interesados”), utilizado para referirse a quienes son afectados o pueden ser afectados en los cambios realizados por la empresa.

**Talend Open Studio for Data Integration**

Talend, permite generar soluciones de integración de datos de alto rendimiento, entre las que se incluyen paquetes de extracción, transformación y carga de datos (ETL) para el almacenamiento de datos

## VIII. **BIBLIOGRAFÍA**

### **Internet**

- <https://wiboomeia.com/mega-tutorial-magento-que-es-y-como-crear-tu-ecommerce-premium/> (Magento, 2015)
- <https://debitoor.es/glosario/definicion-ecommerce> (debitoor, s.f.)
- <http://community.jaspersoft.com/wiki/designing-report-jaspersoft-studio> (Jaspersoft, s.f.)
- <http://churriwifi.wordpress.com/category/business-intelligence/> (churriwifi, s.f.)
- <http://www.gestiopolis.com/recursos/documentos/fulldocs/ger/busint.htm> (gestiopolis, s.f.)
- <https://media.readthedocs.org/pdf/pgsqtutorial/9.1.0/pgsqtutorial.pdf> (Acevedo, 2016)
- <https://community.talend.com/t5/custom/page/page-id/Tutorials> (community, s.f.)