

INTRODUCCION

En este documento se presenta el desarrollo de una solución Inteligencia de Negocio para la empresa Triton Trading S.A. el cual apoya a la toma de decisiones mediante el análisis de información integrando las técnicas Inteligencia de Negocios.

En este trabajo se expone una solución Inteligencia de Negocios factible para otras empresas se describe de la manera más clara posible las técnicas de Inteligencia de Negocio y cómo implementarlas, esto con el fin de comprender qué son y qué se realizó en este trabajo. Además se mencionan algunas herramientas de software disponibles para desarrollar cada una de las técnicas de este tipo de proyectos, esto con motivo de dar a conocer alternativas de software diferentes a las que se abordaron en este proyecto.

Este proyecto se realizó con base en los requerimientos de estudio: Triton Trading S.A., una empresa de tipo comercial con necesidades de análisis de información para mejorar su administración. Se pretende que este documento sirva como guía referencial para poder aplicar este tipo de soluciones en las distintas empresas del grupo.

Finalmente se muestra en el documento los resultados obtenidos con la implementación de la solución y las conclusiones / recomendaciones respectivas.

CAPITULO I

PLANTEAMIENTO DEL PROBLEMA

1.1. Descripción de la Realidad Problemática

Des hace algún tiempo la empresa Triton Trading S.A ha comprendido la importancia que puede tener la información unificada tanto de las ventas como de las compras de maquinarias y repuestos a los distintos clientes y proveedores, no solo es aconsejable sino que se ha vuelto imprescindible para un efectiva toma de decisión con respecto a las ventas y compras.

Él seguir vendiendo a un cliente con mejores precios, es una forma de fortalecer su posición ante competidores y lograr no solo clientes satisfechos sino clientes leales a la empresa esto también está íntimamente relacionado con una efectiva compra .

Esto no se ve reflejado en la empresa Triton Trading S.A, esto se debe a los problemas que enfrenta en la actualidad la empresa que se caracteriza, por una estructura pesada y lenta heredada de prácticas pasadas que impiden poderse mover de forma adecuada y con la creciente complejidad del entorno ha hecho que la información de ventas de maquinarias y repuestos a los distintos clientes y la compra de nuevos artículos y equipos que sustituyan a estos no está unificada si no que se encuentra diversificada en el sistema que

La empresa maneja. Haciendo compleja la toma de decisión de ventas y el abastecimiento de repuestos para el estoqueo del almacén comprometiendo el soporte de la atención de los pedidos diarios que realizan los distintos clientes y nuevos clientes que se les ha vendido o no equipos. Esto también conlleva a tomar decisiones no tan claras que pueden comprometer los nuevos proyectos o los proyectos de venta de la empresa ante las nuevas oportunidades que se abren en la actualidad. Como se sabe las ventas son los ingresos que tiene la empresa este es el sostén que tiene la empresa para seguir realizando su operación y crecer organizacionalmente todo este crecimiento está ligado a una buena gestión y sobre todo a la toma de decisión oportuna tanto en ventas como en compras con los repuestos, equipos con la cantidad adecuada, en la actualidad ya sea por una mala gestión o por el temor de quedarse sin stock en los almacén se hicieron compras de artículos y equipos para el estoqueo, por decirlo la empresa se sentía segura para no dejar de atender a sus clientes, esto por una parte está bien pero por el otro se hizo compra de productos que tiene poca rotación en el mercado esto genera gastos de almacenaje que perjudican el buen desempeño económico de la empresa, en este caso la toma de decisión con respecto a las compras no fue clara y tampoco tenían clara su entorno de ventas, es por ello la importancia de tener la información adecuada en el tiempo adecuado para la toma de decisión.

1.2. JUSTIFICACION DEL PROBLEMA

El desarrollo de una solución Inteligencia de Negocios en la empresa Triton Trading S.A. contribuye con la toma adecuada y asertiva de las decisiones correspondientes a la información actual de las compra y ventas de maquinarias y repuestos haciendo competentes con precios adecuados, fidelizando a los clientes a la empresa, también contribuye a la toma de oportunidades más adecuadas en el momento y tiempo adecuados sacando

ventaja a sus competidores, invirtiendo las cantidades y tiempo adecuadas para una obtención de utilidades provechosa. Toma el pulso en tiempo real de la empresa y el crecimiento que está experimentando. Todo esto ayuda al crecimiento de la organización con mejores ventas y compras que en la actualidad está experimentando la empresa Triton Trading S.A en la compleja coyuntura sacando el máximo provecho a las circunstancias.

1.3. DELIMITACIONES DE LA INVESTIGACIÓN

1.3.1. Limitaciones

A. Limitaciones espaciales

La aplicación de este desarrollo de Inteligencia de Negocios en la ventas de maquinaria y repuestos se realizara en la empresa Triton Trading S.A. ubicada en el kilómetro 17.5 de la antigua panamericana en el distrito de villa el salvador de parlamento de Lima

B. Delimitaciones sociales

La empresa Triton Trading S.A constituye en una fuente de puestos de trabajo, que emplea alrededor de 200 personas actualmente.

La dirección conformada por:

- Gerencia Comercial. Sr. Renzo Gamboa
- Gerente comercial SPV: Sr. Miguel Olavide
- Gerente técnico: Sr. Fernando Jiménez
- Coordinado de importaciones: Sra. Claudia Serna
- Coordinadora de soporte comercial: Sra. Marlene Zavala .

Tienen como rol contribuir al desarrollo económico de la empresa Triton Trading S.A.

C. Delimitaciones Temporales

El presente proyecto de investigación será elaborado y coordinado durante los 6 meses

1.4. FORMULACIÓN DEL PROBLEMA

1.4.1. Problema Principal

¿De qué manera el desarrollo de una solución Inteligencia de Negocios contribuirá en toma de decisiones en el área de compra y ventas de la empresa Triton Trading S.A

1.5 OBJETIVO

Desarrollo de una solución Inteligencia de Negocios que contribuya en toma de decisiones en el área de compra y ventas de la empresa Triton Trading S.A.

CAPITULO II

MARCO TÉORICO

2.1. ANTECEDENTES

La tesis del Sr. Manuel Javier Valdiviezo Basauri en sus tesis “Análisis y Diseño de una herramienta de desarrollo de soluciones para inteligencia de negocios” menciona que hoy en día las empresas necesitan depositar mucha confianza en la toma de decisiones sobre sus negocios, basándose en la información que poseen sobre sus operaciones y lo que ocurre en el mercado. Estas decisiones, que son tomadas en un contexto de alta y creciente competencia, se toman cada vez más utilizando la mayor cantidad de información que la empresa posee. Una mala decisión puede afectar significativamente a la organización. (Lima 2011)

También podemos entrar en la tesis de la señorita Keller Gladys Rodríguez Cabanillas “Análisis Diseño E Implementación De Una Solución De Inteligencia De Negocios Para El Área De Compras Y Ventas De Una Empresa Comercializadora De Electrodomésticos” explica que es muy impórtate la ayuda de la toma decisiones por parte de la metodología y herramientas de inteligencia de negocios

ya que asegura un información en tiempo real a las distintas gerencias o gerencia para la tomar decisiones adecuadas y sobretodo oportunas con llevando así una ventaja competitiva en el mercado

2.2. BASES TEÓRICA

2.2.1. ADMINISTRACIÓN DEL CONOCIMIENTO

La administración del conocimiento es una disciplina que articula personas, procesos, contenido y tecnología, se origina y reside en el cerebro de las personas, por lo que la generación, transferencia y aplicación del conocimiento debe ser fomentada y recompensada, dado que la administración del conocimiento es más un reto cultural y organizacional que un asunto de tecnología.

En ese sentido, es el proceso mediante el cual el aprendizaje individual y la experiencia pueden ser representados, compartidos y utilizados para fomentar el mejoramiento del conocimiento individual y el valor organizacional. Por otro lado, en ciencias de la información se define al conocimiento como al conjunto de datos e información destinados a resolver un determinado problema. Ver figura 1.



Figura 1. Esquema sobre el conocimiento desde el punto de vista de las ciencias de la información,

Fuente: Obtenido del Proyecto de Aplicación de BI. Universidad Rey Juan Carlos, España, 2010

En ese sentido se señala que cualquier empresa se enfrenta a diferentes retos a la hora de almacenar los datos que dificultan alcanzar una inteligencia de negocios a nivel de empresa. La información normalmente se encuentra en lo que se podrían denominar silos de información, que pueden ser utilizados por distintas unidades de negocio e incluso se pueden encontrar en plataformas dispares.

2.2.2. DATOS

Los datos son la mínima unidad semántica, y se corresponden con elementos primarios de información que por sí solos son irrelevantes como apoyo a la toma de decisiones. También se pueden ver como un conjunto discreto de valores, que no dicen nada sobre el porqué de las cosas y no son orientativos para la acción.

Un número telefónico o un nombre de una persona, por ejemplo, son datos que, sin un propósito, una utilidad o un contexto no sirven como base para apoyar la toma de una decisión. Los datos pueden ser una colección de hechos almacenados en algún lugar físico como un papel, un dispositivo electrónico, o la mente de una persona. En este sentido las tecnologías de la información han aportado mucho a recopilación de datos.

Como cabe suponer, los datos pueden provenir de fuentes externas o internas a la organización, pudiendo ser de carácter objetivo o subjetivo, o de tipo cualitativo o cuantitativo, etc.

2.2.3. INFORMACIÓN

La información se puede definir como un conjunto de datos procesados y que tienen un significado (relevancia, propósito y contexto), y que por lo tanto son de utilidad para quién debe tomar decisiones, al disminuir su incertidumbre. Los datos se pueden transformar en información añadiéndoles valor:

- Contextualizando: se sabe en qué contexto y para qué propósito se generaron.
- Categorizando: se conocen las unidades de medida que ayudan a interpretarlos.
- Calculando: los datos pueden haber sido procesados matemática o estadísticamente.
- Corrigiendo: se han eliminado errores e inconsistencias de los datos.
- Condensando: los datos se han podido resumir de forma más concisa (agregación).

Por tanto, la información es la comunicación de conocimientos o inteligencia, y es capaz de cambiar la forma en que el receptor percibe algo, impactando sobre sus juicios de valor y sus comportamientos.

Información = Datos + Contexto (añadir valor) + Utilidad (disminuir la incertidumbre)

2.2.4. CONOCIMIENTO

El conocimiento es una mezcla de experiencia, valores, información, que sirve como marco para la incorporación de nuevas experiencias e información, y es útil para la acción. Se origina y aplica en la mente de los conocedores. En las organizaciones con frecuencia no sólo se encuentra dentro de documentos o almacenes de datos, sino que también está en rutinas organizativas, procesos, prácticas, y normas.

El conocimiento se deriva de la información, así como la información se deriva de los datos. Para que la información se convierta en conocimiento es necesario realizar acciones como:

- Comparación con otros elementos.

- Predicción de consecuencias.

2.2.5. CICLO DE LA INTELIGENCIA DE NEGOCIOS

Para Peña (2006) la Inteligencia de Negocios es mucho más que una actitud empresarial o una tecnología a disposición de las organizaciones; es un marco de referencia para la gestión del rendimiento empresarial, un ciclo continuo por el cual las compañías definen sus objetivos, analizan sus progresos, adquieren conocimiento, toman decisiones, miden sus éxitos y comienzan el ciclo nuevamente. En tal sentido, ayuda a los gerentes a tomar mejores decisiones más rápidamente en los niveles estratégico y operativo.

El ciclo se compone de cuatro fases o etapas a saber: análisis, reflexión, acción y medición. Para lo cual el análisis comienza por determinar los datos a recopilar, la selección se basa en un entendimiento básico y en supuestos de cómo opera la organización, considerando aquello que es relevante a los clientes, proveedores, empleados, los factores que afectan los insumos, entre otros.



Figura 2 ciclo de vida de negocio

Fuente: Elaboración Propia

Tal como se aprecia en la figura 2, implicará poner en relación, comparar y valorar la información obtenida en la fase de análisis y reflexión, apoyados en el benchmarking, esta actividad servirá para ayudar a la decisión de desarrollar su propio marco de referencia, con la vista puesta tanto en la estrategia a largo plazo como en la gestión a corto plazo, para una generar respuestas cada vez más rápidas y ciclos óptimos de realimentación, lo cual proporciona mayores oportunidades de realizar las mencionadas pruebas orientadas a la acción.

- **ENTORNO DE TRABAJO**

Peña (2006) explica que las empresas pueden tener muchas razones diferentes para abordar el entorno de trabajo, a través de la planeación y control de sus procesos en apoyo de sus objetivos estratégicos, modernización, innovación y expansión, y además pretenderán hacerlo a la vista de la ampliación hacia nuevos mercados desatendidos o el desarrollo de su gama de productos, para aumentar sus cuotas de mercado.

2.2.6. MODELOS DE LA SOLUCIÓN

Los modelos de solución de la Inteligencia de Negocios procuran aprovechar las bondades de la tecnología de cómputo para resolver un problema de administración de conocimiento.

Al respecto las aplicaciones de la inteligencia de negocios tienden a crear sistemas especializados en una función específica de la empresa, que contribuya a la eficiencia en el diagnóstico de una situación y tomar la decisión adecuada

para la solución mediante la sistematización de los datos, refinamiento de la información y representación del conocimiento (Hernandez¹, 2004).

2.2.7. INTELIGENCIA DE NEGOCIOS

Es una estrategia empresarial que persigue incrementar el rendimiento de la empresa o la competitividad del negocio, a través de la organización inteligente de sus datos históricos (transacciones u operaciones diarias), usualmente residiendo en Data Warehouse corporativos o Data Marts departamentales.

El concepto de BI no es nuevo, desde que la idea fue introducida a mediados de los años 60, no ha dejado de evolucionar a soluciones más efectivas y adaptadas al nuevo entorno tecnológico imperante. Con el precio del hardware en franco descenso, procesadores más potentes, la hegemonía de Internet-Web y software de gestión más eficientes, el concepto de inteligencia de negocio (BI) se coloca al alcance de muchas organizaciones modernas quienes están interesadas en maximizar sus inversiones en el área informática.

El DSS (Decision Support Systems) fue el origen de todo, luego aparecieron conceptos similares tales como los EIS (Executive Information Systems), hasta llegar al estado del arte actual, los BIs y BI-Web. Los pioneros del campo fueron el Dr. Ralph Timbal, considerado el Dr. del DSS, y Bill Inmon², considerado el padre del Data Warehouse.



Figura 3 desarrollo de BI

Fuente: Introducción a la minería de datos. Madrid: Pearson

¹ Hernandez, J. (2004). *Introducción a la minería de datos*. Madrid: Pearson.

² Inmon, W. (2005). *Building the Datawarehouse*. Usa: Wiley

- **Importancia de BI**

Generalmente, en las organizaciones se genera una gran cantidad de datos e

información que en muchos de los casos el análisis de la misma se convierte en un verdadero problema para los directivos.

Las tecnologías y los sistemas de BI permiten realizar un análisis mucho más ágil y comprensible para la toma de decisiones empresariales, las aplicaciones BI buscan incrementar la eficiencia en la organización. Podemos decir que la información, correctamente analizada e interpretada, es la mayor fuente de poder de las empresas, ya que da pistas muy claras acerca del camino a seguir en futuras acciones.

- **Beneficios de BI**

Entre los beneficios más importantes que brinda una aplicación BI a las organizaciones, se puede mencionar los siguientes:

Minimiza el tiempo de carga de datos, debido a que todos los datos se encuentran en un mismo repositorio o fuente de información.

Los procesos de extracción y carga de la información son automáticos debido

al uso de procesos definidos y metodologías.

- Las herramientas BI permiten realizar análisis, y establecer comparaciones para la toma de decisiones.
- Permite a los usuarios no depender de reportes o informes programados, porque los mismos serán generados de manera dinámica.

- Posibilita la formulación preguntas y respuestas que son claves para el desempeño de la organización.
- Permite acceder y analizar directamente los indicadores de éxito.
- Identifica cuáles son los factores que inciden en el buen o mal funcionamiento de la organización.
- Permite detectar situaciones fuera de lo normal.
- Predice el comportamiento futuro con un alto porcentaje de certeza, basado en el entendimiento del pasado.
- Permite consultar y analizar los datos de manera sencilla e intuitiva.
- Elimina los gastos innecesarios en la organización.
- Aumenta ingresos y producción de bienes.
- Mejora el conocimiento de la empresa sobre clientes, empleados, costos etc. Lo que facilita la gestión de los directivos.

2.2.8. DATAWAREHOUSING

William Inmon³ quien es considerado padre del datawarehouse, lo define como una colección de datos diseñada para apoyar la toma de decisiones. Los datawarehouses integran datos que han sido seleccionados y organizados de manera histórica, sobre los que se realizan análisis que ayuden a justificar las decisiones estratégicas tomadas en las empresas. (Inmon, 2005)

Las características principales de un datawarehouse según (Inmon, 2005) son las siguientes:

- El datawarehouse está orientado a un contexto, organiza y orienta los datos desde la perspectiva del usuario final.

³Inmon, W. (2005). *Building the Datawarehouse*. Usa: Wiley

- Administra grandes cantidades de información. La mayoría de los datawarehouses contienen información histórica que se retira con frecuencia de las bases de datos operacionales.
- Comprende múltiples versiones de esquemas de datos, debido a que el datawarehouse tiene que guardar información histórica, y como esta ha sido manejada en distintos momentos por diferentes versiones de esquemas, debe poder administrar información originada en diferentes bases de datos.
- Condensa y agrega información. Con frecuencia, es muy alto el nivel de detalle de la información guardada. Un datawarehouse condensa y agrega la información para presentarla en forma comprensible a las personas.

Un datawarehouse se compone de los siguientes elementos (Kimball⁴, 2008):

- Fuentes de datos de sistemas operacionales: se refieren a las bases de datos operacionales que contienen información recopilada de las aplicaciones operacionales, que pueden venir en diversos esquemas tales como modelos relacionales, no relacionales o basados en archivos, y pueden ser tanto internas como externas a la organización.
- Área mediación de datos: en esta área se realiza un conjunto de procesos conocidos como de extracción, transformación y carga (ETL por sus siglas en inglés). Esta área comprende todo aquello que se encuentra entre la fuente de datos operacionales y el área de presentación de datos.
- Área de presentación de datos: es el lugar en donde los datos son organizados, almacenados y están disponibles para las consultas, reportes y resúmenes que realicen los usuarios.
- Herramientas de acceso a los datos: todas las herramientas de acceso a los datos realizarán consultas a partir del área de presentación de datos.

⁴ Kimball, R. (2008). *The datawarehouse lifecycle*. Usa: Wiley.

Una herramienta de acceso puede consistir desde una simple consulta, hasta minería de datos.

Para mostrar de manera gráfica los aspectos comunes mencionados anteriormente, se propone la figura 1, que representa en sí la arquitectura general de un datawarehouse.

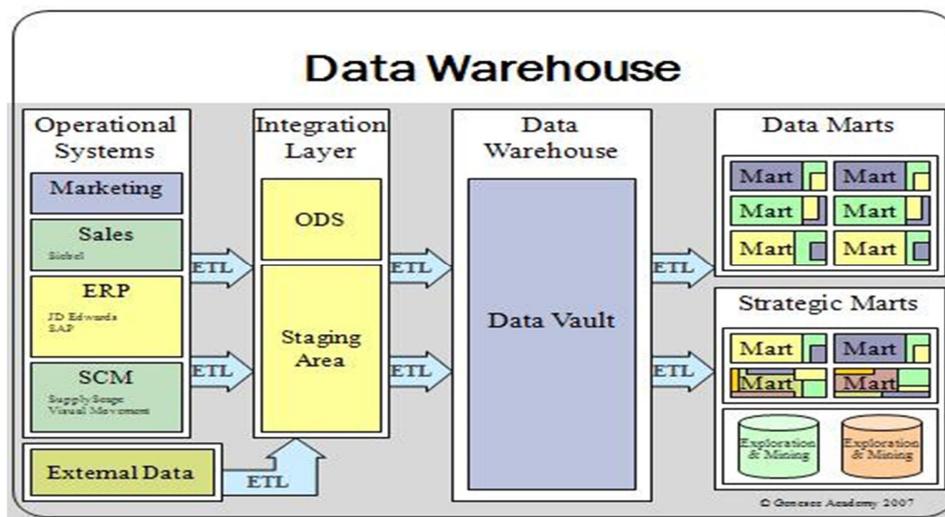


Figura 4 Arquitectura general de un datawarehouse

Fuentes: Extraído de la página de internet ⁵Lima, Perú

Como podemos observar en la figura 4 existe un proceso intermedio entre las fuentes de datos y el datawarehouse. A este proceso se lo conoce como ETL por sus siglas en ingles que significan Extract Transformation Load. Este proceso es muy importante para la construcción de un datawarehouse, ya que aquí es donde se van a limpiar y estandarizar los datos para su posterior almacenamiento. En un datawarehouse nunca se debe almacenar datos basura o inconsistentes, considerando que de aquí se tomaran los datos para realizar el posterior análisis de información.

⁵ http://es.wikipedia.org/wiki/Almac%C3%A9n_de_datos

De manera muy simplificada el proceso ETL consiste :

- **Extraer**

Convierte los datos a un formato preparado para iniciar el proceso de transformación. Se analiza los datos extraídos, de lo que resulta un chequeo para verificar si los datos cumplen la estructura que se esperaba. De no ser así los datos son rechazados.

- **Transformación**

Se aplica reglas de negocio sobre los datos extraídos para convertirlos en datos que serán cargados. Por ejemplo, se selecciona solo ciertas columnas para su carga y las con valores nulos se descartan.

- **Carga**

los datos son cargados en el datawarehouse de destino. Consiste en realizar un resumen de todas las transacciones comprendidas en el período de tiempo seleccionado y transportar el resultado como una única transacción hacia el datawarehouse.

A pesar de que se puede pensar que un datawarehouse es semejante a un OLTP, la verdad es que existen diferencias muy significativas entre ellos como podemos ver en la tabla 1.

Tabla 1 Diferencias entre una base de datos transaccional y un datawarehouse

Aspectos	BD operacional	Data Warehouse
Objetivo	De tipo operativo (operaciones del día a día)	Análisis y toma de decisiones
Proceso	De transacciones. Repetitivo y conocido.	De consultas masivas. Puntual y no conocido.
Actividad	Predomina la actualización	Predomina la consulta
Rendimiento	Importancia del tiempo de respuesta de la transacción instantánea	Importancia de la respuesta masiva.
Explotación	Explotación de la información relacionada con la operativa de cada aplicación	Explotación de toda la información interna y externa relacionada con el negocio
Volatilidad	Actualizable	Carga, pero no actualización
Usuarios	Usuarios de perfiles medios o bajos	Usuarios de perfiles altos
Organización	Estructura normalmente relacional	Visión multidimensional
Granularidad	Datos generales desagregados, al detalle	Datos en distintos niveles de detalle y agregación
Horizonte histórico	30 a 90 días	5 a 10 años

Fuentes: Elaboración propia

Una vez comprendida la definición de un datawarehouse, se puede comenzar a diseñarlo e implementarlo. Para construir un datawarehouse existen 2 tipos de modelado de datos los cuales son:

2.2.9. MODELO RELACIONAL

El modelo relacional se puede usar para implementar un datawarehouse aplicando sobre el modelo de datos de la empresa, un proceso de transformación de datos de 8 pasos, Los 8 pasos a considerarse son los siguientes:

1. Selección de los datos de interés. El modelo de datos de la empresa es una de las entradas en este paso, además existen otras como reportes, prototipos, consultas y requerimientos de información. Se debe tener cuidado en la selección de los datos para no sobresaturar la información que se almacene en el datawarehouse.
2. Añadir la dimensión de tiempo a las llaves. Debido a que el modelo del datawarehouse representa la información a lo largo del tiempo, se debe agregar el tiempo o fecha a la llave de cada entidad de interés.
3. Añadir datos derivados. Los datos derivados se obtienen como resultado de aplicar operaciones matemáticas a otros datos. Es necesario incluirlos en el datawarehouse por razones de optimización y para asegurar la consistencia de los datos.
4. Determinar el nivel de granularidad. El nivel de detalle que se almacenará en el datawarehouse puede variar dependiendo la perspectiva del negocio, técnica o del proyecto. Antes de determinar el nivel de granularidad, es necesario considerar varios factores tales como las necesidades del negocio, de los procesos de data mining, el costo de almacenamiento y el desempeño.
5. Totalizar los datos. Se utiliza con fines de rendimiento en la entrega de los resultados. Comúnmente se totalizan los datos por periodos de tiempo.
6. Mezclar entidades. Combinar 2 o más entidades en una sola, es decir, se desnormalizan los datos, para esto deben tener una llave en común.

7. Crear arreglos. Se utiliza poco, pero si es necesario, puede mejorar considerablemente la población de los datawarehouses.
8. Separar los datos. Se separan tablas con base en su estabilidad y uso.

Estos 8 pasos pueden dividirse en 2 categorías. Los primeros 4 pertenecen a la etapa de creación del datawarehouse y los 4 siguientes pasos sirven para mejorar el rendimiento y optimizar el tiempo de respuesta. Sobre el modelo relacional se pueden aplicar las técnicas de reporte general y permite almacenar el histórico de la información.

2.2.10. MODELO MULTIDIMENSIONAL

El modelo multidimensional, permite tener datos organizados entorno a hechos, que tienen unos atributos o medidas, que pueden verse con mayor o menor detalle según ciertas dimensiones. Los conceptos importantes que se manejan dentro del modelo multidimensional son:

- Hecho: corresponde a la actividad de la empresa que se desea representar, por ejemplo, las ventas.
- Medidas: son el conjunto de indicadores del hecho que se escogió para representar. Generalmente responden a la pregunta ¿Cuánto? Retomando el ejemplo anterior, las medidas para el hecho de las ventas podrían ser: ¿Cuántos productos se vendieron?, ¿Cuánto fue el total de la venta? ¿Cuánto costaron esos productos vendidos?
- Dimensiones: son las que van a caracterizar al hecho y responden a las preguntas ¿Dónde? ¿Cuándo? ¿Qué? Siguiendo con el ejemplo, las dimensiones para el hecho de las ventas podrían ser: la fecha de la venta, la hora, o el lugar.

- Granularidad: corresponde al nivel de detalle que será almacenado en las dimensiones. Por ejemplo: para la dimensión de tiempo podemos tener: año, semestre, trimestre, mes, semana, día, hora.

Una vez comprendidos estos conceptos, podemos definir los elementos de los que se compone el modelo multidimensional para su implementación: las tablas y los esquemas (Kimball⁶, 2008).

Existen dos tipos de tablas, que se muestran gráficamente en la figura 5. Tabla

- hecho: en ella se almacenan las medidas y las claves de las tablas de dimensión u otras medidas derivadas, conocidas como funciones de agregación.
- Tabla de dimensión: contiene los datos descriptivos de cada dimensión, también conocidos como atributos de la dimensión.

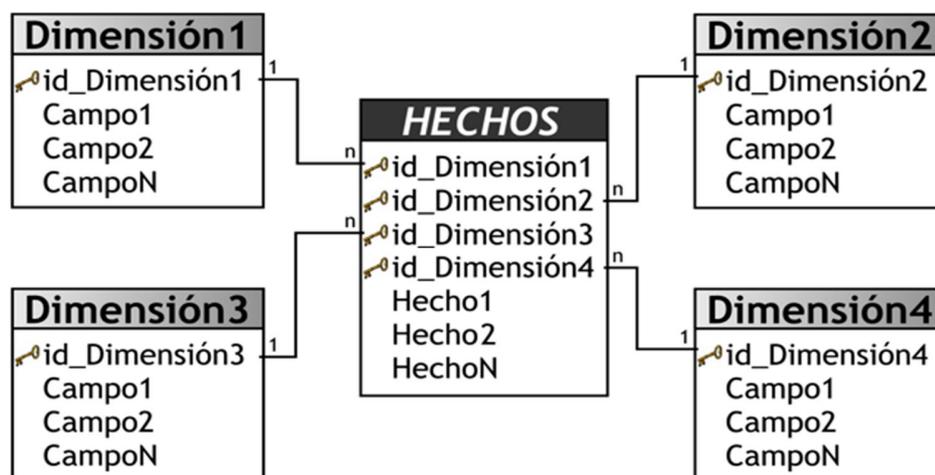


Figura 5 Ejemplo de tablas hecho y dimensión

Fuentes: Business Intelligence Competir con información, Banesto, España-2013

⁶ Kimball, R. (2008). *The datawarehouse lifecycle*. Usa: Wiley.

Los esquemas son colecciones de tablas y pueden ser de dos tipos (Kimball, 2008):

- **Esquema Estrella:** el centro de la estrella consiste en una tabla hecho y las puntas de la estrella son las tablas de dimensión, las cuales tienen una sola conexión a la tabla hecho a través de su llave primaria, que debe ser un atributo único, y no hay caminos alternativos en las dimensiones, es decir no existen jerarquías. Como vemos en la figura 6. Este esquema es una representación sencilla de los datos que agiliza el tiempo de respuesta en las consultas multidimensionales, sin embargo, dificulta el proceso de actualización de datos, porque usualmente está desnormalizado, lo que provoca duplicación de datos o sustitución de llaves por valores de registros y esto ocupa más espacio de almacenamiento.

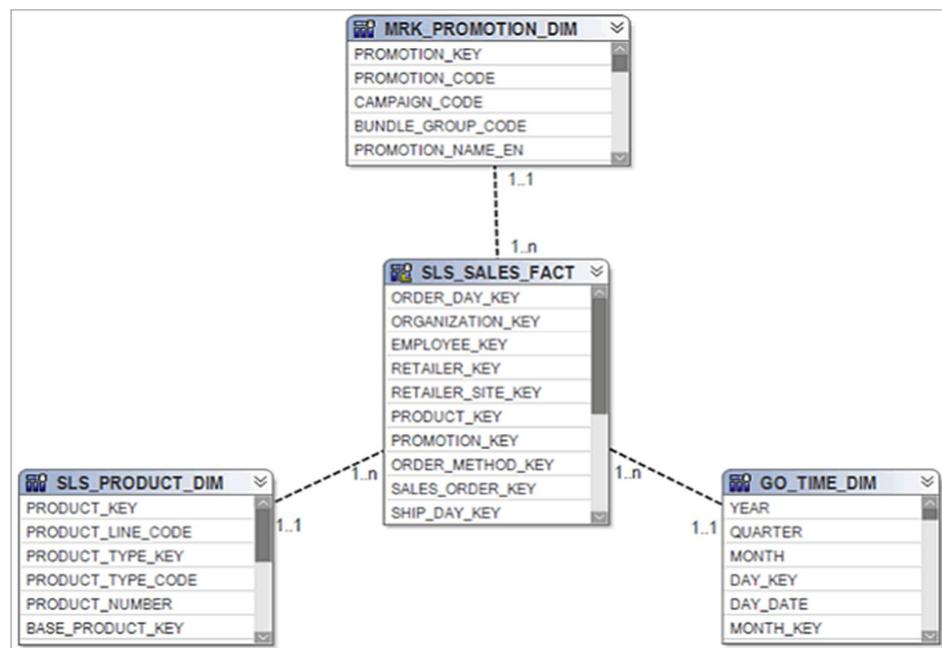


Figura 6 Ejemplo de esquema estrella

Fuente: Tesis ,Análisis Diseño E Implementación De Una Solución De Inteligencia Negocios Pontificia Universidad Católica Del Perú, 2011

- **Esquema Copo de Nieve:** La diferencia con el esquema estrella es que sí existen caminos alternativos en las dimensiones, es decir jerarquías en las dimensiones. Una de las ventajas de este esquema es que facilita la actualización de los datos del datawarehouse y ahorra espacio de almacenamiento. Sin embargo, es una representación de datos más compleja que disminuye el tiempo de respuesta en las consultas multidimensionales. La figura 7 es un ejemplo del esquema copo de nieve.

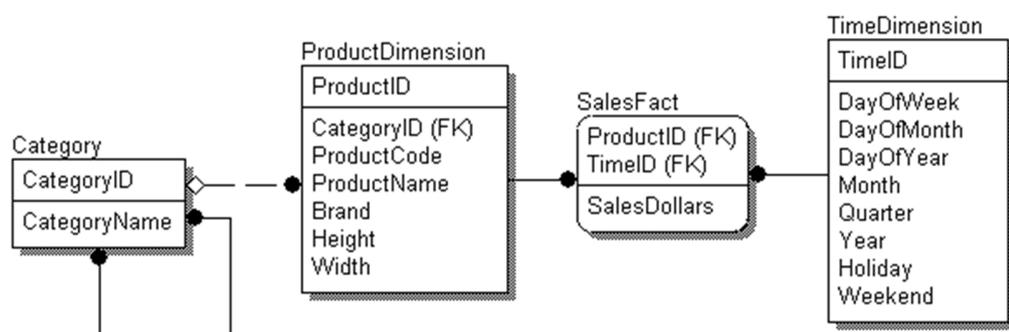


Figura 7 Ejemplo de esquema copo de nieve

Fuente: Tesis ,Análisis Diseño E Implementación De Una Solución De Inteligencia Negocios, Pontificia Universidad Católica Del Perú, 2011

- **Constelación:** es un conjunto de esquemas estrella o copos de nieve que comparten dimensiones, como podemos ver en la figura 8.

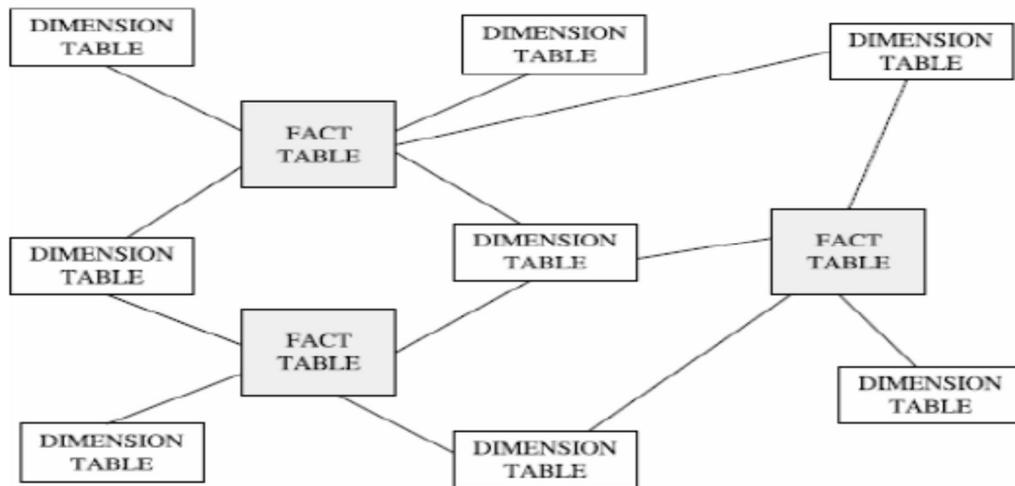


Figura 8 Ejemplo de esquema constelación

Fuente: Tesis ,Análisis Diseño E Implementación De Una Solución De Inteligencia Negocios, Pontificia Universidad Católica Del Perú, 2011

Existen 4 pasos importantes que se deben considerar para la elaboración de un modelo multidimensional (Universidad de Concepción Chile, 2008):⁷

1. Selección del proceso a modelar.
2. Seleccionar el hecho central y el gránulo máximo que se va a necesitar sobre él.
3. Identificar las dimensiones que caracterizarán el dominio.
4. Determinar y refinar las medidas y atributos a almacenar sobre el proceso.

El modelo multidimensional también permite almacenar el histórico de la información, una ventaja importante es que los esquemas estrella ayudan a enfocarse en hechos particulares y permiten crecer tanto como se quiera. La figura

⁷ Universidad de Concepción Chile. (2 de Febrero de 2008). Departamento de Ingeniería Informática.

9 presenta gráficamente las diferencias entre ambos modelos (Inmon⁸, 2005) (Kimball⁹, 2008)

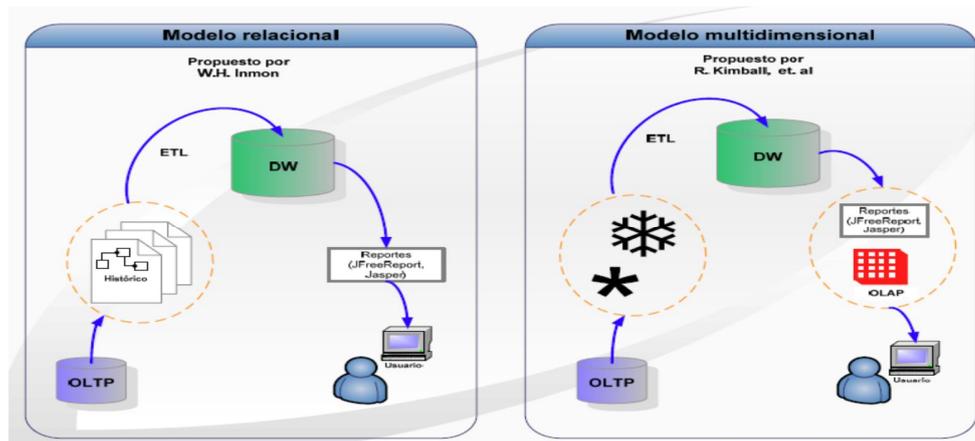


Figura 9 Diferencia entre los modelos relacional y multidimensional

Fuentes: Tesis, Desarrollo de una Aplicación de Business Intelligence Para la Empresa Empaqplast, Ecuador, 2012

A diferencia del modelo relacional, sobre el modelo multidimensional se pueden construir cubos de OLAP, que serán explicados en la siguiente sección.

2.2.11. OLAP

On Line Analytical Processing (OLAP) permite analizar grandes cantidades de datos a través del modelo multidimensional, explicado en la sección anterior.

Esta representación permite mostrar los datos al usuario final de una manera más sencilla y tiene la flexibilidad necesaria para cambiar las perspectivas de visión de la información. OLAP permite realizar análisis históricos complejos con amplia manipulación de los datos (Ponniiah¹⁰, 2007).

⁸ Inmon, W. (2005). *Building the Datawarehouse*. Usa: Wiley

⁹ Kimball, R. (2008). *The datawarehouse lifecycle*. Usa: Wiley.

¹⁰ Ponniiah, P. (2007). *Data modeling fundamentals*. Usa: Wiley.

El análisis de la información se realiza mediante cubos, que son colecciones de dimensiones y medidas, alrededor de un hecho particular, sobre los cuales se aplican distintos operadores para dar los resultados a las consultas que se ejecuten. En la figura 10 se puede observar que un cubo se compone de ejes, representados por las dimensiones y celdas que son las medidas que se quieren analizar.

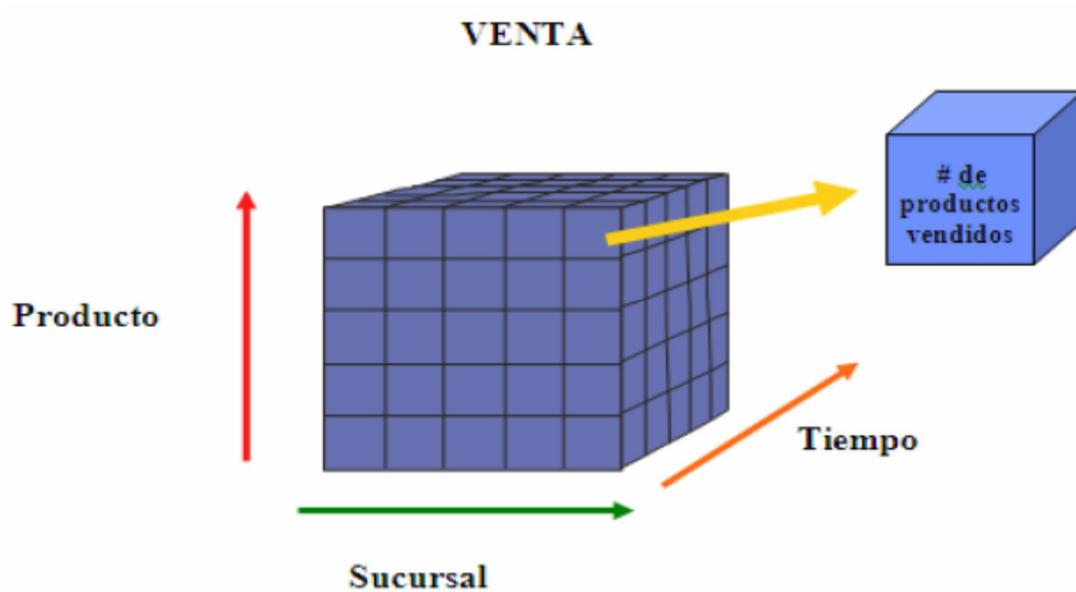


Figura 10 Cubo de OLAP

Fuente: Tesis Desarrollo De Una Solución Business Intelligence En La Pyme Comercializadora De Calzado Stefy Rall, Ecuador, 2013

En la figura 11 podemos observar ejemplos de reportes obtenidos con los cubos de OLAP.



Figura 11 Ejemplos de reportes de OLAP

Fuente: Tesis Desarrollo De Una Solución Business Intelligence En La Pyme Comercializadora De Calzado Stefy Rall, Ecuador, 2013

Existen 2 técnicas de almacenamiento o implementación de cubos de OLAP, como se muestra en la figura 12, que son (Harjinder, 1996):¹¹

- ROLAP: físicamente el datawarehouse se construye sobre una base de datos relacional. Una ventaja de este tipo de esquema es que se pueden utilizar los sistemas de administración de bases de datos relacionales, que son muy populares para el OLTP, y sus herramientas asociadas, además de que el costo necesario para la implementación es mucho menor.
- MOLAP: físicamente el datawarehouse se construye sobre estructuras basadas en matrices multidimensionales. Las ventajas de este esquema son la especialización y la correspondencia entre el nivel lógico y el nivel

¹¹ Harjinder, G. (1996). *La integración de información para la mejor toma de decisiones*. Mexico: Prentice Hall.

físico. Por esto generalmente MOLAP es más eficiente, debido a que es una solución ad-hoc.

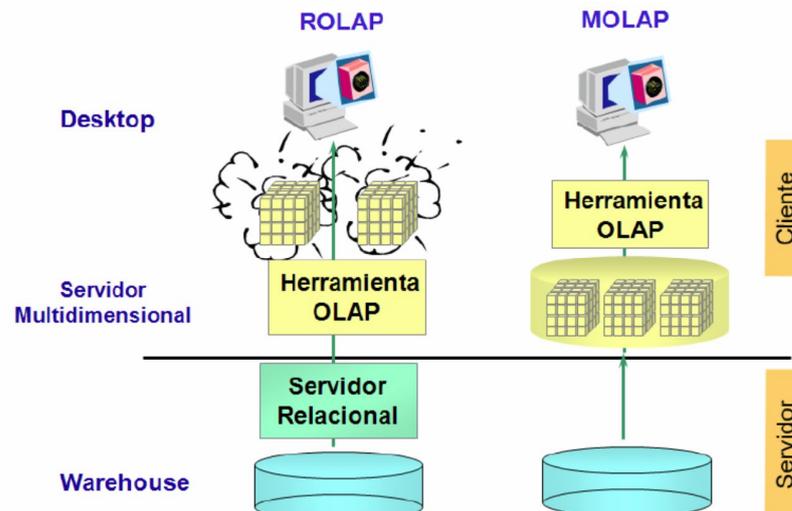


Figura 12 Técnicas de almacenamiento de cubos ROLAP y MOLAP

Fuente: Tesis Desarrollo De Una Solución Business Intelligence En La Pyme Comercializadora De Calzado Stefy Rall, Ecuador, 2013

La diferencia entre ambas técnicas radica en la implementación física y no en la manera en que las herramientas muestren los resultados de las consultas.

Entre las características principales de OLAP se encuentran las siguientes:

- Presenta una visión multidimensional lógica de los datos en el datawarehouse. La visión es independiente de cómo se almacenan los datos.
- Incluye siempre la consulta interactiva y el análisis de los datos. Por lo regular la interacción es de varias pasadas, lo cual comprende la profundización en niveles cada vez más detallados o el ascenso a niveles superiores de resumen y adición.

- Ofrece opciones de modelado analítico, incluyendo un motor de cálculo para obtener proporciones, desviaciones, etc., que comprende mediciones de datos numéricos a través de muchas dimensiones.
- Crea resúmenes, adiciones, jerarquías y cuestiona todos los niveles de adición y resumen en cada intersección de las dimensiones.
- Maneja modelos funcionales de pronóstico, análisis de tendencias y análisis estadísticos.
- Recupera y exhibe datos tabulares en dos o tres dimensiones, cuadros y gráficas. Esto permite analizar los datos desde diferentes perspectivas.

Para explotar la información almacenada en el datawarehouse una vez definido el cubo, se pueden aplicar distintos operadores que se muestran en la tabla 2.

Tabla 2 Operadores de OLAP

Operador	Significado
Drill	Ofrece mayor nivel de detalle y menos agregación.
Roll	Lo contrario a drill, se tiene mayor agregación y menor nivel de detalle.
Slice & Dice	Se proyecta datos de área específicas por selección, no por agregación.
Pivot	Se reorientan las dimensiones, es decir, las columnas ocupan el lugar de filas y las filas el de las columnas.
Drill-down/roll-up	Se aumentan o disminuyen agregación dentro de una consulta y predefina.
Drill-across/roll-across	Se obtiene agregaciones en otras dimensiones que no hayan sido comprendidas inicialmente, o se desaparecen dimensiones.

Fuente: Elaboración Propia

En la figura 13 podemos ver un ejemplo de la aplicación de un operador de OLAP.

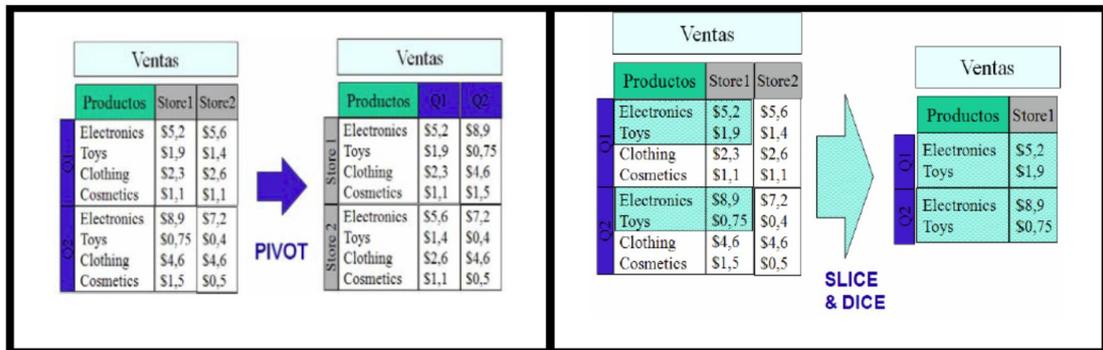


Figura 13 Operadores de OLAP

Fuentes: obtenido de la Tesis Desarrollo de una Aplicación de Business Intelligence para la empresa Empaqplast. Escuela Politécnica del Ejército. Ecuador, 2012.

Una vez que hemos visto la definición de OLAP y los operadores de manipulación de datos que existen, podemos pasar a la realización de consultas, las cuales se definen en lenguaje SQL que se explicará en la siguiente sección.

- **SQL STRUCTURED QUERY LENGUAJE**

El lenguaje de consulta estructurado o SQL, es un lenguaje de acceso a bases de datos relacionales que permite especificar diversos tipos de operaciones. Una de sus características es el cálculo relacional, permitiendo efectuar consultas con el fin de recuperar información de interés de una base de datos, así como también hacer cambios sobre ella (Microsoft, 2010).

Un ejemplo sencillo de consulta en lenguaje SQL luce como se muestra en la figura 14 y cuyo resultado se muestra en la figura 15. (Usando la base de datos Northwind)

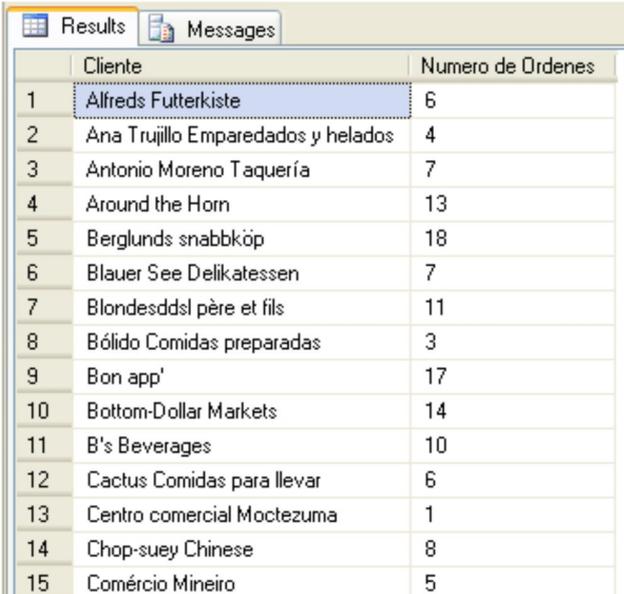
```

/*cuantas ordenes ha hecho cada cliente*/
select    c.companyname    as    [Cliente],
count(o.orderid) as [Numero de Ordenes]
from customers c join orders o on c.customerid =
o.customerid

```

Figura 14 Consulta en lenguaje SQL

Fuentes: Elaboración Propia



	Ciente	Numero de Ordenes
1	Alfreds Futterkiste	6
2	Ana Trujillo Emparedados y helados	4
3	Antonio Moreno Taquería	7
4	Around the Horn	13
5	Berglunds snabbköp	18
6	Blauer See Delikatessen	7
7	Blondesddsl père et fils	11
8	Bólido Comidas preparadas	3
9	Bon app'	17
10	Bottom-Dollar Markets	14
11	B's Beverages	10
12	Cactus Comidas para llevar	6
13	Centro comercial Moctezuma	1
14	Chop-suey Chinese	8
15	Comércio Mineiro	5

Figura 15 Resultado de la consulta SQL

Fuente: Elaboración propia

En conclusión, el objetivo de OLAP es ayudar al usuario final a entender lo que está sucediendo en la empresa, ya que permite mostrar un análisis concentrado de los datos de la empresa. Esto ayuda a los administradores a mantenerse informados de la situación actual de la organización. Otro tipo de análisis es la

técnica de data mining, que también utiliza los datos almacenados en el datawarehouse y que veremos con más detalle en la siguiente sección.

2.2.12. DATA MINING

Es el proceso de extraer información útil y comprensible, previamente desconocida, desde grandes cantidades de datos almacenados en distintos formatos (University of California, 2010).¹²

Para que este proceso sea efectivo debería ser automático o semiautomático y el uso de los patrones descubiertos debería ayudar a tomar decisiones más seguras. Por lo tanto, el objetivo de data mining es descubrir patrones válidos, novedosos, interesantes y comprensibles, que reporten algún beneficio a la organización.

Dos conceptos importantes en data mining son las tareas y los métodos. En las secciones siguientes se explican con más detalle los tipos de tareas y métodos que existen para resolverlas.

- **TAREAS**

Una tarea es un tipo de problema de data mining. Por ejemplo: clasificar productos en defectuosos, no defectuosos, defectuosos reparables y defectuosos no reparables es una tarea.

Existen dos grandes grupos en los que se pueden dividir las tareas, como se pueden ver en la tabla 3.

¹² University of California. (17 de noviembre de 2010). UCLA. Recuperado el 19 de abril de 2012

Tabla 3 Tareas de data mining

Tipo de tarea	Descripción	Ejemplo
Predictivas	Se trata de problemas en los que hay que predecir uno o más valores para uno o más ejemplos	Clasificación
		Clasificación suave
		Estimación de probabilidad de clasificación
		Categorización
		Preferencias o priorización
		Regresión
Descriptivas	Los ejemplos se presentan como un conjunto de datos sin ordenar ni etiquetar de ninguna manera. Por lo tanto, el objetivo, no es predecir nuevos datos sino describir los existentes	Clustering
		Correlación y factorizaciones
		Reglas de asociación
		Árboles de decisión de MS
		Detección de valores e instancias anómalas

Fuente: Elaboración Propia

De todas estas, la tarea que se aborda en esta tesis es descriptiva, específicamente Árboles de Decisión de Microsoft. Tomando en cuenta que la herramienta que usamos para minería de datos es Microsoft Excel, ya que Excel posee un complemento exclusivo para tareas de minería de datos, que permite utilizar la tarea mencionada.

2.3 MARCO CONCEPTUAL

2.3.1. DATOS DE LA INSTITUCIÓN

- **Razón Social:** Triton Trading S.A.
- **Ubicación:** Km 17.5 de antigua panamericana sur
- **Rubro Económico:** Venta de equipo (logísticos, portuarios, construcción) y repuestos
- **Ruc:** 20101369898
- **Página Web:** <http://www.triton.com.pe/>
- **Organigrama:**

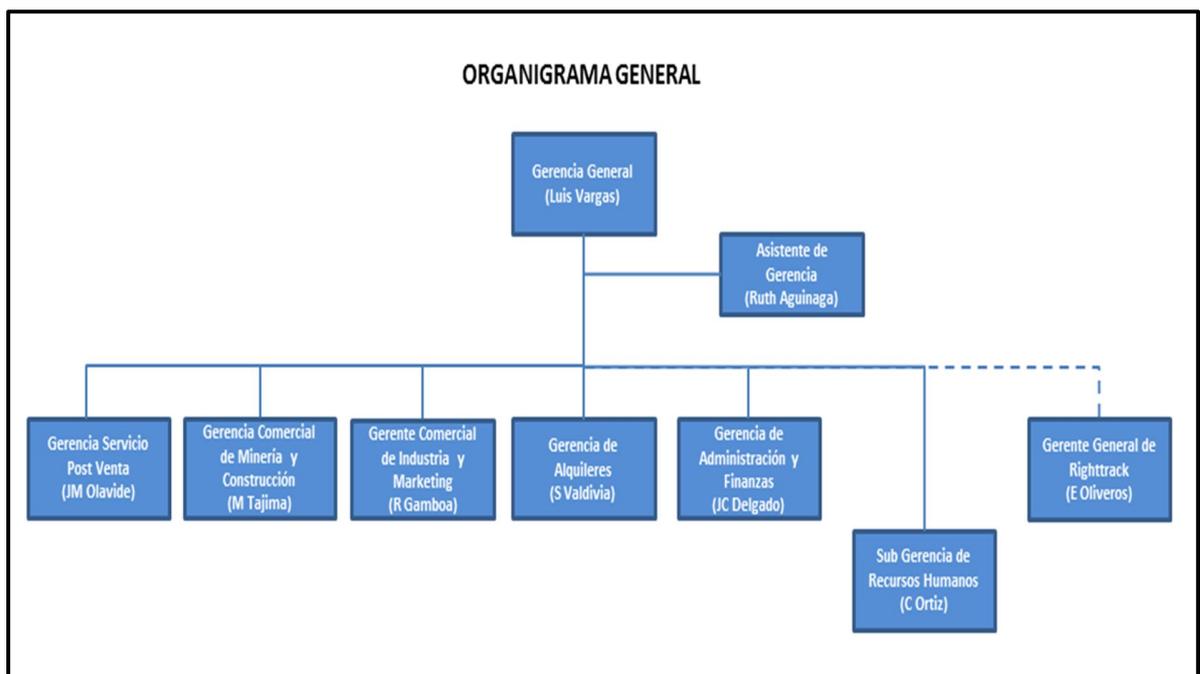


Figura 16 de organigrama de la empresa

Fuentes: RRHH de la empresa Triton Rading SA, Lima Perú

2.3.2. RESEÑA DE LA EMPRESA

Triton Trading S.A. inició operaciones en Septiembre del año 1990 con la visión de satisfacer el sector portuario con equipos y servicio técnico especializado. A través de los años han mantenido un crecimiento sostenido que les ha permitido desarrollar y consolidar la División de venta equipos y repuestos para los rubros de Industria, Minería y Construcción.

Loas grades pilares o líneas de negocio de la empresa son:

- Venta de equipo
- Venta de repuestos
- Venta de servicios
- Alquiler de equipos

En la actualidad la empresa cuenta con tres sucursales a nivel nacional.

La primera sucursal y la más grande se encuentra en la ciudad de Piura atendiendo al sector logístico y portuario de zonal.

La segunda sucursal se encuentran en ciudad de Cajamarca a tendiendo al sector minero de la zona.

La tercera y la más pequeña se encuentran en la ciudad de Arequipa atendiendo al sector minero y logístico.

A de más cuenta con punto de servicio en de parlamento de lima en el distrito de San Martin de Porres y la provincia constitucional del Callo.



Figura 17 sucursales

Fuente: Área Comercial de la empresa Triton Trading SA., Lima, Peru

La experiencia de más de 23 años les ha permitidos ser representantes de marcas líderes a nivel mundial, que ofrecen total respaldo, confianza, garantía de fábrica, y servicio Post-Venta:



Figura 18 marcas presentadas por la empresa

Fuente: Área Comercial de la empresa Triton Trading SA., Lima, Peru

En la actualidad la cantidad de equipos colocados entre alquiler y venta, que asciende a más de 1,500 unidades a nivel nacional con alta satisfacción y fidelidad de sus nuestros clientes.¹³

¹³ Extraído de <http://www.triton.com.pe/>

2.3.3. UNIDADES DE NEGOCIO

Las técnicas de inteligencia de negocios son universalmente aplicadas a cualquier área de negocios y los procesos departamentales inservibles pueden ser significativamente mejorados o mejor gestionados con la rápida información que proporciona la inteligencia de negocios. Según Vitt, Luckevich y Misner (2003) es una estructura organizacional, en la cual un conjunto coherente de actividades funcionales se ejecutan dentro de una línea de negocios.

2.3.4. TOMA DE DECISIONES

Es el proceso mediante el cual se realiza una elección entre las alternativas o formas para resolver diferentes situaciones de la vida”, ya que se pueden presentar en diferentes contextos: a nivel laboral, familiar, sentimental o empresarial. Por otra parte, Méndez (2006) explica seis criterios para tomar una decisión eficaz:

- Concentrarse en lo que realmente importa.
- Realizar el proceso de forma lógica y coherente.
- Considerar tanto los elementos objetivos como los subjetivos y utilizar una estructura de pensamiento analítica e intuitiva.
- Recoger la información necesaria para optar o elegir.
- Recopilar las informaciones, opiniones, que se han formado en torno a la elección.
- Ser directos y flexibles antes, durante y después del proceso.

- **ESTRUCTURA JERÁRQUICA DE LA TOMA DE DECISIONES EN LAS ORGANIZACIONES**

El rol que ocupa el personal de las empresas se agrupa en alguno de los tres niveles jerárquicos de autoridad y responsabilidad, otorgándole atribuciones específicas dentro del marco de competencias del proceso administrativo. Según Peña (2006) ¹⁴tradicionalmente los niveles jerárquicos se representan como una pirámide dividida en tres partes, para destacar el orden ascendente de autoridad, el impacto de las decisiones y competencias de las funciones que cada rango realiza. Estos niveles jerárquicos se conocen como:

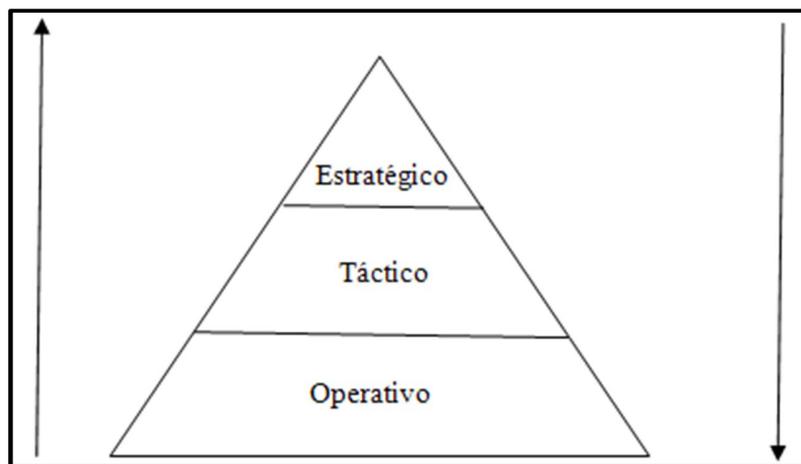


Figura 19 Nivel Jerárquico de una empresa por Peña ,2006.

Fuentes: Tesis Inteligencia De Negocios Y Toma De Decisiones De Los Gerentes En La Banca Universal En Venezuela, 2011

¹⁴ <http://www.joaquinpenasiles.com/blog/category/empresa-2-0/>

- **EL PROCESO DE LA TOMA DE DECISIONES**

El proceso de TD puede definirse de acuerdo a las etapas y resultados que se ilustran en la Figura 20, denominada "El proceso de toma de decisiones", cuya representación y explicación se ofrece a continuación:

La Investigación es una tarea de discernimiento e interpretación compuesta por: Identificación de problemas, el rol del tomador de decisiones y la formulación de problemas, de acuerdo con la siguiente descripción.

Identificación de problemas. Busca alguna diferencia entre la situación existente y un estado deseado. Es decir, compara el modelo del estado esperado con el existente, precisa y evalúa las diferencias para determinar si existe un problema esperado con el existente, precisa y evalúa las diferencias para determinar si existe un problema. Por ejemplo, Pounds usa cuatro modelos para desarrollar expectativas frente a las cuales se compara la realidad.

- Históricos. Las expectativas se delinear como resultado de las experiencias anteriores.
- De planeación. La expectativa está definida por el plan.
- Otras personas en la empresa. Son las expectativas de terceros.
- Extra organizacionales. Se derivan de la competencia, clientes y mercado entre otros.

El rol del tomador de decisiones. Es la función que está vela no solamente de la aparición de las diferencias entre la expectativa y la realidad, sino también a prevenir a que esto no ocurra; asumiendo entonces un papel pro y reactivo de acuerdo con las circunstancias. La formulación de problemas. Para resolver un problema es indispensable identificar su origen, desarrollo y resultados que se han producido o están por suceder. Esta definición debe ser clara, procurando reducir la complejidad conforme a las siguientes estrategias:

- Precisión de los límites. Identifica claramente los elementos implicados en el problema.
- Examen de los argumentos. Los cuales pueden haber precipitado el problema.
- Descomposición del problema. En varios problemas más pequeños y específicos.
- Concentración. En los elementos controlables.

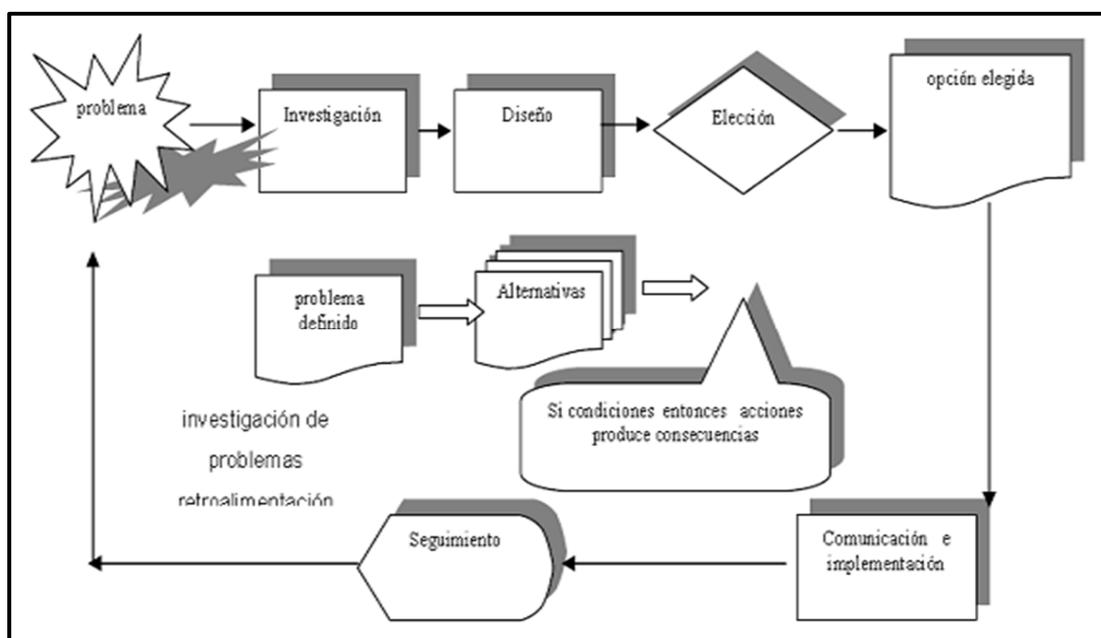


Figura 20 toma de decisiones

Fuentes: Tesis Inteligencia De Negocios Y Toma De Decisiones De Los Gerentes En La Banca Universal En Venezuela, 2011

2.3.5. APLICACIÓN A LA TOMA DE DECISIONES

El área de apoyo a la toma de decisiones constituye un área multidisciplinar cuyo objetivo es la introducción de métodos y/o herramientas que ayuden a las personas en la toma de decisiones clave.

La fase de toma de decisiones usualmente se refiere al proceso necesario para realizar la selección de una opción o alternativa. Este proceso incluye: conocer el problema, recoger información sobre el problema, identificar alternativas, anticipar consecuencias de posibles decisiones, realizar la selección utilizando juicios lógicos y coherentes basados en la información disponible.

Entonces se puede definir el área del apoyo a la toma de decisiones en 2 partes: la primera que concierne a la toma de decisiones por parte del personal involucrado y la segunda que corresponde al estudio de técnicas que asistan a las personas a mejorar las decisiones tomadas (Sprague, 1995). Entre esas técnicas podemos ubicar a: datawarehousing, OLAP y data mining, que nos ayudan a formar un sistema completo de inteligencia empresarial.

Existen diversas áreas de aplicación en las que ya se han incorporado estas técnicas para apoyar la toma de decisiones, las más importantes se muestran en la tabla 4.

Tabla 4 Áreas de aplicación de apoyo a la toma de decisiones

Area de aplicación	Ejemplos
Aplicaciones financieras	Obtención de patrones de uso fraudulento de tarjetas de crédito, determinación del gasto en tarjeta de crédito por grupos, cálculo de correlaciones entre indicadores financieros, análisis de riesgo en créditos.
Análisis de mercado, distribución y comercio	Análisis de la canasta básica de mercado, evaluación de campañas publicitarias, análisis de la fidelidad de los clientes, estimación de inventarios, costos y ventas.
Seguros y salud privada	Determinación de clientes potencialmente caros, identificación de patrones de comportamiento para clientes con riesgo, identificación de comportamiento fraudulento, predicción de clientes que podrían ampliar su póliza.
Educación	Selección o captación de estudiantes, detección de abandonos y fracasos, estimación de tiempo de estancia en la institución.
Procesos industriales	Extracción de modelos sobre comportamiento de compuestos, detección de piezas con defectos, predicción de fallos y accidentes, estimación de composiciones óptimas en mezclas, extracción de modelos de costos, extracción de modelos de producción.
Medicina, biología, bioingeniería y otras ciencias	Diagnóstico de enfermedades, detección de pacientes con riesgo de sufrir una enfermedad concreta, recomendación priorizada de fármacos para una misma enfermedad, predecir si un compuesto químico causa cáncer, clasificación de cuerpos celestes, predicción del recorrido y distribución de inundaciones, modelos de calidad de aguas.
Telecomunicaciones	Establecimiento de patrones de llamadas, modelos de carga en redes, detección de fraude.

Fuente: Elaboración Propia

2.3.6. HERRAMIENTAS Y SUITES DE INTELIGENCIA DE NEGOCIOS

En la actualidad se han desarrollado diversas herramientas que ofrecen soluciones a los problemas de Inteligencia de Negocios y que implementan todas o alguna de sus tecnologías. Para poder entender mejor el conjunto de herramientas para Inteligencia de Negocios, en esta tesis se ha realizado una clasificación de dichas herramientas, las cuales se muestran en la tabla 5 en donde los sistemas se dividen en 2 clases:

Tabla 5 Clasificación de sistemas de inteligencia empresarial

Clasificación	Definición
Herramientas simples	Herramientas especializadas en una sola técnica de inteligencia empresarial
Suites	Ambientes completos que ofrecen la implementación de las 3 técnicas de inteligencia empresarial: datawarehousing, OLAP y data mining.

Fuente: Elaboración Propia

A continuación veremos las herramientas simples y las suites más populares que existen en el mercado para cada una de las clasificaciones propuestas, junto con sus características principales.

- **HERRAMIENTAS SIMPLES**

Las herramientas simples son aquellas que dan soporte a las tecnologías de la inteligencia empresarial individualmente. Estas herramientas han sido clasificadas a su vez de acuerdo a la tecnología de inteligencia empresarial a la que dan soporte, en seguida mencionaremos las más importantes y populares para cada una.

- **Herramientas de integración de datos**

En la figura 21 vemos algunas de las herramientas más populares que apoyan la integración de datos, las cuales se describirán con más detalle en seguida.



Figura 21 Herramientas de integración de datos (ETL)

Fuente: obtenido de la Tesis Desarrollo de una Aplicación de Business Intelligence para la empresa Empaqplast. Escuela Politécnica del Ejército. Ecuador, 2012.

- **Clover ETL**

Es un ambiente de transformación de datos de código abierto basado en Java, para datos estructurados, capaz de funcionar como aplicación independiente (standalone) o estar incluida en otra aplicación.

- Características principales: Maneja todas las bases de datos con conexión JDBC disponible para ellas.
- Realiza transformaciones gráficas basadas en XML para la descripción de los metadatos de los registros.
- Soporta valores nulos.

- Puede ejecutarse en múltiples CPU utilizando una estrategia llamada “pipeline- parallelism”.
- Se distribuye bajo la licencia LGPL.
- Clover ETL soporta 4 diferentes tipos de datos: string, numeric, date, bytes.

- **Kettle**

Conocido actualmente como Pentaho Data Integration, es un proyecto belga de código abierto, ahora adoptado por Pentaho BI, (del cual hablaremos con más detalle en la sección de suites), que incluye un conjunto de herramientas para realizar ETL. Uno de sus objetivos es que el proceso de ETL sea fácil de generar, mantener y desplegar .

Se compone de 4 herramientas:

- Spoon: permite diseñar de forma gráfica la transformación ETL.
- Pan: ejecuta un conjunto de transformaciones diseñadas con SPOON, conocidas como trabajos (Jobs), creando dependencias entre dichas transformaciones.
- Chef: permite mediante una interfaz gráfica, diseñar la carga de datos incluyendo un control de estado de los trabajos.
- Kitchen: permite ejecutar los trabajos batch diseñados con Chef.

Características principales:

- Funciona en Windows, UNIX and Linux.

- Tiene una interfaz visual con indicadores de las transformaciones.
 - Es una aplicación escrita en Java con algunas características avanzadas escritas en JavaScript.
 - Ofrece una licencia pública GPL.
 - Soporta Oracle, DB2, SQL Server y Sybase así como MySQL, Postgres, Hypersonic, FireBird SQL e Ingres. También soporta la conectividad con SAP con una tarifa de licencia.
 - Con respecto a escalabilidad, soporta la arquitectura de procesamiento en paralelo para distribuir las tareas de ETL a través de múltiples servidores.
- **Talend**

Talend Data Integration, es una herramienta de ETL de código abierto, que efectúa operaciones tales como alimentación de un datawarehouse, sincronización de bases de datos, transformación y verificación de la integridad de datos. Su interfaz gráfica de usuario está basada completamente en Eclipse RCP (Rich Client Platform) e incluye numerosos componentes para procesos de modelado de negocios, así como implementaciones técnicas para extracción de información, transformación y mapeo del flujo de datos. Las funciones generales de Talend incluyen modelador de negocios, diseñador del trabajo y administrador de metadatos.

Características principales:

Compatible con Microsoft SQL Server Integration Services (SSIS).

Funciona en Windows, Unix y Linux. Interfaz de diseño drag and drop para componentes, conectores y relaciones. Genera componentes en Perl, Java o SQL para ETL o transformaciones ETL.

Tiene una comunidad de soporte en Talend wiki, Talend Forum y un rastreador de errores.

Conectividad con Oracle, DB2, MySQL, Sybase y Postgres y conectividad ODBC para otras bases de datos.

- **Herramientas para OLAP**

En la figura 22 se muestran algunos de los proveedores de OLAP más populares que se describen más detalladamente a continuación



Figura 22 Herramientas de OLAP

Fuente: obtenido de la Tesis Desarrollo de una Aplicación de Business Intelligence para la empresa Empaqplast. Escuela Politécnica del Ejército. Ecuador, 2012.

- **Mondrian**

Es un motor de ROLAP desarrollado en Java, que permite analizar grandes conjuntos de datos que se encuentran almacenados en el datawarehouse. Mondrian se considera un motor, porque se encarga de recibir consultas dimensionales en lenguaje MDX y devolver los datos del cubo que correspondan a la consulta. El cubo se representa como un conjunto de metadatos que definen cómo se han de mapear estas consultas dimensionales a sentencias SQL para obtener de la base de datos la información necesaria para satisfacer la consulta dimensional. Utiliza una memoria caché para almacenar los resultados de las consultas que se acceden múltiples veces

Mondrian es usado para:

- Alto desempeño, análisis interactivo de grandes o pequeños volúmenes de información.
- Exploración dimensional de los datos, por ejemplo analizando ventas por marcas de productos, región o periodo de tiempo.
- Parseo de expresiones en lenguaje MDX a expresiones en SQL para recuperar respuestas a consultas dimensionales.
- Cálculos avanzados utilizando las expresiones de cálculo del lenguaje MDX.

Para definir el esquema que utilizará Mondrian existe una herramienta denominada Cube Designer, que está basada en un ambiente gráfico que permite diseñar un documento XML que representa un cubo o hipercubo para analizar la información almacenada de forma multidimensional. Este esquema será interpretado por Mondrian para obtener la información de las consultas de MDX ejecutadas.

- **Palo**

Es un servidor de base de datos multidimensional capaz de centralizar y administrar casi un número infinito de hojas de cálculo. El sistema opera en tiempo real, soporta la consolidación de jerarquías así como numerosas funciones de inteligencia empresarial y es un servidor de código abierto.

Palo es un servidor de datos multidimensional (MOLAP) orientado a celdas, específicamente desarrollado para almacenamiento y análisis de datos en hojas de cálculo.

- **Olap4j**

Es una interfaz de aplicación para el ambiente Java 2 Platform, Enterprise Edition, que soporta la creación, almacenamiento y administración de datos para una aplicación OLAP. Hyperion, IBM, y Oracle iniciaron el desarrollo de OLAP4J con la intención de que fuera un equivalente a la conexión JDBC pero específica para OLAP, que se liberó en 2009 y es compatible con Mondrian.

- **Herramientas para data mining**

En la figura 23 vemos algunos de los proveedores más populares para data mining. Estos y otros más se describen con más detalle enseguida.



Figura 23 Herramientas de data mining

Fuente: obtenido de la Tesis Desarrollo de una Aplicación de Business Intelligence para la empresa Empaqplast. Escuela Politécnica del Ejército. Ecuador, 2012.

- **Clementine**

La empresa SPSS, además de su conocido paquete estadístico, distribuye Clementine, uno de los sistemas de data mining más populares del mercado. Se trata de una herramienta visual inicialmente desarrollada por ISL (Integral Solutions Limited). En la actualidad esta herramienta, comercializada por SPSS, posee una arquitectura distribuida cliente / servidor.

Sus características principales son:

- Acceso a datos: fuentes de datos ODBC, tablas Excel, archivos planos ASCII y archivos SPSS.

- Pre procesado de datos: pick & mix, muestreo, particiones, reordenación de campos, nuevas estrategias para la fusión de tablas, etc.
 - Técnicas de aprendizaje: árboles de decisión, redes neuronales, agrupamiento, reglas de asociación, regresión lineal y logística, combinación de modelos.
 - Técnicas para la evaluación de modelos guiadas por las condiciones especificadas por el experto.
 - Visualización de resultados: ofrece un potente soporte gráfico que permite al usuario tener una visión global de todo el proceso, que comprende desde el análisis del problema hasta la imagen final del modelo aprendido.
 - Exportación: generación automática de informes (HTML y texto), volcado de los resultados del ejercicio de data mining en bases de datos, exportación de los modelos a distintos lenguajes (C, SPSS, HTML, estándar PMML, SQL para árboles de decisión y reglas).
- **Rapid Miner**

Es un entorno para la realización de experimentos de aprendizaje automático implementado en Java por la Universidad de Dortmund. Las operaciones elementales se encapsulan en los llamados operadores, los cuales son configurables mediante archivos XML. Estos archivos a su vez, pueden ser especificados gráficamente

El sistema incluye operadores para:

- Importación y pre-procesamiento de datos.
- Aprendizaje automático: máquinas de vectores de soporte, árboles de decisión, agrupamiento y algoritmos genéticos.
- Validación de modelos.

- **Weka**

Es una herramienta de código abierto desarrollada por un equipo de investigadores de la Universidad de Waikato en Nueva Zelanda, implementada en Java (University of Waikato, 2011). Sus características principales son:

- Acceso a datos: los datos son cargados desde un archivo en formato ARFF (archivo plano organizado en filas y columnas). El usuario puede observar en sus diferentes componentes gráficos, información de interés sobre el conjunto de muestras.
- Pre procesado de datos: existe una gran cantidad de filtros disponibles tales como: selección de atributos, discretización, tratamiento de valores desconocidos, transformación de atributos numéricos.
- Métodos y técnicas: árboles de decisión, tablas de decisión, vecinos más próximos, máquinas de vectores de soporte, reglas de asociación, métodos de agrupamiento y modelos combinados.
- Puede utilizarse como aplicación independiente o dentro de otra aplicación.

- **Xelopes**

Xelopes es una librería con licencia pública GNU para el desarrollo de aplicaciones de data mining, implementada por Prudsys AG en colaboración con Russian MDA specialist Zsoft Ltd. La librería implementa de una manera eficiente la mayoría de los algoritmos de aprendizaje, para desarrollar aplicaciones particulares de data mining. Además Xelopes es extensible, es decir, el usuario puede incorporar sus propios métodos a la propia librería (Prudsy¹⁵, 2011). Sus características principales son:

- Acceso a datos: existe una clase especial que permite dar uniformidad a todos los modos de accesos de datos permitidos. Así, el usuario puede acceder a archivos .log, archivos de bases de datos o implementar su propio formato de datos.
- Modelos: árboles de decisión lineales y no lineales, máquinas de vectores de soporte, redes neuronales, métodos de agrupamiento, métodos de reglas de asociación.
- Exportación de datos: existen métodos para exportar los modelos y sus resultados a otros entornos de data mining, soportando el estándar PMML.

- **SUITES**

Las suites, son sistemas que ofrecen la implementación de todas las tecnologías de Inteligencia de Negocios . En la figura 24 podemos ver algunas de las más populares que se describen a continuación.

¹⁵ Prudsys. (8 de Marzo de 2011). Prudsys. Recuperado el 14 de enero de 2014, de Prudsys: <http://www.prudsys.de/en/technology/xelopes/>



Figura 24 Suites de Inteligencia de Negocios

Fuente: obtenido de la Tesis Desarrollo de una Aplicación de Business Intelligence para la empresa Empaqplast. Escuela Politécnica del Ejército. Ecuador, 2012.

- **Pentaho BI**

Pentaho BI es la única suite que existe de código abierto, ya que ha adoptado las herramientas de Kettle, Mondrian y WEKA para conformar una sola plataforma de Inteligencia de Negocios. Tiene las funcionalidades de reportes, análisis, data mining e integración de datos (Pentaho, 2011). Los proyectos involucrados en Pentaho BI son:

- Mondrian: servidor de OLAP.
- JFreeReport: reporteador.
- Kettle: integración de datos (ETL).
- Pentaho: plataforma de inteligencia empresarial.
- WEKA: data mining.

Permite a los desarrolladores de Java diseñar componentes que pueden ser rápidamente ensamblados en soluciones de inteligencia empresarial y a los usuarios finales desplegar rápidamente las soluciones existentes de Inteligencia de Negocios .

A pesar de que Pentaho BI se conforma ya como una suite que satisface las necesidades de inteligencia empresarial, sus componentes no se integran perfectamente ya que surgieron de manera individual. Sin embargo la compañía Pentaho ofrece soporte comercial para todas ellas como conjunto, por eso se considera una suite.

- **Oracle**

Oracle BI es la plataforma más completa para Inteligencia de Negocios , cubre un amplio rango de necesidades, incluidos los tableros interactivos, el análisis ad-hoc, alertas e inteligencia proactivas, publicación e informes avanzados, análisis predictivo en tiempo real, y análisis de tecnología móvil (Oracle, 2012).¹⁶

Existen varias versiones de las plataformas:

¹⁶ Oracle. (13 de Enero de 2012). Oracle. Recuperado el 20 de enero de 2014, de Oracle: <http://www.oracle.com/us/solutions/business-analytics/business-intelligence/index.html?origref=http://www.oracle.com/us/solutions/ent-performance-bi/enterprise-edition-066546.html>

- Oracle BI Suite Enterprise Edition (EE): es una plataforma completa e innovadora que ofrece bases para crear soluciones empresariales desde fuentes de datos heterogéneas para la distribución de datos, con sistemas Oracle y que no son de Oracle.
- Oracle BI Standard Edition (SE): incluye Oracle Discoverer, está optimizado para trabajar con los datos y las aplicaciones Oracle, y ofrecer análisis e inteligencia avanzados al menor costo.
- Oracle BI Publisher: también denominado Oracle XML Publisher, ofrece la solución más eficiente y escalable para informes y publicaciones, disponible para entornos complejos y distribuidos.
- Disponible con Oracle BI Suite EE o como solución independiente. Oracle Real-Time Decision: combina los requerimientos comerciales y de información del cliente para hacer la mejor recomendación en cada interacción con el cliente y en cada decisión operacional al adaptarse de manera inteligente, la información en constante cambio.

La versión Enterprise cuenta con las herramientas Oracle Data Integrator con las capacidades de ETL para la creación de datawarehouses, Oracle OLAP para el análisis de datos multidimensional y Oracle Data mining para el análisis y aplicación de algoritmos de data mining. Actualmente estas 3 herramientas se distribuyen también por separado.

- **SAS**

SAS Inteligencia de Negocios integra los datos de la empresa y tiene un servicio de análisis y reportes, posee amplias capacidades de inteligencia empresarial y la aplicación de análisis e integración de datos para crear una solución completa y de fácil interacción SAS BI ofrece:

Reportes: wizards intuitivos, reportes ilimitados, reportes ad-hoc, integración de resultados y procesos analíticos.

- Consultas y análisis: consultas para todos los niveles de usuarios basados en sus necesidades y habilidades, y se pueden realizar sobre múltiples fuentes de datos utilizando datos relacionales o multidimensionales.
 - OLAP: interfaces fáciles de utilizar para construir y mantener los cubos de OLAP. Los cubos pueden ser creados con base en datos almacenados en cualquier base de datos.
 - Visualización: sumatorias y presentación de datos usando una gran variedad de gráficos.
 - Data mining con Enterprise Miner.
- **SQL Server 2012**

Constituye la alternativa de Microsoft a otros potentes gestores de bases de datos como son Oracle o MySQL. Generalmente no se lo asocia como una suite de BI, sin embargo, posee potentes herramientas de productos integrados que ofrecen cobertura en la toma de decisiones. Estos productos son: Integration Services, Analysis Services y Reporting Services, que junto al alcance de SQL Server se convierten en el motor de potentes aplicaciones de inteligencia de negocios para empresas, manteniendo los costos bajo control.

Entre sus principales características tenemos:

- Integra bases de datos en almacenes de datos (datawarehouses) y aumenta la capacidad de gestión de grandes tablas.
- Permite la construcción de soluciones Inteligencia de Negocios robustas y de alto rendimiento para el análisis de datos.

- Posee el módulo SQL Server Inteligencia de Negocios Development Studio para desarrollar, probar, implementar y mantener segura, escalable la integración de datos. Además de generar reportes y análisis optimizado de soluciones.
- SQL Server Analysis Services permite el diseño de cubos de datos y facilita el diseño de dimensiones.
- SQL Server Integration Services proporciona las características y el rendimiento necesario para construir aplicaciones de clase empresarial de integración de datos desde diversas fuentes de datos.
- SQL Server Reporting Services extiende la plataforma de almacenes de datos para llevar la información que el trabajador necesita y tener acceso a los datos del negocio.
- Genera directamente reportes a herramientas familiares como Microsoft Excel y Microsoft Word.
- Soporta código en transact-sql y ansi-sql.

2.3.7. CLASIFICACIÓN DE SUITES EN EL MERCADO

La figura 18 muestra una clasificación general de las suites comerciales de acuerdo a la opinión de los principales proveedores de software de Inteligencia de Negocios a nivel mundial.

La figura 25 muestra 4 cuadrantes: los líderes, retadores, visionarios y aplicaciones de nichos de oportunidad.

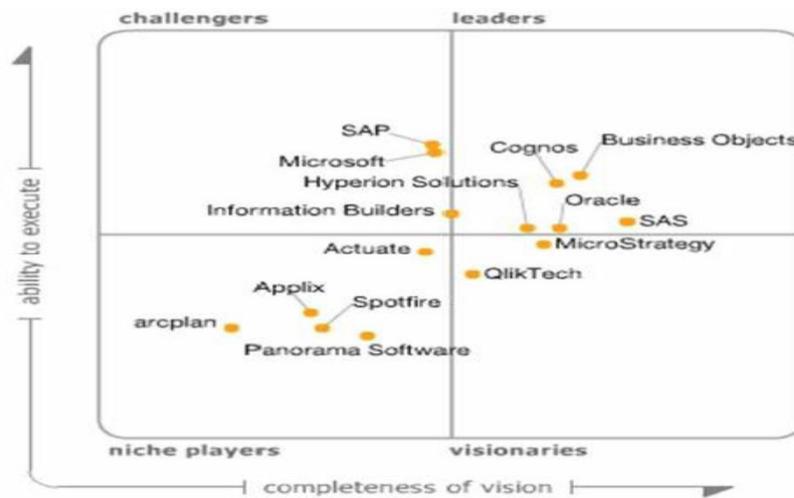


Figura 25 Clasificación de suites comerciales

Fuente: obtenido de la Tesis Desarrollo de una Aplicación de Business Intelligence para la empresa Empaqplast. Escuela Politécnica del Ejército. Ecuador, 2012.

2.3.8. METODOLOGÍA DE RALPH KIMBALL

Ralph Kimball es considerado uno de los representantes más importantes del Data Warehouse y Business Intelligence. Su metodología ha sido probada en muchos escenarios de negocio y se podría decir que se ha llegado a convertir en un estándar de proyectos BI.

En el año 1998 se publica la primera edición del libro “The Data Warehouse Lifecycle Toolkit” donde se expone dicha metodología.

En la siguiente figura 26 se muestra el ciclo de vida propuesta por Ralph Kimball y sus fases.

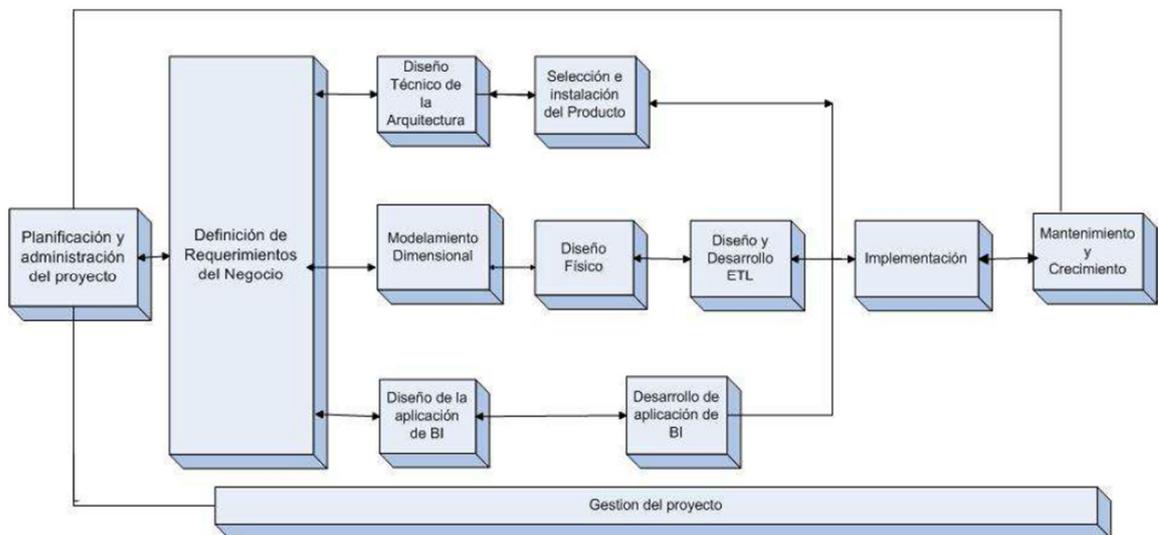


Figura 26 ciclo de vida de la metodología de Ralph Kimball

Fuente: Obtención de internet ¹⁷, Kimball Group, EE.UU

2.3.8.1. Planificación y Administración del Proyecto

La planificación busca identificar la definición y el alcance del proyecto, incluyendo las justificaciones del negocio y las evaluaciones de factibilidad.

La metodología en esta etapa propone identificar el alcance basado en requerimientos de negocios y no en fechas establecidas. Así el cumplimiento del proyecto será directamente relacionado con el negocio de la empresa.

- **Definición del proyecto**

Como paso inicial para el desarrollo de un proyecto de BI se debe identificar de donde proviene la necesidad de información la cual puede darse de un sector específico como los directivos o gerentes de la organización o a su vez puede provenir de varios sectores de la empresa.

¹⁷ <http://www.kimballgroup.com/data-warehouse-business-intelligence-resources/kimball-techniques/dw-bi-lifecycle-method/>,

- **Administración del proyecto**

En esta etapa se verifica mediante reuniones con los involucrados en el proyecto el avance del mismo y el cumplimiento de los requerimientos

- **2.3.8.2. Definición de los Requerimientos del Negocio**

El punto clave para el proceso de desarrollo un Data Warehouse es la manera como se interpretan y se analizan los requerimientos la interpretación correcta de los diferentes niveles de requerimientos proporcionara la visión de cómo se realizara el diseño del Data Warehouse.

Para iniciar con el proceso de obtención de requerimientos se debe hablar con los principales encargados del negocio, se debe enfocar el proceso de toma de requerimientos en cosas como: en base a qué información se toman decisiones, en qué consisten sus trabajos, que información es la que se maneja con más frecuencia, en que actividades invierte más tiempo.

Los diseñadores del Data Warehouse deben comprender cuales son los factores y las reglas que dirigen el negocio, cual es el funcionamiento y los procesos que se llevan a cabo para la actividad de negocio, determinar y analizar los requerimientos para traducirlos en consideraciones de diseño apropiadas.

- **Entrevistas**

Es de vital importancia dirigir las entrevistas a los directivos debido a que son los que tienen un mejor conocimiento de las reglas de negocio, las entrevistas también deben ser realizadas al departamento de tecnología para conocer con qué recursos se cuenta para el desarrollo del proyecto.

2.3.9. Diseño Técnico de la arquitectura

El diseño de un Data Warehouse requiere la integración de varias tecnologías por lo que se debe tener en cuenta elementos como: requerimientos del negocio, el entorno técnico y las estrategias de diseño.

2.3.8.3. Modelo dimensional

El Modelado Dimensional, según su creador Ralph Kimball, “es el diseño físico y lógico que transformará las antiguas fuentes de datos en las estructuras finales del Data Warehouse, a través de una técnica que busca la presentación de los datos en un marco de trabajo estándar que es intuitivo y permite un acceso de alto desempeño. Cada modelo dimensional está compuesto de una tabla que tiene una llave compuesta llamada tabla de hechos y un conjunto de tablas más pequeñas llamadas dimensiones. Cada tabla dimensión tiene una llave primaria simple, que corresponde exactamente a una de las partes de la llave compuesta en la tabla de hechos. Esta estructura característica es usualmente llamada esquema estrella”¹⁸

2.3.8.4. Técnicas de Modelado Dimensional

“La idea fundamental del modelado dimensional es que casi cada tipo de datos de negocio puede ser representado como un tipo de cubo de datos, donde cada celda del cubo contiene un valor medido y las aristas del cubo definen las dimensiones naturales de los datos”¹⁹

- Seleccionar el proceso de Negocio
- Declaración de la granularidad

¹⁸ Kimball98, p144.

¹⁹ Kimball98, p165

- Identificación de las dimensiones
- Identificación de los hechos

2.3.8.5. Diseño Físico

En lo que se refiere a la estructura física la etapa incluye tareas como la configuración de la base de datos debe incluir los nombres de columna físicos, los tipos de datos y constraints.

La creación de tablas para datos y metadatos, donde se alojarán dimensiones y tablas de hechos.

- Creación de secuencias para los procesos ETL.
- Creación de tablas temporales, tablas históricas, tablas para errores.
- Generación de Scripts y diagramas físicos de base de datos.

2.3.10. Diseño y Desarrollo de la Presentación de Datos

Esta etapa está conformada primeramente por las actividades de extracción, transformación y carga de datos (ETL). Se realiza un análisis de los datos como se van a integrar y como se pueden resolver problemas de inconsistencias.

2.3.13.1. Actividades para el proceso ETL

- Realizar un plan cada proceso de ETL.
- Seleccionar una herramienta ETL.
- Estrategias para proceso de carga.
- Carga inicial de Datos(Dimensiones y Hechos)
- Carga incremental de Datos(Dimensiones y Hechos)

- Automatización del proceso ETL

2.3.11. Especificación de Aplicaciones para Usuarios Finales

Cada usuario del Sistema de Inteligencia de Negocio necesita un rol o perfil de usuario para los diferentes tipos de aplicaciones, acceso a carpetas, acceso a cierto tipo de reportes, entre otras por lo que en esta etapa se definen los diferentes tipos de permisos para los usuarios.

2.3.12. Desarrollo de Aplicaciones para Usuarios Finales

Los usuarios acceden al Data Warehouse por medio del BI Server herramienta gráfica, que contiene la información de cada área de negocio, se pueden desplegar reportes, vistas de análisis, tableros de control. Los cubos, reportes y tableros de control son implementados conforme a los requerimientos.

2.3.13. Implementación

La implementación representa la unión de la herramienta de BI, los datos y las aplicaciones de usuarios finales. Existen factores extras que aseguran el correcto funcionamiento de todos los elementos, entre ellos se encuentran la capacitación, el soporte técnico y la comunicación. Todas estas tareas deben tenerse en cuenta antes de que cualquier usuario pueda tener acceso al sistema de BI.

2.3.14. Mantenimiento y crecimiento

Como en todo proyecto se necesita tener actualizaciones de forma constante para poder dar un ciclo de vida adecuado al producto. Es importante establecer las prioridades para poder manejar los nuevos requerimientos de los usuarios y de esa forma poder evolucionar y crecer.

2.3.15. Gestión del proyecto

La gestión del proyecto asegura que las actividades del ciclo de vida se lleven a cabo de manera sincronizada. La gestión del proyecto acompaña todo el ciclo de vida. Entre sus actividades principales se encuentra la monitorización del estado del proyecto y el acoplamiento entre los requerimientos del negocio y las restricciones de los sistemas de información para poder manejar correctamente las expectativas en ambos sentidos.

CAPITULO III

DESARROLLO DE LA METODOLOGIA

3.1 DESARROLLO DE BI PARA LA EMPRESA TRITON TRADING

Para el desarrollo del proyecto se utilizará la metodología de Ralph Kimball lo que permitirá dar un seguimiento adecuado al desarrollo del proyecto.

3.1 PLANEACIÓN Y ADMINISTRACIÓN DEL PROYECTO

En este punto se determinara el propósito del proyecto de inteligencia de negocios, sus objetivos y los alcances del mismo, los principales riesgos y una aproximación inicial a las necesidades de información

- **Definición del Proyecto**

Partiendo de la necesidad de la empresa Triton Trading S.A en el desarrollo de un sistema BI para dar soporte al análisis de la información y la toma de decisiones por los gerentes de las áreas de ventas y compras, sea decidido implementar este proyecto.

- **Preparación para un proyecto de Data Warehouse.**

La empresa Triton Trading es muy consciente de la necesidad de contar con un herramienta como la Inteligencia de Negocios , por lo que se tiene el apoyo de la gerencia de ventas y compras en el desarrollo, el departamento de tecnología de la información (TI), es la encargada de presentar este proyecto como soporte para la toma de decisión para los gerentes de las áreas ya mencionadas

- **Alcance**

El proyecto estará centrado en el desarrollo de una aplicación de Inteligencia de Negocios (BI) para el soporte en la toma de decisiones en la empresa Triton Trading S.A.. El aplicativo abarcará las siguientes áreas:

- Ventas.
- Inventarios.
- Compras.

Para cada una de estas áreas se desarrollará, los cubos de información, siendo los gerentes de cada departamento los responsables del análisis de los cubos. Se desarrollará los Data Mart de las áreas de ventas, inventarios y compras que conformarán el Data Warehouse. Se generarán reportes que se adaptarán a las necesidades de información concerniente a las áreas ya mencionadas.

- **Justificación**

La solución que se plantea para el análisis de la información está basada en la elaboración de una aplicación de Inteligencia de Negocio (BI) que estará conformado por los Data Mart del área ya mencionada.

Se busca relacionar los datos con el negocio para obtener información relevante sobre la situación de la empresa. Tomando en cuenta que la

empresa trabaja con un sistema ERP llamado SAP Business One y un ERP de Gestión de activos Máximo. Se considerará el uso de la herramienta Suite de SQL Server 2008 esta herramienta facilitará el camino para conseguir una completa solución de Inteligencia de Negocio (BI) y una rápida integración con la infraestructura que existe actualmente en la empresa. Aplicando este sistema de Inteligencia de Negocio (BI), se pretende reducir los costos y optimizar los tiempos en lo que se refiere al tratamiento de la información se facilitará la administración de la misma personalizándola y adaptándola a las necesidades lo que permitirá la Integración y depuración de los datos.

- **Descripción de los Stakeholders**

A continuación en la tabla 6 se detalla los Stakeholder que intervendrán en el proyecto del desarrollo del proyecto:

Tabla 6 Descripción de stakeholders

Stakeholders	Cargo	Función que desempeña en el proyecto
Renzo Gamboa	Gerente Comercial	Brindar Información Necesaria sobre la gestión de ventas y compra de maquinarias y repuestos
Jose Miguel Olavide	Gerente de Compras	Brindar Información Necesaria sobre la gestión de compra de maquinarias y repuestos
Victoria Gomez	Coordinadora de Venta	Brindar información de ventas de repuestos
Marlene Zavala	Coordinadora de Compras	Brindar información de compras de repuestos
Neil Alarcon	Analista desarrollador	Análisis de requerimiento y carga de datos al Data Mart
Luis Ticacala	Analista de TI	Soporte tecnología para el desarrollo del negocio
Luis de la Cruz	Analista BI	Brinda información de datos de gestión para la toma de decisión

Fuente: Elaboración propia

- **Requisitos del Proyecto**

- **Humanos**

- **TI**

Tabla 7 recurso de TI

Stakeholders	Cargo	Función que desempeña en el proyecto
Luis Ticacala	Analista de TI	Soporte tecnología para el desarrollo del negocio
Luis de la Cruz	Analista BI	Brinda información de datos de gestión para la toma de decisión

Fuente: Elaboración propia

- **Gestión**

Tabla 8 recurso de gestión

Stakeholders	Cargo	Función que desempeña en el proyecto
Renzo Gamboa	Gerente Comercial	Brindar Información Necesaria sobre la gestión de ventas y compra de maquinarias y repuestos
Jose Miguel Olavide	Gerente de compras	Brindar Información Necesaria sobre la gestión de compra de maquinarias y repuestos
Victoria Gomez	Coordinadora de Venta	Brindar información de ventas de repuestos
Marlene Zavala	Coordinadora de compra	Brindar información de compra de repuestos

Fuente: Elaboración propia

- **Requerimiento Tecnológicos**

- **Hardware**

Tabla 9 recurso de hardware

Equipo	Uso	Cantidad
Laptop	Análisis y desarrollo del proyecto	2

Fuente: Elaboración propia

- **Software**

Tabla 10 recurso de hardware

Software	Uso	Cantidad
Suite de SQL server 2008	Analysis services Integration services Herramienta desarrollo operacional Herramienta de desarrollo dimensional	2

Fuente: Elaboración propia

Riesgo del Proyecto

El riesgo que presenta el proyecto es el tiempo para el desarrollo de la misma otro riesgo para nuestro proyecto es usuario final tiene que participar desde el inicio del proyecto ya que en diferentes roles va a ser el beneficiario de la solución y conceptualmente va a tener que trabajar para acostumbrarse a las formas de estas soluciones.

- **Definición del requerimiento del negocio**

De acuerdo a las reuniones realizadas con el Gerente de Comercial y La coordinadoras de ventas y compras que son los dueños del proceso de ventas en la empresa Tritón Trading S.A. se establecieron requerimiento para el control de ventas y las asignación de cuotas para distintas sucursales y vendedores, ya que se necesita el control oportuno para la toma de decisión adecuada que conlleva una oportunidad de negocio para la empresa. La cuales se describe en un matriz

Tabla 11 Requerimiento de negocio

Proceso de Negocios	Proveedor	compras	Clientes	Productos	Tipo producto	Vendedor	Sucursal	alancen
Gestión de Venta		x	x	x	x	x	x	x
Control de cuota		x		x	x	x	x	x
Control de clientes TOP		x	x	x	x			
Ventas por productos				x	x			
Rotación de inventario				x	x			x
Gestión de compras	x	x		x	x		x	x
Tiempo de respuesta proveedor	x	x		x				
Cantidad solicitada VS cantidad entregada	x	x						
Compra por categoría de proveedor	x	x						
Respuestas de proveedores estratégicos	x	x		x	x			x
Valorización de Stock				x	x			x
Movimiento de Inventario				x	x		x	x

Fuente: Elaboración propia

3.3 DISEÑO Y ARQUITECTURA

Así como es importante tener un plano para un construcción de un puente o de un casa, para saber cómo se está estructurando o que elementos serán

incorporados en proyecto, para que en el futuro estos sean modificados más fácilmente. Es por ello que es también importante tener un diseño de nuestra estructura del proyecto BI ya que es importante tener desarrollo un plano cómo guía para que más adelante hacer remodelaciones o hacer modificaciones. Esta estructura nos permite obtener información necesario de diversas fuentes ya se de una base de datos de distintos o de un archivo Excel en nuestro caso obtendremos la información de la base de datos SQL de la empresa Triton Trading el cual gracias a Integración SSIS podemos realizar el llenado filtrado y conversión de datos a nuestras tablas de destino o nuestras tablas de dimensión. Una bes procesado y llenado las tablas de dimisión y nuestra tablas de hechos con SSIS procederemos a la construcción de nuestro cubos para ello usaremos la herramienta de SQL server SSAS, una vez hecho nuestros cubos procederemos a la presentación de nuestra información a los usuarios finales lo cual lo realizaremos con la herramienta de análisis de inteligencia de negocios de la aplicación Microsoft Excel.

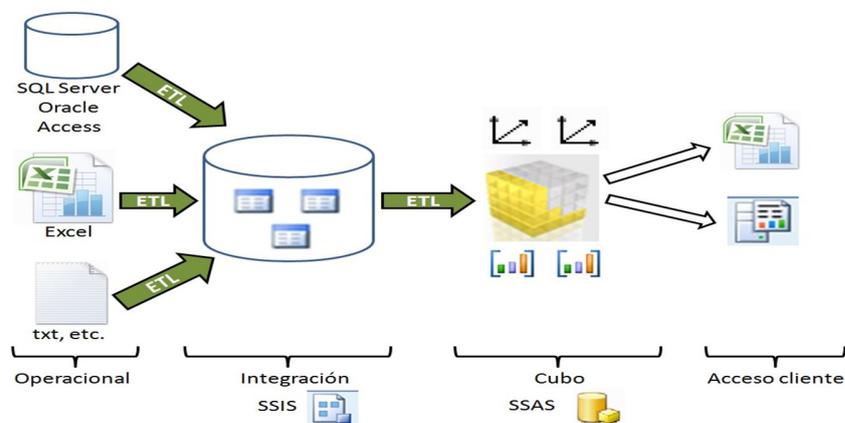


Figura 27 Estructura de proyecto

Fuente: obtenido de la Tesis Desarrollo de una Aplicación de Business Intelligence para la empresa Empaqplast. Escuela Politécnica del Ejército. Ecuador, 2012.

3.4 INSTALACION Y SELECCIÓN DEL PRODUCTO:

En esta etapa se instala y configura el Gestor de Base de Datos y la Herramienta de Bussines Intelligence que se va a utilizar para la implementación del Data Mart.



Figura 28 de instalación de sql server 2008

Fuente: MS SQL Server 2012



Figura 29 Carga de sql server 2008

Fuente: MS SQL Server 2012

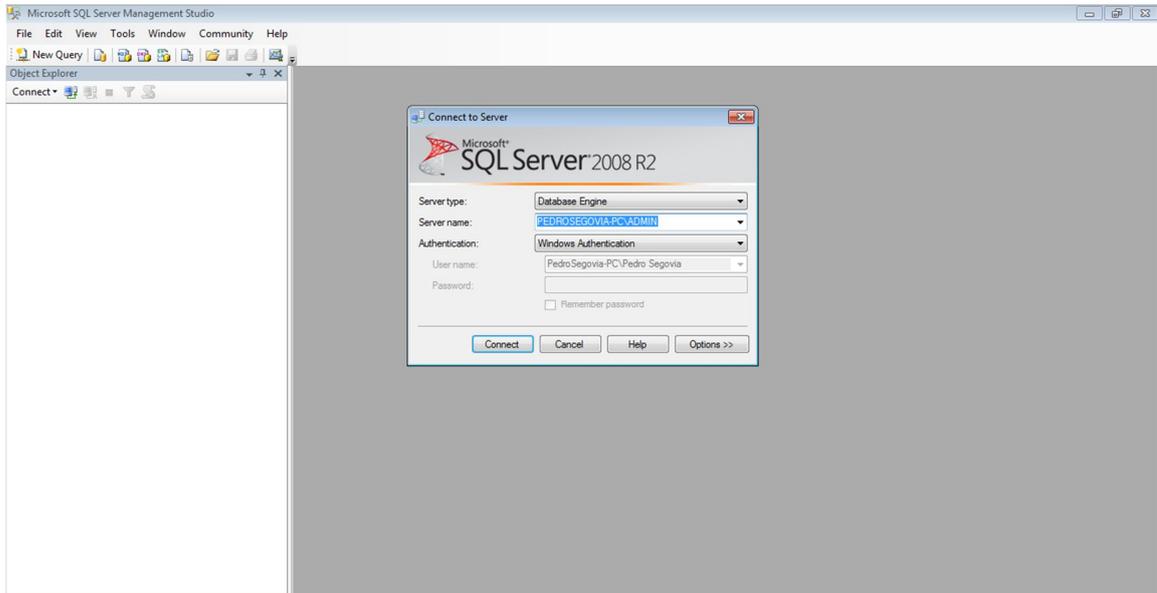


Figura 30 Conexión a la base de datos

Fuente: MS SQL Server 2012

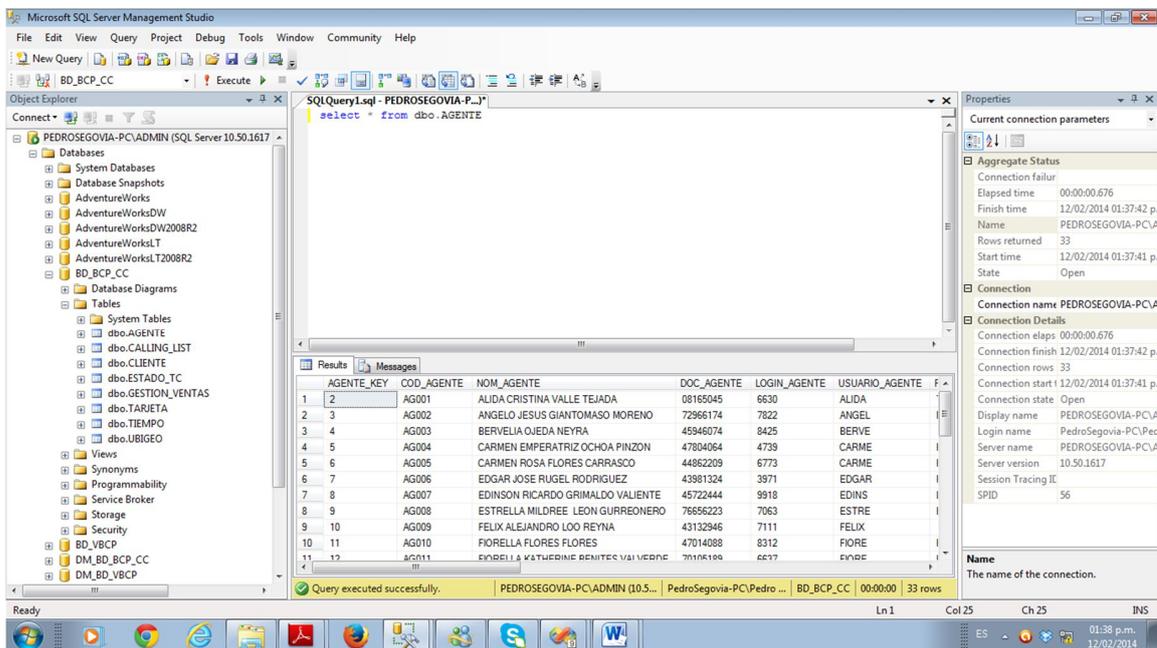


Figura 31 Área de trabajo de sql server 2008

Fuente: MS SQL Server 2012

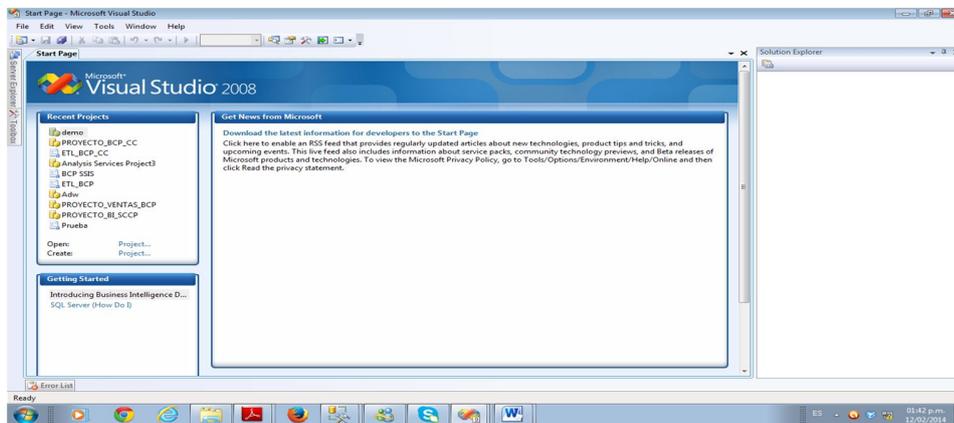


Figura 32 Herramienta de trabajo de inteligencia de negocio

Fuente: Elaboración Propia

3.5 MODELO DIMENSIONAL

El modelo dimensional utilizado para nuestro proyecto es un modelo estrella con copo de nieve, aquí se presenta y se organiza los datos definiéndolos en dimensiones. De esta manera se permite analizar la información.

Dentro de cada dimensión se puede definir los niveles de agregación para cada análisis a estos niveles se les conoce como atributos.

- **Modelos transaccional de empresa Triton Trading**

La toma de información de información para nuestro proyecto se realizó la extracción de datos para ello utilizamos las siguientes tablas

- Tablas entrada a almacén (OPDN Y PDN1)
- Tabla factura (OINV Y INV1)
- Tabla productos y grupos de producto (OITM y OITB)

- Tabla de almacén y productos por almacén (OITW Y OWHS)
- Tabla de vendedores (OSLP)
- Tablas cliente y proveedores (OCRD Y CRD1)
- Tabla de marca (@marca)
- Tablas orden de compra (OPOR Y POR1)

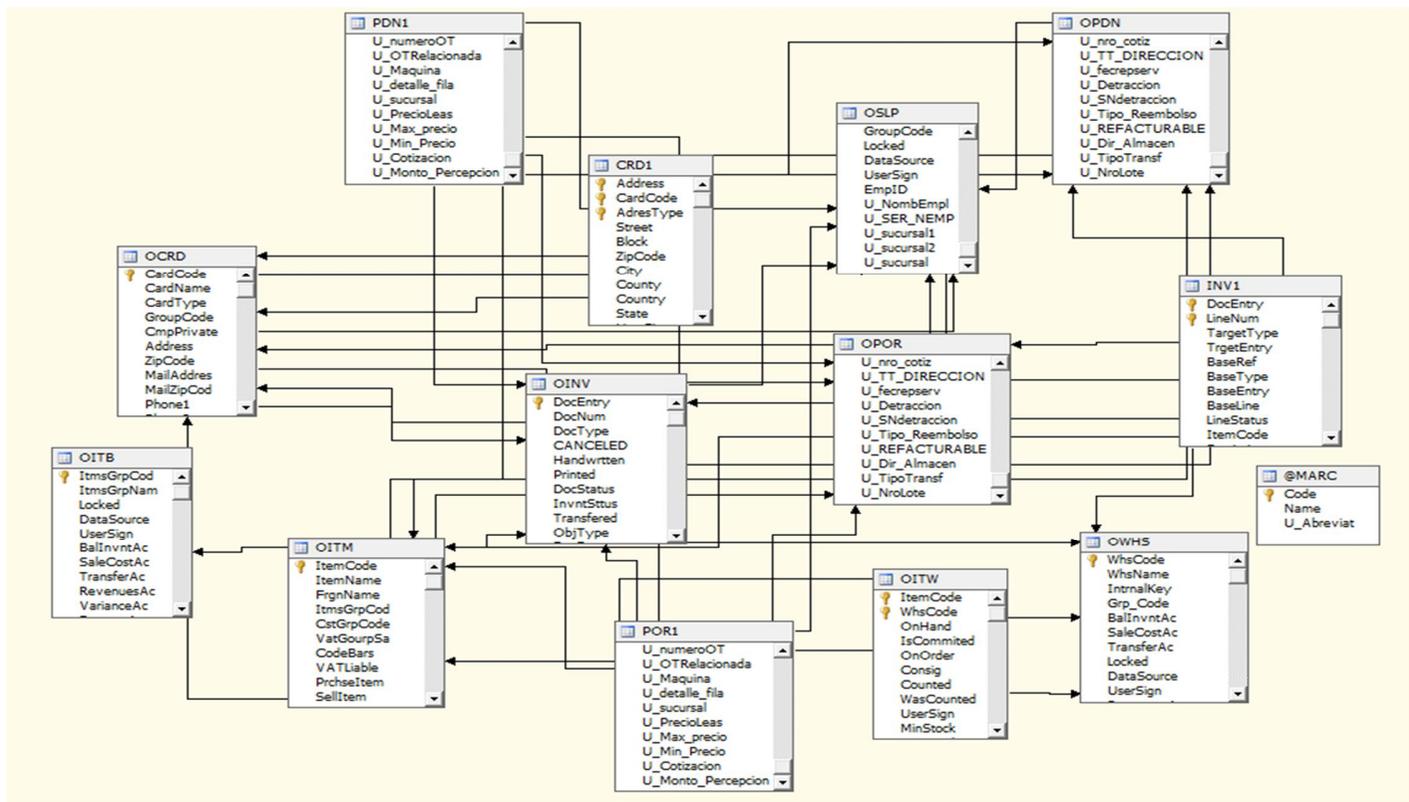


Figura 33 Modelo transaccional de la base de datos de Trion Trading

Fuente: Elaboración Propia

Diagrama Dimensional

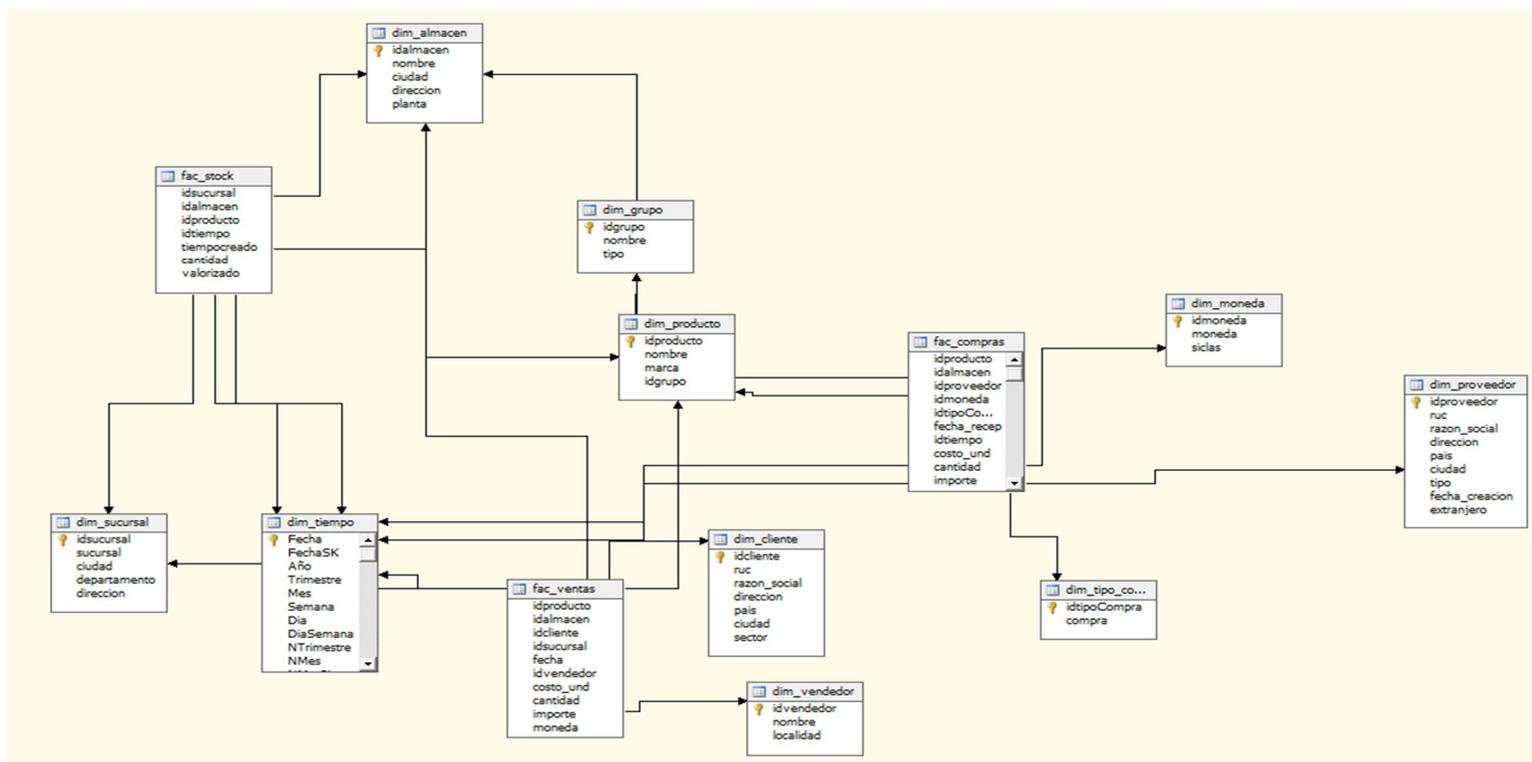


Figura 34 Diagrama Dimensional

Fuente: Elaboración Propia

- **Estructura de datos de las tablas**

A continuación se muestra las tablas con sus respectivos campos y tipos de datos que serán almacenados

- **Dimensión Producto**

Tabla 12 tablas dimensionales Dim Producto

Tabla	DIM_PRODUCTO							
Tipo de tabla	Dimensión productos							
Descripción	Detalle de productos							
Esquema	DW							
	Destino					Origen		
Columna	Tipo dato	tamaño	clave	Null	Tabla	campo	Tipo dato	tamaño
idProducto	Varchar	20	PK	N	OITM	Itemcode	varcahar	32
nombre	varchar	100		Y		itemname	Varchar	255
marca	Varchar	100		Y		U_codmarca	Varchar	100
idgrupo	numeric			Y		itemGrpcod	smallint	

Fuente: Elaboración Propia

- **Dimensión Tiempo**

Tablas 13 dimensionales Tiempo

Dimensión - Atributo	Tipo de Dato	
DIM_TIEMPO	Fecha	date
	Mes	smallint
	Trimestre	smallint
	dia	smallint
	Año	smallint
	diaSemana	smallint
	NTrimenstre	Char(7)
	Nmes	Char(15)
	Nsemana	Char(10)
	Ndia	Char(6)

Fuente: Elaboración Propia

- **Dimensión Tipo De Compra**

Tablas 14 dimensionales Dim Compra

Tabla	Dim_tio_compra								
Tipo de tabla	Dimensión tipo de compra								
Descripción	Tella de tipo de moneda								
Esquema	DW								
Destino					Origen				
Columna	Tipo dato	tamaño	clave	Null	Tabla	campo	Tipo dato	tamaño	
idtipoCompra	varchar	15			OPOR	U_TipoComp	varchar	32	
nombre	varchar	100			OPOR	U_TipoCompname	varcha	100	

Fuente: Elaboración Propia

- **Dimensión Moneda**

Tablas 15 dimensionales Dim Moneda

Tabla	Dim_moneda								
Tipo de tabla	Dimensión moneda								
Descripción	Destalle de moneda								
Esquema	DW								
Destino					Origen				
Columna	Tipo dato	tamaño	clave	Null	Tabla	campo	Tipo dato	tamaño	
idmoneda	char	3	PK	N	OPOR	currency	varchar	15	
moneda	varchar	30		Y					
siglas	Varchar	10		Y					

Fuente: Elaboración Propia

- **Dimensión Vendedor**

Tablas 16 dimensionales Vendedor

Tabla	Dim_vendedor							
Tipo de tabla	Dimensión vendedores							
Descripción	Detalle de vendedores							
Esquema	DW							
Destino					Origen			
Columna	Tipo dato	tamaño	clave	Null	Tabla	campo	Tipo dato	tamaño
idvendedor	numeric		pk	N	OSLP	slpcode	smallint	
nombre	varchar	100		Y	OSLP	slpname	Nvarchar	32
Localidad	numeric			Y	OSLP	U_sucursal	Nvarchar	max

Fuente: Elaboración Propia

- **Dimensión Proveedor**

Tablas 17 dimensionales Dim Proveedor

Tabla	Dim_proveedor							
Tipo de tabla	Dimensión proveedores							
Descripción	Detalle de proveedores							
Esquema	DW							
Destino					Origen			
Columna	Tipo dato	tamaño	clave	Null	Tabla	campo	Tipo dato	tamaño
Idproveedor	Varchar	15	PK	N	OCRD	cardcode	Varchar	32
ruc	varchar (32)	32		Y	OCRD	lictracnum	varchar	32
Razón_social	varchar (255)	255		Y	OCRD	cardname	varchar	255
direccion	Varchar(255)	255		Y	CDR1	Addrees1	varchar	255
pais	varchar(15)	15		Y	CDR1	country	varchar	50
ciudad	Varchar(50)	50		Y	CDR1	county	varchar	50
Tipo	Varchar(max)	Max		Y	CDR1	U_tipo	varchar	max
extranjero	Varchar	15			OCRD	U_vs_snpro	varchar	15
Fecha creación	datetime			Y	OCRD	createdate	datetime	

Fuente: Elaboración Propia

- **Dimensión Cliente**

Tablas 18 dimensionales Cliente

Tabla	DIM_CLIENTE							
Tipo de tabla	Dimensión Clientes							
Descripción	TABLA QUE CONTINEN DETALLES DE CLIENTES							
Esquema	DW							
Destino					Origen			
Columna	Tipo de dato	tamaño	clave	Null	Tabla	campo	Tipo dato	tamaño
IdCliente	Varchar	15	PK	N	OCRD	cardcode	Varchar	32
ruc	varchar	32		Y	OCRD	lictracnum	varchar	32
Razón_social	varchar	255		Y	OCDR	cardname	varchar	255
direccion	Varchar	255		Y	CRD1	Addrees1	varchar	255
pais	varchar	15		Y	CRD1	country	varchar	50
ciudad	Varchar	50		Y	CRD1	county	varchar	50

Fuente: Elaboración Propia

- **Dimensión Sucursal**

Tablas 19 dimensionales Sucursal

Tabla	Dim_sucursal							
Tipo de tabla	Dimensión sucursal							
Descripción	Detalle de sucursal							
Esquema	DW							
Destino					Origen			
Columna	Tipo dato	tamaño	clave	Null	Tabla	campo	Tipo dato	tamaño
idsucursal	numeric		pk	N	OLCT	code	smallint	
sucursal	varchar	50		Y	OLCT	name	Nvarchar	32
ciudad	numeric	50		Y	OLCT	U_abreviat	Nvarchar	max

Fuente: Elaboración Propia

- **Dimensión Almacén**

Tablas 20 dimensionales Dim Almacén

Tabla	Dim_almacen							
Tipo de tabla	Dimensión almacén							
Descripción	Detalle de almacén							
Esquema	DW							
	Destino				Origen			
Columna	Tipo dato	tamaño	clave	Null	Tabla	campo	Tipo dato	tamaño
idsucursal	numeric		PK	N	OWHS	whscode	Nvarchar	15
sucursal	varchar	50		Y	OWHS	whsname	Varchar	50
ciudad	Varchar	50		Y	OWHS	city	nvarchar	100
Dirección	Varchar	50		Y	OWHS	street	nvarchar	100

Fuente: Elaboración Propia

- **Dimensión Grupo**

Tablas 21 dimensionales Dim Grupo

Tabla	Dim_grupo							
Tipo de tabla	Dimensión grupos							
Descripción	Detalle de grupo							
Esquema	DW							
	Destino				Origen			
Columna	Tipo dato	tamaño	clave	Null	Tabla	campo	Tipo dato	tamaño
Idgrupo	numeric		PK	N	OITB	itmsGrpCod	Smallint	
Nombre	Varchar	100		Y	OITB	ItmsGrpNam	Varchar	100
tipo	varchar	50		Y	OITB	U_tipo	nvarchar	Max

Fuente: Elaboración Propia

- **Tabla de Hechos de Ventas**

Tablas 22 Tabla de hechos Ventas

Tabla de hechos - Ventas	Tipo de Dato	
FAC_VENTA	Idproducto	FK
	Idalmacen	FK
	Idcliente	FK
	Idsucursal	FK
	Fecha	FK
	Idvendedor	FK
	Costo_und	Numeric
	cantidad	Numeric
	importe	Numeric
	moneda	Vachar(3)

Fuente: Elaboración Propia

- **Tabla de Hechos de Compras**

Tablas 23 Tabla de hechos Compras

Tabla de hechos - compras	Tipo de Dato	
FAC_COMPRAS	Idproducto	FK
	Idalmacen	FK
	Idproveedor	FK
	IdtipoCompra	FK
	idmoneda	FK
	idtiempo	FK
	Fecha_recep	FK
	Costo_und	Numeric
	cantidad	Numeric
	importe	Numeric
	Can_recep	numeric

Fuente: Elaboración Propia

- **Tabla de Hechos de Stock**

Tablas 24 Tabla de hechos Stock

Tabla de hechos - Stock	Tipo de Dato	
FAC_STOCK	Idproducto	FK
	Idalmacen	FK
	Idsucursal	FK
	Idtiempo	FK
	Tiempo_creado	FK
	cantidad	Numeric
	valorizado	Numeric

Fuente: Elaboración Propia

3.6 DISEÑO FÍSICO

En este apartado nos centraremos en la arquitectura de base de dato, la selección de las estructuras de almacenamiento necesario para soportar el diseño lógico, así como los métodos que garanticen un acceso adecuado a la base de datos.

Poblando las dimensiones

En este punto procedemos a la llenada de datos de las dimensiones, como podemos observar en la figura 35 tenemos un flujo que empieza con la limpieza de las tablas DIM y FAC para luego llenar cada tabla de dimensión terminando con las tabla FAC al guanos llenados se pueden hacer por medio de tares ya que no tiene muchos datos para el llenados

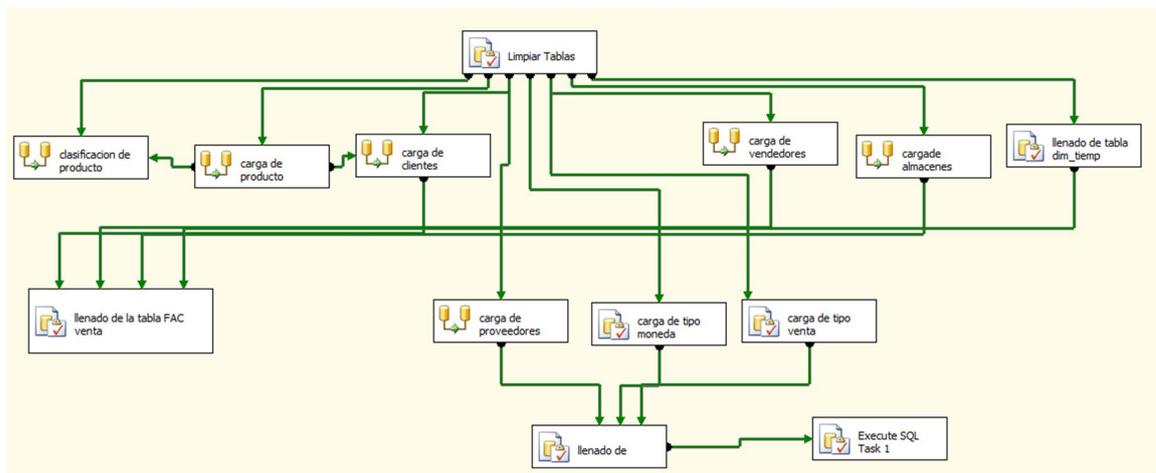


Figura 35 diagrama de flujo

Fuente: Elaboración Propia

Para el llenado de las dimensiones Lo primero que realizamos es establecer una conexión con la base de datos de origen para ellos usamos el siguiente objetos  que nos ayuda discriminar los campos que necesitamos para nuestra tabla DIM, la segunda tarea es convertir la data origen a los datos convenientes de la tabla DIM para ello usamos el objeto  una vez

terminada se realiza la copia de toda la información en una tabla temporal  para luego pasar toda la información a la tabla DIM para ellos usamos el objeto  que nos ayuda con la copia de la información a continuación procederemos a mostrar las figuras de llenado de dimensiones

En figura 36 mostramos el llenado de la dimensiones de tipo clientes.

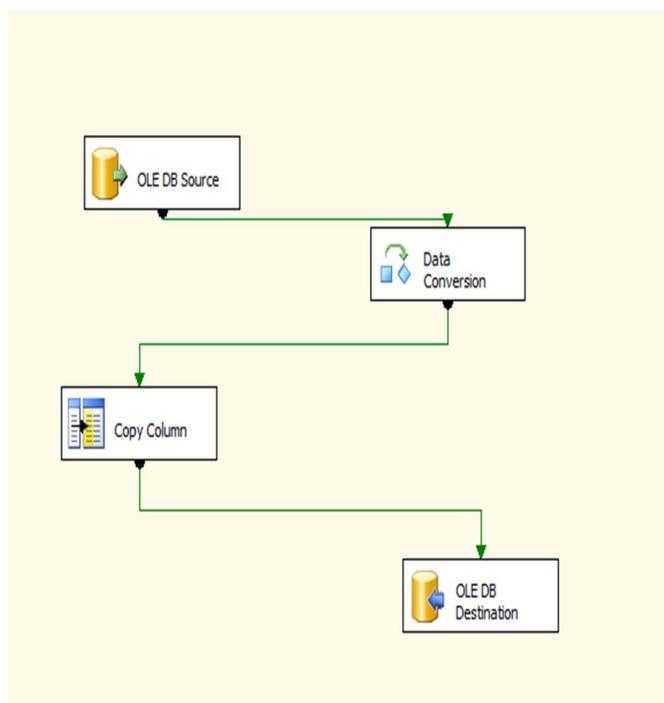


Figura 36 llenado dimensión cliente

Fuente: Elaboración Propia

En figura 37 mostramos el llenado de la dimensiones de tipo vendedores

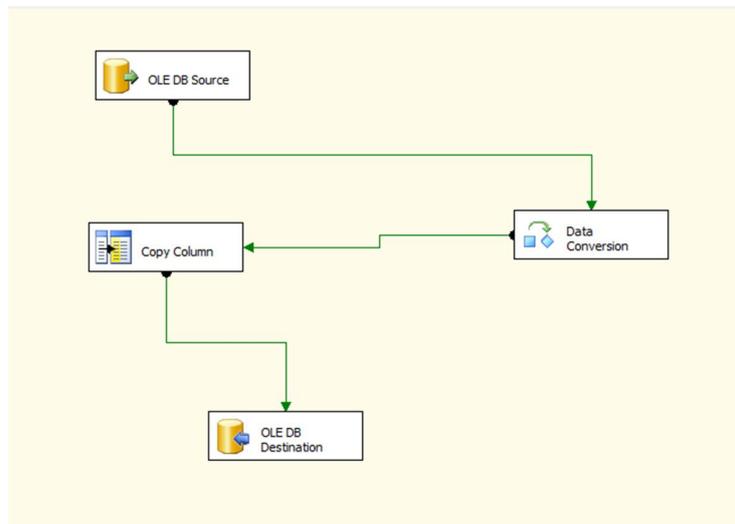


Figura 37 llenado de dimensión vendedores

Fuente: Elaboración Propia

En figura 38 mostramos el llenado de la dimensiones de tipo productos

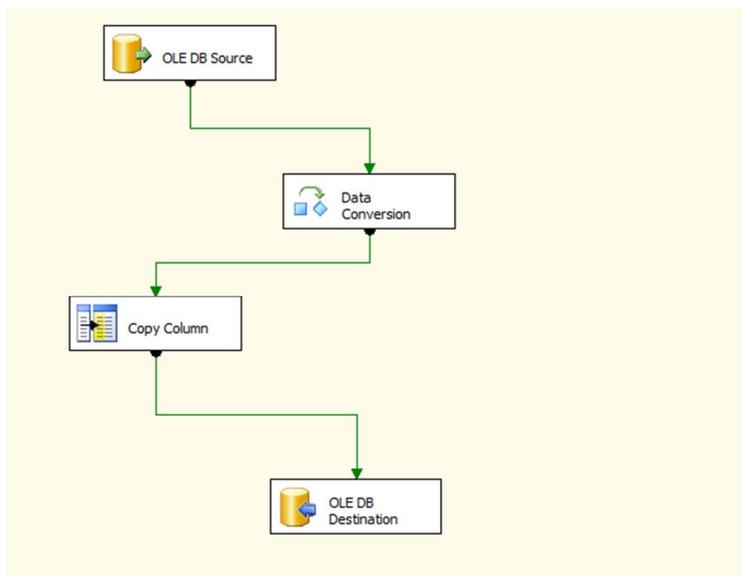


Figura 38 llenado de dimensión productos

Fuente: Elaboración Propia

En figura 39 mostramos el llenado de la dimensiones de tipo grupos

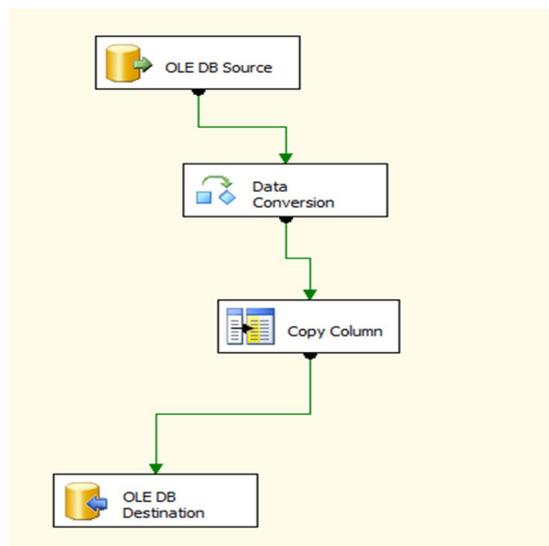


Figura 39 llenado de dimensión grupos
Fuente: Elaboración Propia

En figura 40 mostramos el llenado de la dimensiones de tipo almacén

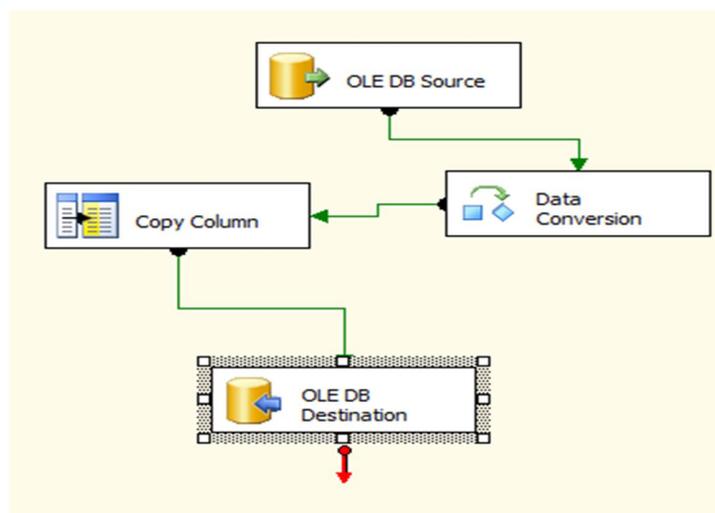


Figura 40 llenado de dimensión almacén
Fuente: Elaboración Propia

En figura 41 mostramos el llenado de la dimensiones de tipo proveedores

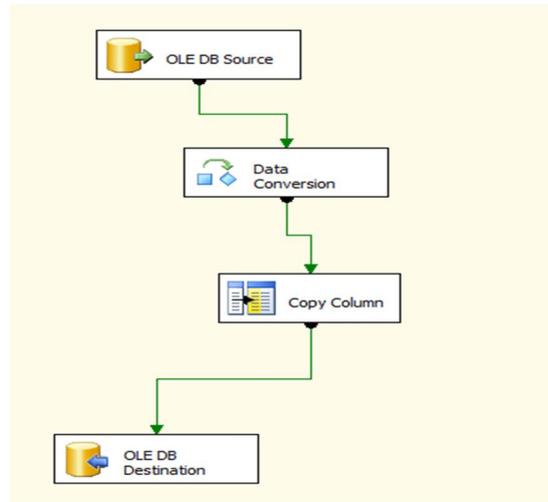


Figura 41 Llenado de dimensión Proveedores
Fuente: Elaboración Propia

Para el llenado de las dimensiones tipo de compra, tiempo y moneda se usaron tareas ya que esta son códigos sql, el uso de esta opción se debe a la poca complejidad extracciones datos

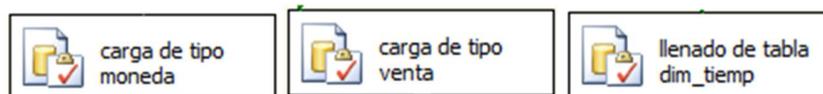


Figura 42 Tareas
Fuente: Elaboración Propia

Llenado de la tabla Fac Ventas se hará por un tarea que se ejecutar por el flujo de y esta se por un procedimiento almacenados

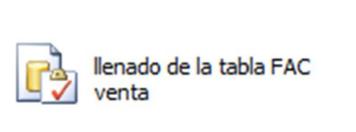


Figura 43 Tareas
Fuente: Elaboración Propia

DESARROLLO DE CUBOS

CUBO DE VENTAS

En esta sección observaremos nuestra creación de cubos de ventas. Este cubo nos presentará la información de ventas necesarios para la gestión de ventas y el análisis de la misma con ello ayudará a la toma de decisión del área correspondiente, para la construcción de nuestro cubo se utilizó las siguientes tablas de dimensión

- dim_producto
- dim_grupo
- Dim_cliente
- Dim_sucursal
- Dim_almacen
- Dim_tiempo

A continuación mostraremos la estructura

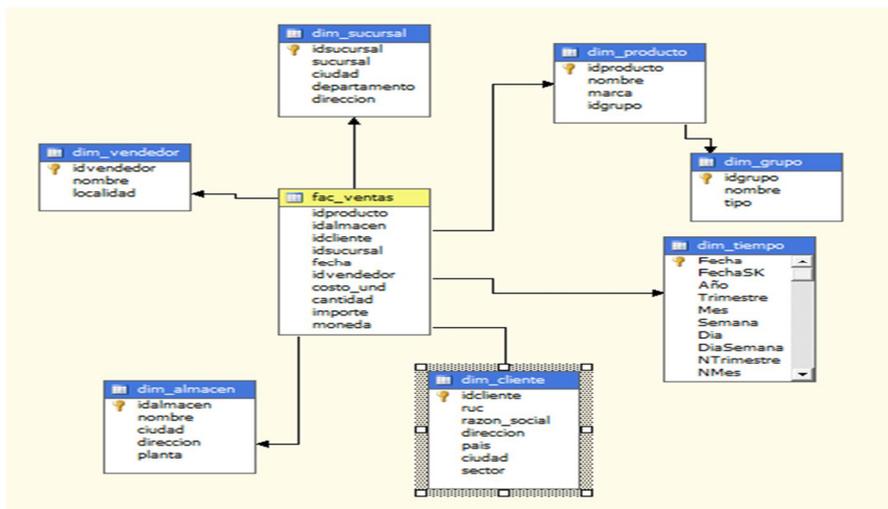


Figura 44 estructura de cubo Venta

Fuente: Elaboración Propia

Resultados del Cubo de Ventas

grupo nombre	Costo Unit	Cantidad	Importe	Fac. Ventas Count
EQUI BATERIAS	169318.793212	33	181930.08	31
EQUI CARGADORES	261955.24	83	261955.24	83
EQUI ALBOKITA	85000.06	20	140706.88	18
EQUI OTROS	18279.09	1	18279.09	1
RPTOS ALLE TECH	12856.077948	1689	43105.67	67
RPTOS BOONMA	110091.33094	487	181875.54	121
RPTOS CASCADE	58724.83	21	68876.95	19
RPTOS CHEM TOOLS	14978.461552	1754	28273.16	1567
RPTOS CHICAGO PHELMIA	124548.403385	1229	231019.99	426
RPTOS CONTRA	1158.278299	538.71	5266.18	94
RPTOS CREOVIN	1935138.43592602	40664	2638932.380000001	17530
RPTOS CLIMAXIS	393669.870240001	5109	532426.860000001	1882
RPTOS DANA	43191.172503	747	85097.9	102
RPTOS FERRITARIO	248263.816000001	20835	495212.81	16602
RPTOS GENIE	29712.26	283	32832.58	206
RPTOS HALLOTTE	11664.67	62	14220.23	46
RPTOS HENKONSA	1128.47	6	1238.17	6
RPTOS HUCK	112496.929479	16993	143002.51	65
RPTOS ILKALAI	53189.63307698	17521	791449.579999999	8147
RPTOS KALMAR	3410560.904145	23575	5716611.59000001	6981
RPTOS KIBOTA	203300.193550001	18803	427446.14	2948
RPTOS LAM TRUCK PART	179799.862852	1487	175487.24	334
RPTOS LIBROSR	471363.488518	4267	717501.14	1112
RPTOS LIANTAS	2258.27	20	5107.62	9
RPTOS LUBRICANTES	408718.171965998	31825.14	783044.68	13148
RPTOS MARIQUETRAHELE	1411.884608	4	2832.77	2
RPTOS NARROW	2384.87	13	2996.55	6
RPTOS NEMIA	128435.49164	1243	461028.499999999	638
RPTOS OTROS	488835.530921001	25790.3	731889.71	5821
RPTOS TP	7138.63	25	0	0
Grand Total	9688923.46271318	213037.19	14526332.94000001	77805

Figura 45 resultado de cubo
Fuente: Elaboración Propia

CUBO DE COMPRAS

Al igual que el cubo de ventas el cubo de compra presenta la información necesaria para la gestión de compras y análisis de la misma, esta herramienta ayudara al o las ares respondientes para una mejor toma de decisión. Nuestro cubo está conformado por las siguientes tablas de dimensiones

- dim_producto
- dim_grupo
- Dim_proveedor
- Dim_tipo_compra
- Dim_almacen
- Dim_tiempo
- Dim_moneda

Resultado de cubo de compras

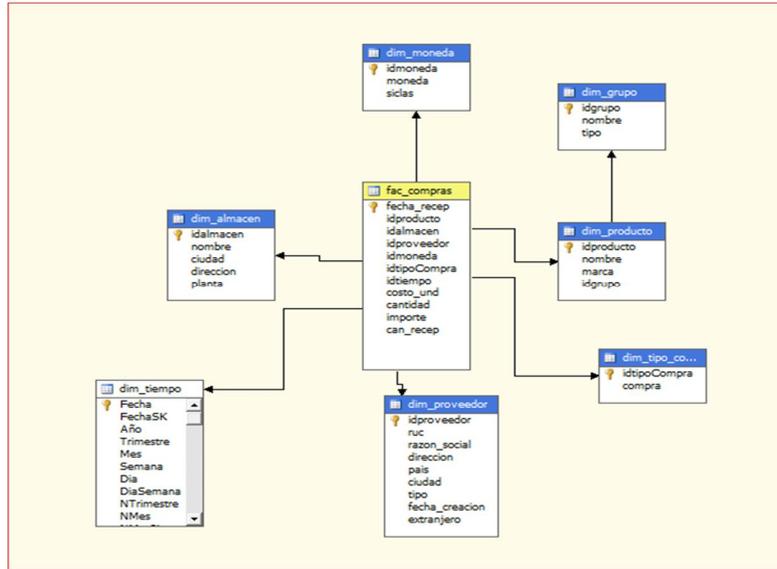


Figura 46 estructura de cubo de compra
Fuente: Elaboración Propia

Resultado de cubos

La imagen muestra una interfaz de usuario de un software de análisis de datos. En la parte superior hay una barra de herramientas con opciones como 'Cube Structure', 'Dimension Usage', 'Calculations', etc. El panel principal muestra un resumen de un cubo de compra con los siguientes campos:

grupo nombre	Costo Und	Cantidad	Importe	Can Recep	Fac Compras Count
EQU BATERIAS	28379.38	277	651067.24	385	75
EQU CARGADORES	142820.22	278	285073.68	417	87
EQU CROWN	1691882.36	520	3253133.16	435	93
EQU CUMMINS	24600	1	24600	1	1
EQU HYUNDAI	679662	61	1160794	120	37
EQU KUBOTA	35164.46	39	105082.13	21	12
EQU LITREBERR	1994948	14	2572844	11	10
EQU OTROS	3013819.66127	202	3534650.65	231	107
EQU ROSSINABLER	24727707	68	49455414	189	34
EQU SANY	3244617.1045	39	4320758.88	39	35
EQU WIANCO	98175	144	946880	86	15
EQU XCMG	75500	1	75500	1	1
RPTOS AXLE TECH	527.78	48	6333.36	10	4
RPTOS BROMMA	112.3	20	1123	11	2
RPTOS CHEM TOOLS	2919.82	46554	784857.8	21182	206
RPTOS CHICAGO PNEUMA	8837.2538	128	10278.74	501	49
RPTOS CONITX	10589.29076	7335.09	48300.69	7145.41	1071
RPTOS CROWN	68972.791743	31599	1067607.83	31538	4855
RPTOS CUMMINS	7043246.576908	274281	13224518.98	263122	82463
RPTOS DANA	4013.28	31	4833.28	31	5
RPTOS FLETIGUARD	327915.291415	85699	2513466.13	85203	10402
RPTOS GENIE	15907.428236	725	42893.48	906	300
RPTOS GETS	273.95	392	3631.2	392	25
RPTOS HALLOTTE	2939.706898	119	10314.82	149	45
RPTOS HINOMISA	531.35518	164	4627.11	164	17
RPTOS HYUNDAI	289732.656202	11048	595159.72	11655	3103
RPTOS KALMAR	905236.880413999	7893	1222481.69	12622	2316
RPTOS KUBOTA	21599.509829	3543	48505.43	2940	480
RPTOS LITREBERR	20393.17	3630	113818.11	1474	214
RPTOS LLANTAS	381.3	2	762.6	2	1
RPTOS LUBRICANTES	2878096.65212101	163638	10317694.54	113483	12187
RPTOS NIELMA	48127.171467	741	200767.1	342	175
RPTOS OTROS	1758484.873808	89545.24	4311559.6	65235.92	13764
Grand Total	49941529.1836572	739679.23	100919252.35	640053.33	132491

Figura 47 resultados de cubo de compra
Fuente: Elaboración Propia

CUBO DE STOCK

Siguiendo con la presentación de cubos no toca presentar el ultimo cubo que es un complemento a las dos ares ya mencionadas para ayuda y análisis de la información referente al stock necesario y óptimos. Este cubo esta confirmado por los siguientes tablas de dimensión.

- dim_producto
- dim_grupo
- Dim_sucursal
- Dim_almacen
- Dim_tiempo

Estructura de cubo Stock

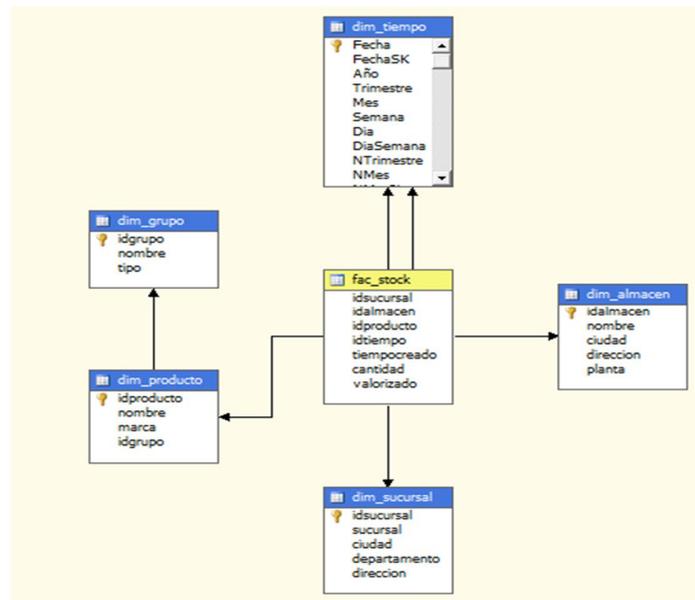


Figura 47 estructura de cubo Stock

Fuente: Elaboración Propia

Resultados de cubo

grupo nombre	Cantidad	Valorizado	Fac Stock Count	Cantidad	Valorizado	Fac Stock Count	Cantidad	Valorizado	Fac Stock Count
EQI CUMBOBI	1	12725.01	1	1	44705.06	1	2	18224.07	1
EQI ATLAS COPO	55	133854.278444	11	55	51219.880303	9	25	51219.880303	9
EQI CARGADORES	26	136619.720111	7	7	149528.83	6	32	566148.520111	13
EQI CHICAGO PNEUMATI	40	518225.921955	10	1	68486.4	1	42	545647.421955	12
EQI CLIMBERS	2	26121.51	2	2	89117.24	2	3	158873.04	3
EQI HINCHONGA	1	490016.06	1	11	3555183.12	6	3	61393.68	1
EQI HYDRAEAT	8	2997240.563333	6	2	804907.03	2	11	4252263.653333	9
EQI LIEBHERR	30	426484.872368	6	6	12424.86	5	262	3049.082405	5
EQI TSLUE	6	13243.86	5	15	209.572	2	15	209.572	2
RPTOS AME TECH	2	5886.23	2	2	5886.23	2	2	5886.23	2
RPTOS BROMMA	39	42.675	3	39	42.675	3	39	42.675	3
RPTOS CASCAD	588	107146.775482	102	205	31140.464699	53	793	138287.240181	155
RPTOS CHEN TOOLS	10.72	71.017941	5	50.72	71.017941	5	60.72	71.017941	5
RPTOS CHICAGO PNEUMA	5740	246088.089708	1023	745	25489.322165	193	6957	27964.773452	1520
RPTOS CONTEX	103.5	26584.201666	33	20	4728.44	10	123.5	31312.641666	43
RPTOS CUMBERS	213	18312.395423	56	120	626.946162	7	433	18933.841585	63
RPTOS DANAI	1244	12512.85873	213	1382	9289.853097	173	2631	21884.711827	389
RPTOS FLEETGUARD	30	11425.5725	20	1	72.92	1	31	11490.4925	21
RPTOS GENIE	90	69.303	3	90	69.303	3	90	69.303	3
RPTOS GETS	1	124.25	1	1	124.25	1	1	124.25	1
RPTOS HALLOTTE	2	325.98	2	59	5806.61275	15	1	136.94	1
RPTOS HEMINGAL	2	325.98	2	59	5806.61275	15	62	6069.53275	18
RPTOS HUCK	214	25.638815	6	214	25.638815	6	214	25.638815	6
RPTOS HURDAL	17	169.14	4	3305	11190.382364	464	255	10930.789403	65
RPTOS KALMAR	92	8226.412415	17	888	179328.268289	163	275	62947.149065	53
RPTOS KALMAR	92	8226.412415	17	888	179328.268289	163	275	62947.149065	53
RPTOS KUBOTA	969	19899.269715	132	8	2651.6	5	1238	21752.861975	195
RPTOS LAMAR	9	854.515	5	8	2651.6	5	17	8476.515	10
RPTOS LIEBHERR	38	43347.39375	14	2012	320955.308672	366	29	11820.81	11
RPTOS LILANTAS	18	1143.621666	4	18	1143.621666	4	18	1143.621666	4
RPTOS LUBRICANTES	9	453.575	4	649.25	12349.943894	25	382.5	8841.109042	43
RPTOS NEMPA	49	6973.92542	10	1	24.49	1	50	6998.08242	11
RPTOS OTSOS	11	1147.82	4	1561.6	54676.270574	188	1223.5	18519.168567	70

Figura 48 resultado de cubo stock
Fuente: Elaboración Propia

4. RESULTADOS DE REPOSTES

4.1. REPORTE DE RANKING DE ALMACENES

Este reporte podemos observar el Ranking de todos los almacenes que la empresa maneja ya sea en Lima y en provincia, podemos observar que el que tiene mayor salida de mercancía es el almacén 01 que se encuentra en la localidad de Lima seguida por el almacén 03 que se encuentra en la localidad de Piura y por último el almacén de Cajamarca que es número 13, como podemos observar el que tiene menor movimiento de todos es el almacén de Arequipa esto ayuda a la gerencia a tomar una mejor decisión de donde es aconsejable estocar de artículo y/o equipos conllevando una mejor distribución de recurso que puedan generar menor costo y pérdidas para la empresa.

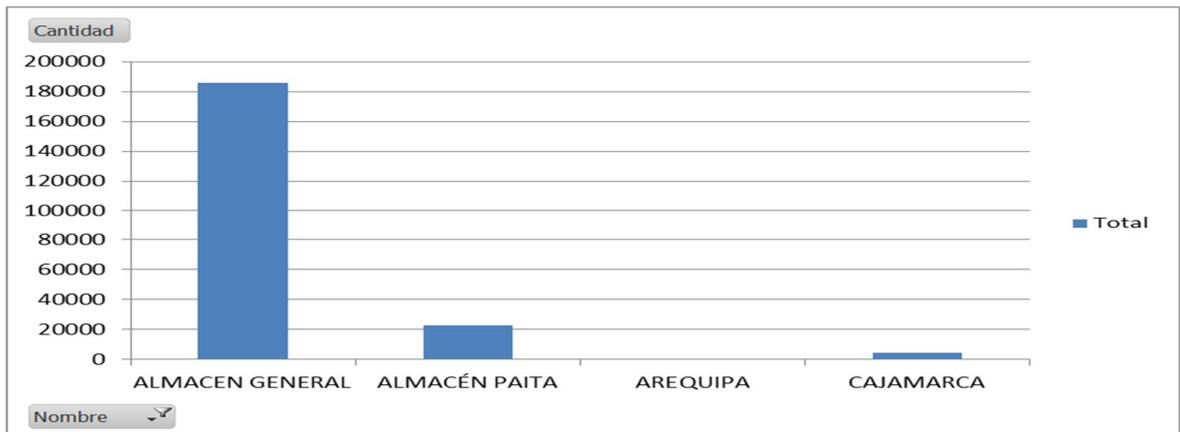


Figura 49 Ranking de almacenes

Fuente: Elaboración Propia

4.2. REPORTE DE COSTO DE ALMACÉN

En este reporte podemos observar el costo de almacén versus la rentabilidad o importe generado, esto ayuda a decidir si es necesario tener un almacén en la zona o ver otras opciones y que esto genera pérdida para la empresa. Esto corrobora el reporte anterior y da una decisión de que al macen es necesario tener activos, todo esto repercute con las ventas de la empresa y el ingreso mensual

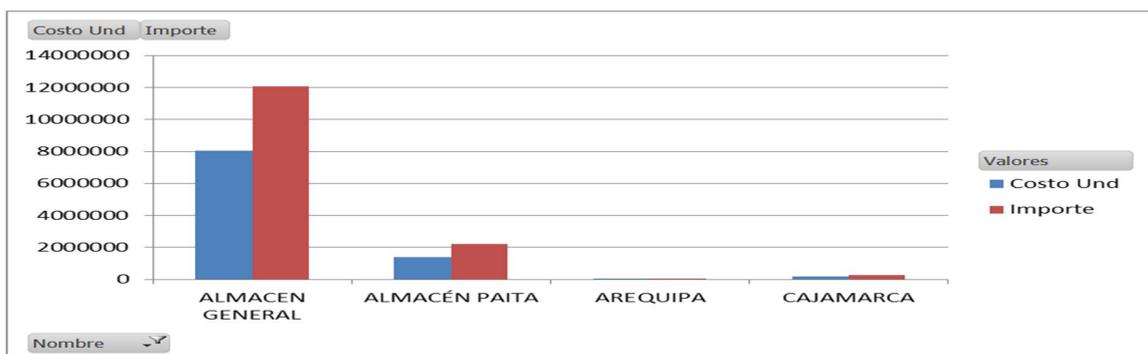


Figura 50 costos de almacén

Fuente: Elaboración Propia

4.3. REPORTE DE HISTORIAL DE VENTAS POR VENDEDOR

Este informe podemos ver el historial de todos los vendedores viendo la cantidad anual vendida, este informe ayuda al área de ventas a realizar el premio que se hace anualmente a los vendedores con más ventas realizadas y también ayuda ver la eficiencia que tienen cada uno de ellos otro aporte de este reporte es que ayuda a la gerencia de venta para ver y sectorizar la ventas y tomar un buena decisión de donde es aconsejable hacer ofertas o dar mejores precios. Esto se puede hacer ya que cada vendedor está distribuido por zona y/o ciudad

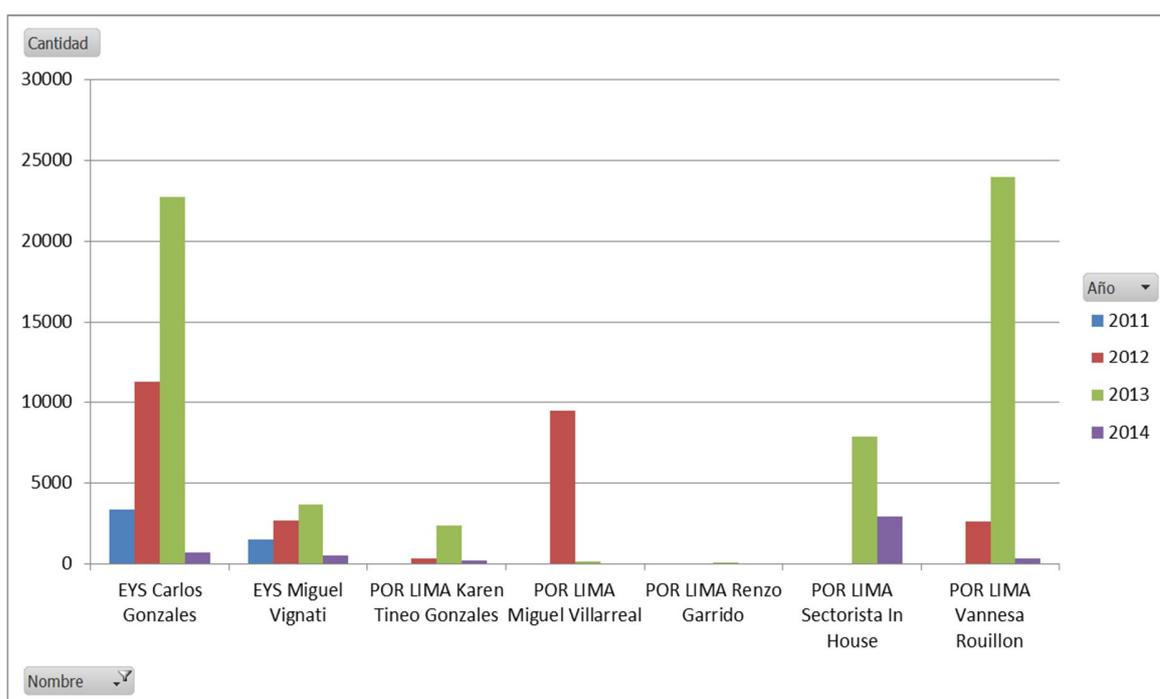


Figura 51 historial de ventas anual de vendedores

Fuente: Elaboración Propia

4.4. REPORTE DE CANTIDAD DE VENTAS MENSUALES POR AÑO Y POR VENDEDOR

Este informe muestra la venta que se tiene por cada vendedor por cada mes, esto permite tomarles el pulso a cada momento a los vendedores y exigirle que cumplan con su cuota de ventas mensual y también ayuda a la gerencia de ventas saber la cantidad de ventas por sector y por temporada haciendo que se tome un adecuada decisión en el estocqueo del almacén en el tiempo adecuado

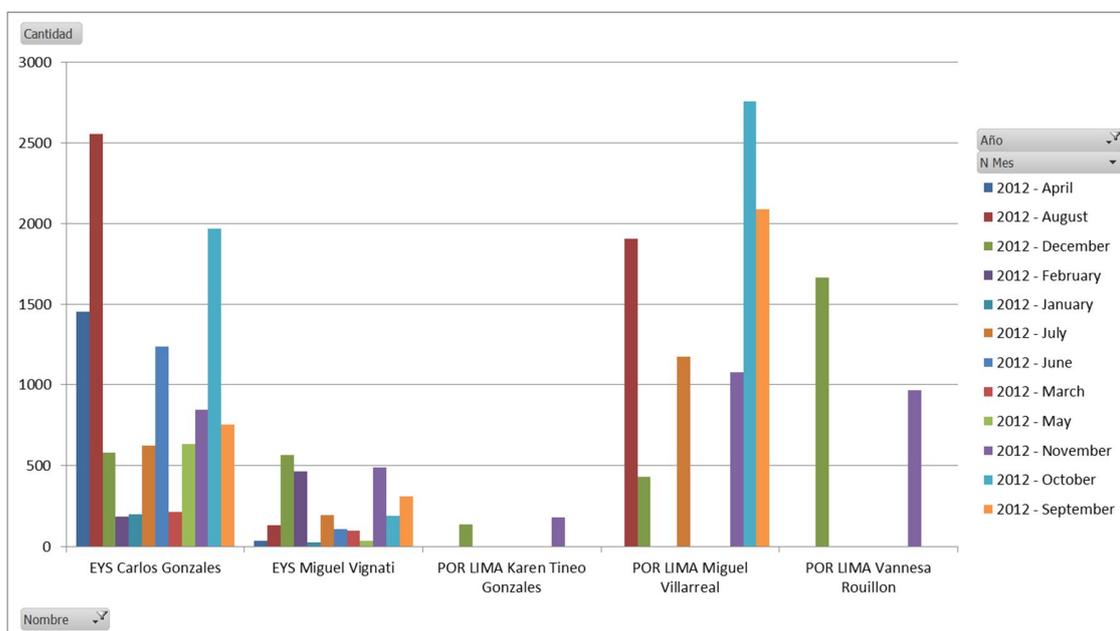


Figura 52 venta mensual por año y vendedor

Fuente: Elaboración Propia

2.4 REPORTE DE ATENCIÓN DE PEDIDOS DE PROVEEDORES EXTRANJEROS ESTRATÉGICOS POR AÑO

En este informe podemos ver cuán efectiva puede ser el servicio Postventa de nuestro proveedor estratégico ya que de esto depende mucho la toma de decisión a adecuada en la solicitud de nuevas compra con el tiempo oportuno y no perjudicando la relación de la empresa con sus clientes con tiempo de espera demasiado largo para la entrega del bien esto podría perjudicar la imagen de la empresa ante nuevo clientes

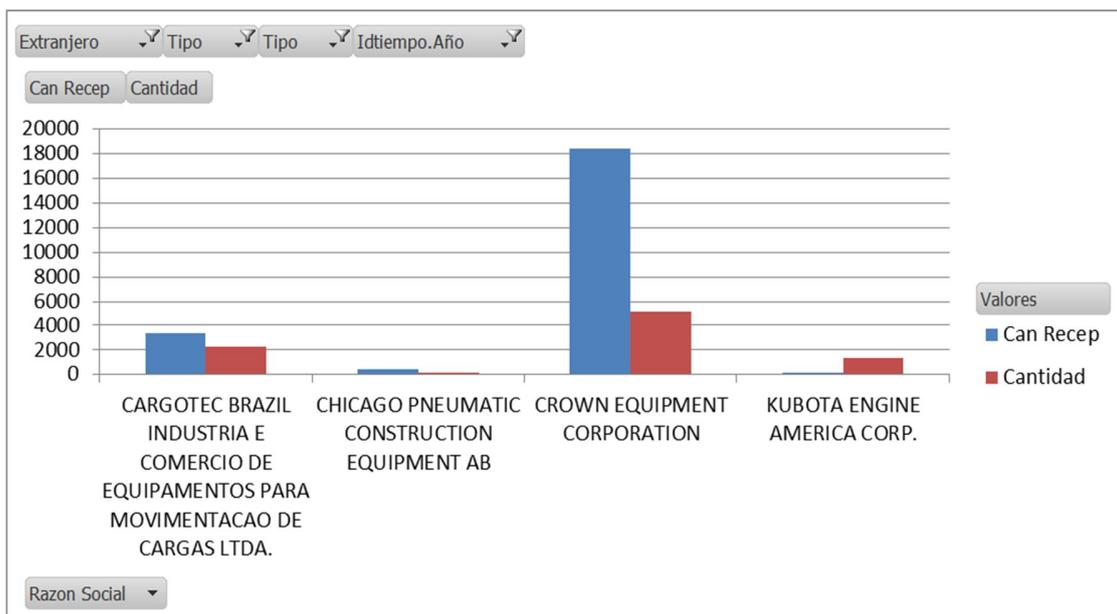


Figura 53 Atención De Pedidos De Proveedores Estratégicos Por Año

Fuente: Elaboración Propia

REPORTE DE ATENCIÓN DE PEDIDOS DE PROVEEDORES LOCALES ESTRATÉGICOS POR AÑO

Tanto los proveedor extranjeros como los proveedores locales son importantes en la toma de decisiones con respecto a si se debe comprar en el extranjero o localmente. Esto se presenta muy a menudo ya que existen clientes que necesitan el reemplazo de los equipos y estos no pueden estar mucho tiempo parados por falta de stock por parte de la empresa Triton. Es ahí donde podemos contar con los proveedores extranjeros para ello se evalúa la efectividad de la respuesta ante los pedidos.

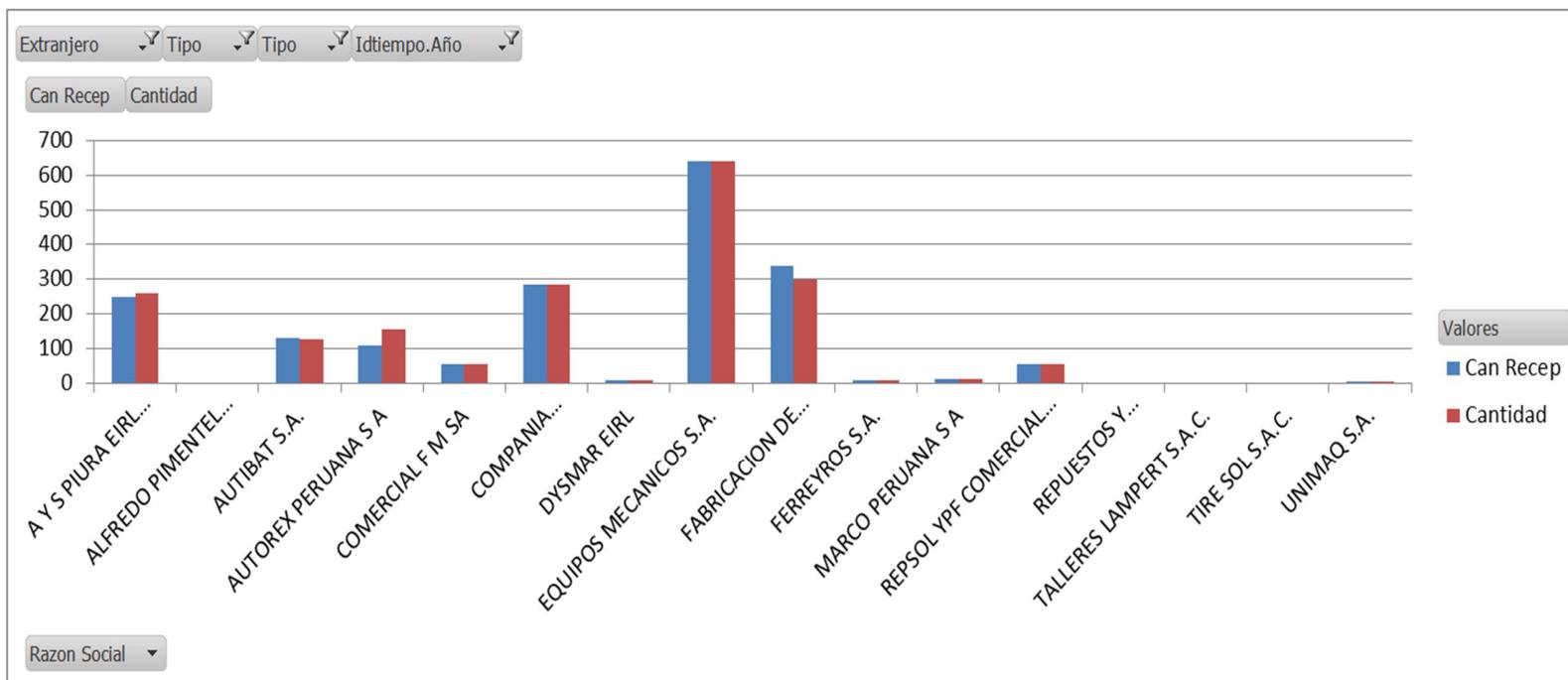


Figura 54 Atención De Pedidos De Proveedores Locales Estratégicos Por Año

Fuente: Elaboración Propia

4.5. REPORTE DE VENTAS ANUALES

Este reporte podemos ver el crecimiento de ventas que tiene la empresa Triton Trading, se puede observar que la empresa tiene un crecimiento sostenido anual esto traduce en una buena posición en el mercado peruano. Podemos ver que el 2014 siendo solo el mes de enero ya tiene una significativa venta, podemos concluir que las en 2014 podrán ser mucho mejor que el año 2013, esto aclara la visión dela empresa en la toma de decisión de incrementar l las compras para sostener el crecimiento que presenta la empresa.

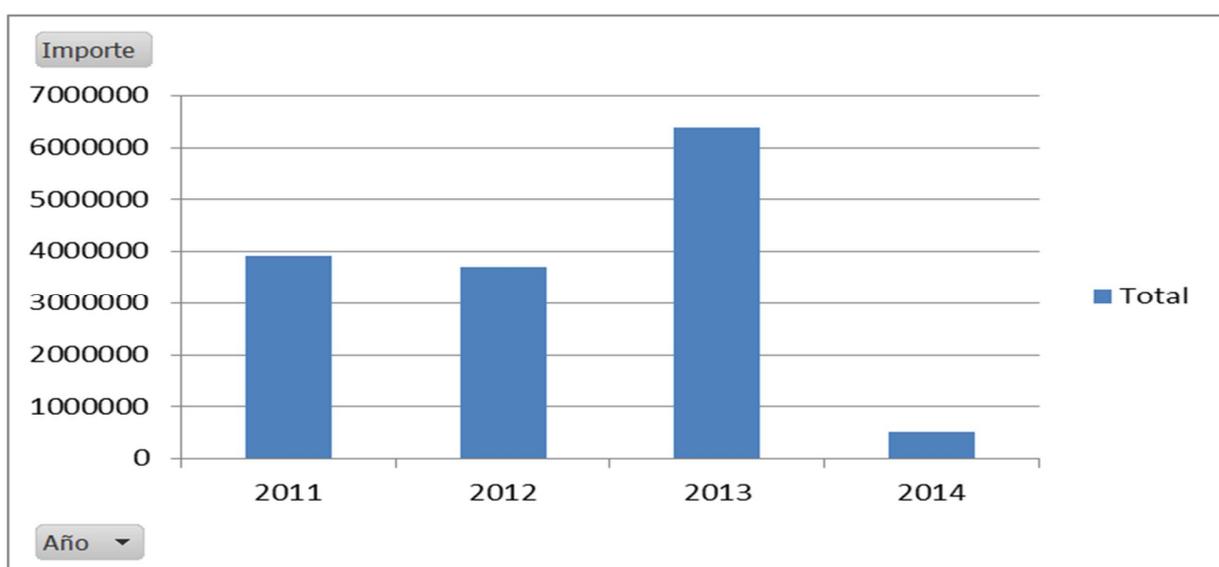


Figura 55 ventas anuales
Fuente: Elaboración Propia

4.6. REPORTE DE MOVIMIENTO DE STOCK POR AÑO Y SUCURSAL

En este reporte podemos observar los movimientos que se hicieron en todo el año y que podemos observar un flujo constante de stock ayuda al gerente del área de compras tomar la decisión de si es necesario el estoqueo de las sucursales para que puedan desarrollar las actividades de venta sin ningún problema. Esto ayuda a ver mejor el panorama de ventas.

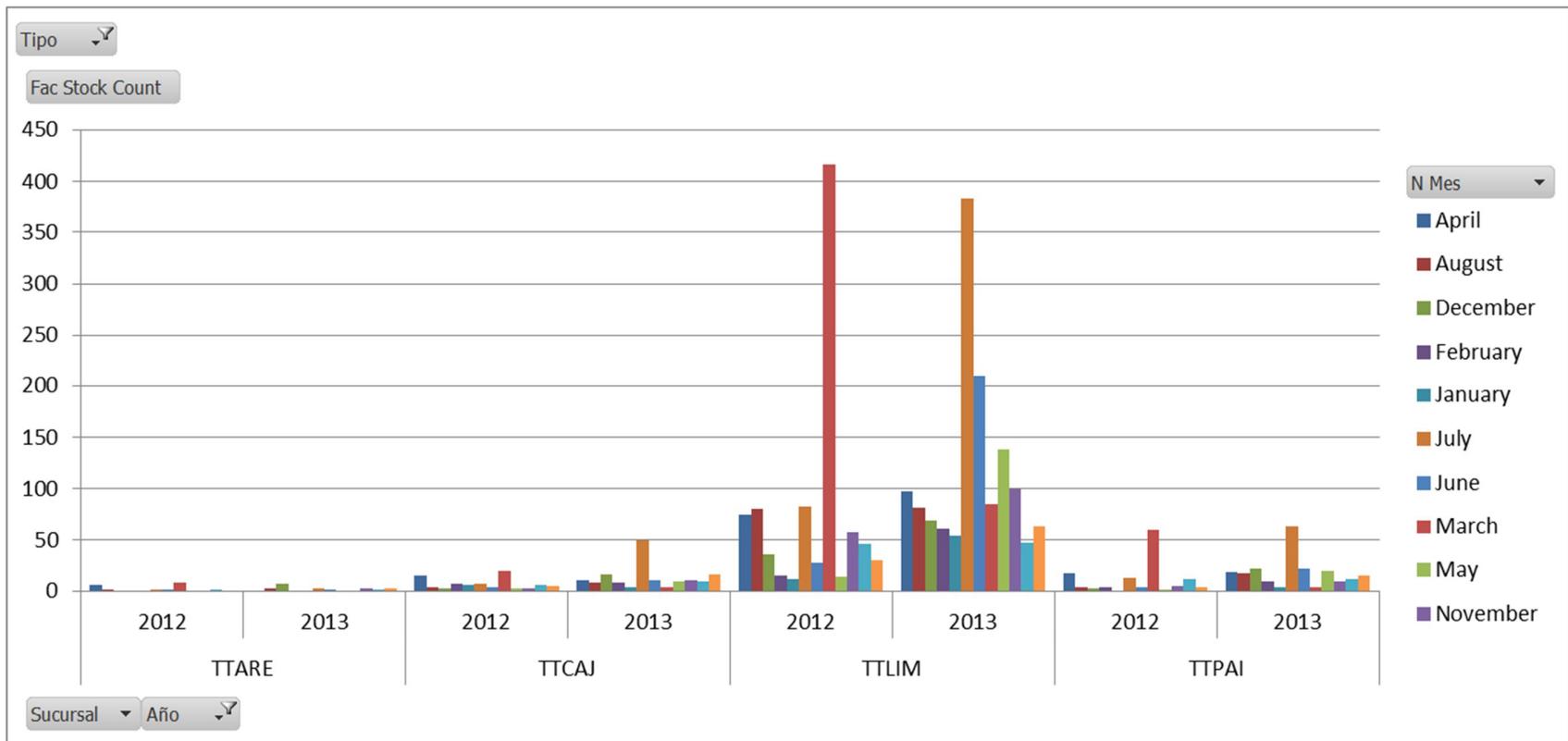


Figura 56 Movimiento de estock por años

Fuente: Elaboración Propia

4.7. VENTAS DE EQUIPOS ANUAL POR AÑO Y EMPRESA

En este reporte podemos ver las ventas anuales de empresa que compraron maquinaria, esto se puede traducir en nuevas ventas de repuesto para la maquina compradas y tambien servicio de mano de obra la decisio si es necesario la ompra o el estokeo de repuesto para las maquina mas vendidas.

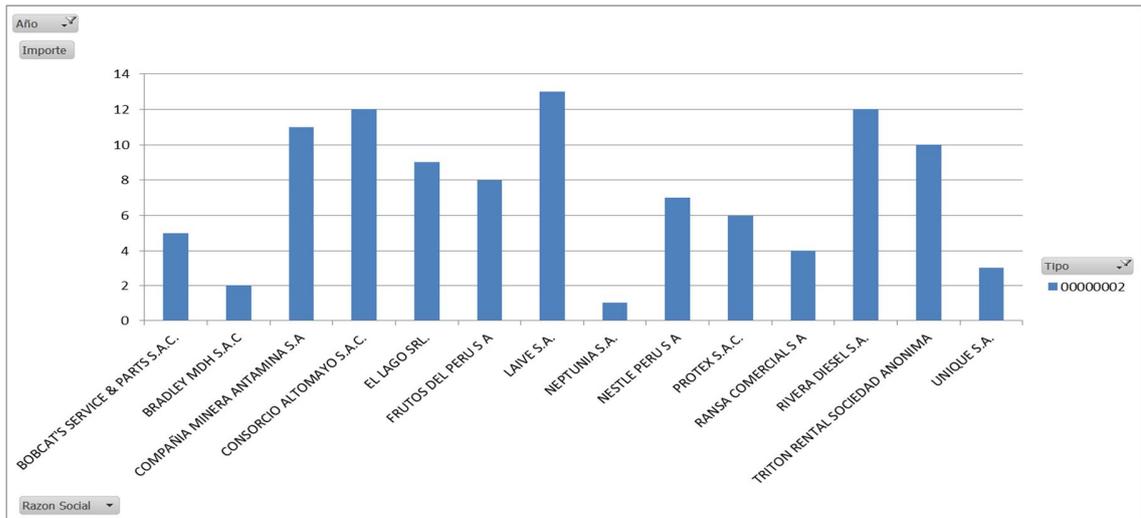


Figura 57 Venta de equipo por año y empresas

Fuente: Elaboración Propia

COCLUSIONES

- El trabajo de investigación demuestra en forma clara y concisa de como los avances tecnológicos pueden influenciar en los procesos de las organizaciones y en especial en la gestión empresarial.
- En base a las características propias de la empresa el uso de la metodología de Ralph Kimball resulta una solución eficaz en tiempo y recursos debido a que abarca la solución al problema en un corto plazo.
- Esta solución sirve de base para la futura construcción de un Data Warehouse que mejoraría la toma decisión ya no respecto a las áreas de ventas y compra sino en toda la empresa haciendo mejor el control y administración de la empresa.
- Las aplicaciones de Inteligencia de Negocio permiten la integración de la empresa con sus proveedores, empleados y socios. B.I. es un concepto que trata de englobar todos los sistemas de información de una organización para obtener de ellos no solo información o conocimiento, si no una verdadera inteligencia que le confiera a la organización una ventaja competitiva por sobre sus competidores.
- Se puede decir satisfactoriamente que el gerente cuentan con una información en tiempo real del proceso de venta y compra cual les permite tomar decisiones mucho más acertadas y adecuadas para la Gestión de ventas y compra.

RECOMENDACIONES

La aplicación, basada en Inteligencia de Negocio, también puede ser usada por todo tipo de organizaciones y empresas que quieran ver masificados su productividad y tener el conocimiento en tiempo real, de esta forma se elimina el factor distancia y se reducen costos.

Se debe brindar capacitación al personal con el fin de garantizar la eficiencia y productividad en su trabajo.

Cabe señalar como recomendación que las redes actuales disponen de equipos optimizados para funcionar con el protocolo IPv4, lo cual no implica su imposibilidad de ser actualizados a la nueva versión, sin embargo la performance se ve degradada en forma significativa. Los equipos modernos, por el contrario funcionan mejor con la nueva versión.

BIBLIOGRAFIA

Clover ETL. (1 de Julio 14 de 2010). CloverETL. Recuperado el 30 de Mayo de 2012, de CloverETL: <http://www.cloveretl.com/products/desktop-edition>

Hernandez, J. (2004). Introducción a la minería de datos. Madrid: Pearson.

Hyde, J. (19 de Agosto de 2006). Olap4j. Recuperado el 7 de Junio de 2012, de Olap4j: http://www.olap4j.org/olap4j_fs.html

Ponniah, P. (2007). Data modeling fundamentals. Usa: Wiley.

Inmon, W. (2005). Building the Datawarehouse. Usa: Wiley.

Kimball, R. (2008). The datawarehouse lifecycle . Usa: Wiley.

http://www.cse.hcmut.edu.vn/~chauvtn/data_mining/Texts/%5B6%5D%202006%20Data%20Mining%20Methods%20and%20Models.pdf

Harjinder, G. (1996). La integración de información para la mejor toma de decisiones. Mexico: Prentice Hall.

University of California. (17 de noviembre de 2010). UCLA. Recuperado el 19 de abril de 2012, de UCLA:

http://www.anderson.ucla.edu/faculty/jason.frand/teacher/technologies/palace/data_mining.htm

Prudsys. (8 de Marzo de 2011). Prudsys. Recuperado el 14 de Junio de 2012, de Prudsys: <http://www.prudsys.de/en/technology/xelopes/>

Oracle. (13 de Enero de 2012). Oracle. Recuperado el 20 de enero de 2014, de Oracle: <http://www.oracle.com/us/solutions/business-analytics/business-intelligence/index.html?origref=http://www.oracle.com/us/solutions/enterprise-performance-bi/enterprise-edition-066546.html>

ANEXOS

CÓDIGOS SQL PARA CREACIÓN DE TABLA FAC

```
use [datamart]

create table dim_sucursal
(idsucursal NUMERIC primary key NOT NULL,
sucursal varchar (50),
ciudad varchar (50)
)

create table dim_tiempo(
idtiempo int identity (1,1) primary key not null,
año integer,
mes integer,
fecha datetime,
)

create table dim_grupo(
idgrupo numeric primary key not null,
nombre varchar (100),
tipo varchar(50)
)

create table dim_cliente(
idcliente varchar (15) primary key not null,
ruc varchar (32) ,
razon_social varchar (255),
direccion varchar(255),
pais varchar (15),
ciudad varchar(50))

create table dim_producto(
idproducto varchar (20) primary key not null,
nombre varchar (100),
marca varchar (100),
idgrupo numeric not null,
--foreign key (idgrupo) references dim_grupo(idgrupo)
)

create table dim_almacen(
idalmacen varchar (8) primary key not null,
nombre varchar (100)
)

create table dim_vendedor
(idvendedor numeric primary key not null,
nombre varchar (100),
localidad NUMERIC
)
```

```

create table fac_ventas
(idproducto varchar(20) not null,
idalmacen varchar(8) not null,
idcliente varchar(15) not null,
idsucursal int not null,
idtiempo int not null,
idvendedor int not null,
costo_und numeric (19,6),
cantidad numeric ,
importe numeric,

foreign key (idcliente) references dim_cliente(idcliente),
foreign key (idproducto) references dim_producto(idproducto),
foreign key (idalmacen) references dim_almacen(idalmacen),
foreign key (idsucursal) references dim_sucursal(idsucursal),
foreign key (idtiempo) references dim_tiempo(fecha),
foreign key (idvendedor) references dim_vendedor(idvendedor),

)

USE [ETL_datamar]
INSERT INTO fac_ventas
select cast (x2.itemcode as varchar (20)) 'IDPRODUCTO', x2.whscode
'ALMACEN'

, x1.cardcode 'cliente',

CAST ((case x1.U_siteidot
when 'TTLIM' THEN 3
WHEN 'TTARQ' THEN 1
WHEN 'TTCAJ' THEN 2
WHEN 'TTPAI' THEN 4 END) as numeric) 'SUCURSAL'

, cast (x1.docdate as date) 'FECHA', CAST (x1.slpcode AS
NUMERIC) 'vendedor' ,x2.price ,x2.quantity 'cantidad' , x2.totalfrgn ,
cast(x2.currency as varchar(3))

from [febo].[SBO_TRITON].dbo.oinv x1

left join [febo].[SBO_TRITON].dbo.invl x2 on x2.docentry=x1.docentry
where x2.itemcode in (select idproducto collate
SQL_Latin1_General_CP850_CI_AS From dim_producto)
and x2.whscode in (select idalmacen collate
SQL_Latin1_General_CP850_CI_AS from dim_almacen)
and x1.cardcode in(select idcliente collate
SQL_Latin1_General_CP850_CI_AS from dim_cliente)
and x1.U_siteidot in (select sucursal collate
SQL_Latin1_General_CP850_CI_AS from dim_sucursal)

```

```
ORDER BY x1.docdate
```

```
create table dim_proveedor(  
idproveedor varchar (15) primary key not null,  
ruc varchar (32) ,  
razon_social varchar (255),  
direccion varchar(255),  
pais varchar (15),  
ciudad varchar(50),  
tipo nvarchar(max),  
fecha_creacion datetime )
```

```
create table dim_tipo_compra(  
idtipoCompra varchar (15) primary key not null,  
compra varchar (32) )
```

```
create table dim_centro_costo(  
idcentroCosto nvarchar(8) primary key not null,  
nombre nvarchar(30)  
)
```

```
create table dim_moneda(  
idmoneda varchar (5) primary key not null,  
moneda varchar (30),  
siclas varchar(10))
```

```
create table fac_compras  
(idproducto varchar(20) not null,  
idalmacen varchar(8) not null,  
idproveedor varchar(15) not null,  
idcentroCosto nvarchar(8) not null,  
idmoneda varchar (5) not null,  
idtipoCompra varchar (15) not null,  
fecha_recep date not null,  
idtiempo date not null,  
costo_und numeric (19,6),  
cantidad numeric (19,6) ,  
importe numeric (19,6),  
can_recep numeric (19,6),  
impor_recpc numeric(19,6),
```

```
foreign key (idproveedor) references dim_proveedor(idproveedor),  
foreign key (idcentroCosto) references dim_centro_costo(idcentroCosto),  
foreign key (idmoneda) references dim_moneda(idmoneda),  
foreign key (idtipoCompra) references dim_tipo_compra(idtipoCompra),  
foreign key (fecha_recep) references dim_tiempo(fecha),  
foreign key (idproducto) references dim_producto(idproducto),  
foreign key (idalmacen) references dim_almacen(idalmacen),  
foreign key (idtiempo) references dim_tiempo(fecha)
```

)

```
create table fac_stock(
```

```
idsucursal numeric (18,0) not null,
```

```
idalmacen varchar(8) not null,
```

```
idproducto varchar(20) not null,
```

```
cantidad numeric(19,6),
```

```
valorizado numeric(19,6),
```

```
foreign key (idsucursal) references dim_sucursal(idsucursal),
```

```
foreign key (idproducto) references dim_producto(idproducto),
```

```
foreign key (idalmacen) references dim_almacen(idalmacen)
```

)