

**UNIVERSIDAD PRIVADA ANTENOR ORREGO**  
**FACULTAD DE INGENIERÍA**  
**ESCUELA PROFESIONAL DE INGENIERÍA DE COMPUTACIÓN**  
**Y DE SISTEMAS**



**TRABAJO DE TESIS PARA OBTENER EL TITULO PROFESIONAL DE**  
**INGENIERO DE COMPUTACION Y SISTEMAS**

---

**SISTEMA DE SOPORTE DE DECISIONES BASADA EN LA**  
**ARQUITECTURA BUSINESS INTELLIGENCE (BI) DE MICROSOFT**  
**AZURE Y MS SQL SERVER 2014 PARA EL AREA DE VENTAS DE LA**  
**EMPRESA COPY VENTAS S.R.L. TRUJILLO PARA EL AÑO 2016.**

---

**Área de Investigación:**

0403 0703a. Desarrollo de modelos, aplicaciones y arquitecturas de sistemas para toma de decisiones y dirección estratégica.

**Autor:**

**BR. JORGE JHOEL VÁSQUEZ TIRADO**

**Asesor:**

**ING. AGUSTÍN ULLÓN RAMÍREZ**

**2016**

**“SISTEMA DE SOPORTE DE DECISIONES BASADA EN LA  
ARQUITECTURA BUSINESS INTELLIGENCE (BI) DE MICROSOFT  
AZURE Y MS SQL SERVER 2014 PARA EL AREA DE VENTAS DE LA  
EMPRESA COPY VENTAS S.R.L. TRUJILLO PARA EL AÑO 2016”**

**Elaborado por:**

Br. Jorge Jhoel Vásquez Tirado

**Aprobada por:**

---

**Ing. Heber Gerson Abanto Cabrera  
Presidente  
CIP: 106421**

---

**Ing. Freddy Henry Infantes Quiroz  
Secretario  
CIP: 139578**

---

**Ing. José Arturo Castañeda Saldaña  
Vocal  
CIP: 48234**

---

**Ing. Agustín Eduardo Ullón Ramírez  
Asesor  
CIP: 137602**

## **PRESENTACIÓN**

Señores Miembros del Jurado:

De conformidad y en cumplimiento de los requisitos estipulados en el reglamento de grados y Títulos de la Universidad Privada Antenor Orrego y el Reglamento Interno de la Escuela Profesional de Ingeniería de Computación y Sistemas, pongo a vuestra disposición el presente Trabajo de Tesis: “SISTEMA DE SOPORTE DE DECISIONES BASADA EN LA ARQUITECTURA BUSINESS INTELLIGENCE (BI) DE MICROSOFT AZURE Y MS SQL SERVER 2014 PARA EL AREA DE VENTAS DE LA EMPRESA COPY VENTAS S.R.L. TRUJILLO PARA EL AÑO 2016” para obtener el Título Profesional de Ingeniero de Computación y Sistemas.

El contenido de la presente tesis ha sido desarrollado tomando como marco de referencia los lineamientos establecidos por la Facultad de Ingeniería, la Escuela Profesional de Ingeniería de Computación y Sistemas y los conocimientos adquiridos durante mi formación profesional, además de consulta de fuentes bibliográficas.

El Autor.

## **DEDICATORIA**

A mi esposa por el apoyo y por ser la fuerza para alcanzar mis metas como profesional y en la formación como persona.

A mi familia por el cariño y apoyo en todo momento para el desarrollo de mi vida como estudiante y profesional



## **AGRADECIMIENTO**

Quiero expresar mi agradecimiento al personal de las diferentes bibliotecas de la Universidad Privada Antenor Orrego y docentes e ingenieros, quienes me brindaron todas las facilidades para conocer más sobre las herramientas analizadas en la presente tesis.

También agradezco al el Ing. Agustín Ullón, por su apoyo y asesoramiento en el desarrollo de la presente Tesis.

Y a todos aquellos que prefirieron permanecer en el anonimato pero que estuvieron siempre conmigo durante todo el camino de este trabajo de tesis.

Muchas Gracias.

**El autor.**

## **RESUMEN**

# **“SISTEMA DE SOPORTE DE DECISIONES BASADA EN LA ARQUITECTURA BUSINESS INTELLIGENCE (BI) DE MICROSOFT AZURE Y MS SQL SERVER 2014 PARA EL AREA DE VENTAS DE LA EMPRESA COPY VENTAS S.R.L. TRUJILLO PARA EL AÑO 2016”**

Por:

Br. Jorge Jhoel Vásquez Tirado

Competitividad e innovación son nociones estrechamente ligadas a la Inteligencia de Negocios y, en este sentido, la empresa moderna debe estar consciente de la necesidad de crear un ambiente propicio para lograr que la efectiva transferencia de la información y que la aplicación del conocimiento obtenido se efectúe de una forma transparente en todas las actividades de sus trabajadores y, en todas las instancias de la estructura organizacional.

Es aquí donde entra una de las herramientas más mencionadas últimamente los sistemas de soporte de decisiones: Business Intelligence en la nube va permitir ayudar a las empresas a adquirir un mejor entendimiento de ellas mismas. Esto gracias a la capacidad de explotar su información, con la intención de poder manipularlos de una manera más sencilla y entender el porqué de nuestro desempeño o, mejor aún, plantear escenarios a futuro, lo cual nos ayudará a tomar mejores decisiones.

La Empresa Copy Ventas S.R.L. es una empresa que se dedica a la comercialización de Libros y útiles escolares para instituciones educativas y también para personas naturales. La empresa tiene ya 10 años en este rubro y se ha convertido en un espejo para otras muchas empresas por su sistema de gestión, cuidado, moderno y fácilmente adaptable a las circunstancias del mercado, primando, antes, ahora y después, las necesidades del cliente, otorgándole a éste, la oportunidad de conseguir cultura al mejor precio.

La empresa ha decidido evaluar la implementación de un sistema de soporte de decisiones capaz de dar una solución a la gestión del flujo de información y toma de decisiones. Herramientas que dan apoyo a esta problemática son las Aplicaciones de Inteligencia de Negocios las cuales pueden ser utilizadas para mejorar la gestión de la Fuerza de Venta.

## **ABSTRACT**

### **DECISION SUPPORT SYSTEM BASED ARCHITECTURE BUSINESS INTELLIGENCE (BI) OF MICROSOFT AZURE AND MS SQL SERVER 2014 FOR AREA SALES COMPANY SALES COPY S.R.L. TRUJILLO 2016**

By:

Br. Jorge Jhoel Vásquez Tirado

Competitiveness and innovation are concepts closely linked to Business Intelligence and in this sense, the modern enterprise must be aware of the need to create an enabling environment to achieve the effective transfer of information and the application of knowledge gained is made in a transparent manner in all activities of their workers and, in all instances of the organizational structure.

This is where comes some of the tools mentioned lately the decision support system: Business Intelligence in the cloud will help enable companies to gain a better understanding of themselves. This is due to the ability to exploit your information, with the intention to manipulate them in a simpler way and understand why our performance or, even better, raise future scenarios, which will help us make better decisions.

Copy Sales Company S.R.L. It is a company dedicated to the marketing of books and school supplies for educational institutions and for individuals. The company has 10 years in this area and has become a mirror for many other companies for their management system, care, modern and easily adaptable to market conditions, prevailing before now and then customer needs , giving it the opportunity to get the best price culture.

The company has decided to evaluate the implementation of a decision support system capable of providing a solution to the management of information flow and decision making. Tools that support this problem are the Business Intelligence Applications which can be used to improve the management of the Sales Force.

# ÍNDICE DE CONTENIDOS

	Pág.
Carátula	
Hoja para firma de Jurados y Asesor .....	ii
Presentación.....	iii
Dedicatorias.....	iv
Agradecimiento.....	v
Resumen .....	vi
Abstract .....	vii
Índice de contenidos .....	viii
Índice de Figuras .....	xi
Índice de tablas.....	xii
<b>INTRODUCCIÓN.....</b>	<b>xiii</b>
<b>CAPITULO I: FUNDAMENTO TEÓRICO, MATERIALES Y METODOS</b>	
<b>I. FUNDAMENTO TEÓRICO .....</b>	<b>01</b>
<b>1.1.1. Integración de datos .....</b>	<b>01</b>
<b>1.1.2. Proceso de Toma de Decisiones .....</b>	<b>03</b>
<b>1.1.3. Tecnologías de Información.....</b>	<b>05</b>
<b>1.1.4. Sistema de Información.....</b>	<b>06</b>
<b>1.1.5. Procesamiento Transaccional en Línea (OLTP).....</b>	<b>10</b>
<b>1.1.6. Datawarehouse .....</b>	<b>11</b>
<b>1.1.7. Introducción a Datamarts .....</b>	<b>13</b>
<b>1.1.8. Arquitectura del Datawarehouse.....</b>	<b>13</b>
<b>1.1.9. Introducción al Procesamiento Analítico en Línea.....</b>	<b>14</b>
<b>1.1.10. Arquitectura OLAP .....</b>	<b>16</b>
<b>1.1.11. El modelo de datos OLAP .....</b>	<b>16</b>
<b>1.1.12. Implementación del OLAP .....</b>	<b>19</b>
<b>1.1.13. CUBOS OLAP .....</b>	<b>21</b>
<b>1.1.14. Agregados.....</b>	<b>24</b>
<b>1.1.15. Business Intelligence (BI).....</b>	<b>25</b>
<b>1.1.16. Microsoft SQL Server 2014.....</b>	<b>26</b>
<b>1.1.17. Analysis Services 2014 .....</b>	<b>28</b>
<b>1.1.18. Microsoft Azure .....</b>	<b>30</b>
<b>II. METODOLOGÍA .....</b>	<b>31</b>

1.II.1. Metodología de Ralph Kimball .....	31
--	----

## CAPITULO II: DESARROLLO DEL PROYECTO

<b>I. PLANEACION Y ADMINISTRACION DE PROYECTO .....</b>	<b>38</b>
2.I.1. El negocio .....	38
2.I.2. Selección de la Estrategia de Implementación .....	41
2.I.3. Selección de la Metodología de Desarrollo .....	42
2.I.4. Selección del Ámbito de Implementación.....	42
2.I.5. Selección del enfoque arquitectónico .....	43
2.I.6. Desarrollo de un Programa y del Presupuesto del Proyecto.....	44
2.I.7. Desarrollo del escenario del uso empresarial.....	46
<b>II. DETERMINACIÓN DE LOS REQUERIMIENTOS.....</b>	<b>47</b>
2.II.1. Requerimientos del Propietario .....	47
2.II.2. Requerimientos del Usuario Final .....	48
2.II.3. Requerimientos no Funcionales.....	49
2.II.4. Análisis de los requerimiento .....	49
<b>III. DISEÑO TÉCNICO DE LA ARQUITECTURA.....</b>	<b>53</b>
2.III.1. Modelo Startnet.....	53
2.III.2. Nivel de Datos .....	54
2.III.3. Nivel Técnico .....	55
<b>IV. MODELADO DIMENSIONAL .....</b>	<b>56</b>
2.IV.1. Identificación de los Componentes del Modelo .....	56
2.IV.2. Diagrama de la Tabla de Hechos .....	60
2.IV.3. Esquema Estrella .....	65
<b>V. DISEÑO FÍSICO.....</b>	<b>67</b>
2.V.1. Determinación de las agregaciones .....	73
2.V.2. Construcción de las Tablas y la Base de Datos en SQL.....	74
<b>VI. PROCESO DE EXTRACCIÓN, TRANSFORMACIÓN Y CARGA DE DATOS .....</b>	<b>77</b>
<b>VII. SELECCIÓN DE PRODUCTOS .....</b>	<b>91</b>
2.VII.1. Hardware .....	91
2.VII.2. Software .....	91

<b>VIII. ESPECIFICACIÓN DE LA APLICACIÓN DEL USUARIO FINAL .....</b>	<b>92</b>
<b>2.VIII.1.Estructura de los Cubos .....</b>	<b>92</b>
<b>IX. DESARROLLO DE LA APLICACIÓN DEL USUARIO FINAL .....</b>	<b>98</b>
<b>X. REPORTES EN POWERPIVOT .....</b>	<b>101</b>
<b>XI. IMPLEMENTACION EN MICROSOFT AZURE .....</b>	<b>106</b>
<b>CAPITULO III: DISCUSIÓN DE LA HIPOTESIS .....</b>	<b>110</b>
<b>CONCLUSIONES.....</b>	<b>122</b>
<b>RECOMENDACIONES.....</b>	<b>123</b>
<b>REFERENCIAS .....</b>	<b>124</b>
<b>ANEXOS .....</b>	<b>127</b>

## ÍNDICE DE FIGURAS

	Pág.
<b>Figura N° 1:</b> Tipos de Sistemas de Información .....	07
<b>Figura N° 2:</b> Modelo de un ESS.....	10
<b>Figura N° 3:</b> Estructura básica Datawarehouse.....	14
<b>Figura N° 4:</b> Arquitectura básica para OLAP .....	16
<b>Figura N° 5:</b> Un esquema de estrella .....	17
<b>Figura N° 6:</b> Un esquema de copo de nieve.....	18
<b>Figura N° 7:</b> Cubos OLAP ... ..	22
<b>Figura N° 8:</b> Esquemas Relacionales y cubo .....	22
<b>Figura N° 9:</b> Jerarquías de los cubos OLAP ... ..	23
<b>Figura N° 11:</b> Medidas de un cubo .....	24
<b>Figura N° 12:</b> Business Intelligence .....	25
<b>Figura N° 13:</b> Metodología de implementación.....	32
<b>Figura N° 14:</b> Cronograma de actividades .....	47
<b>Figura N° 15:</b> Diagrama E-R de la base de datos .....	56
<b>Figura N° 16:</b> Enfoque Arquitectónico del Datamart.....	57
<b>Figura N° 17:</b> Tabla de Hechos del Data Mart .....	62
<b>Figura N° 18:</b> Esquema Estrella del Data Mart .....	68
<b>Figura N° 19:</b> Proceso ETL.....	82
<b>Figura N° 20:</b> Diseño de Cubo atenciones .....	95

## ÍNDICE DE TABLAS

	Pág.
<b>Tabla N° 1:</b> Diferencia entre Sistemas Transaccionales y Datawarehouse .....	12
<b>Tabla N° 2:</b> Metodología Kimball.....	36
<b>Tabla N° 3:</b> Personal involucrado en el Proyecto .....	48
<b>Tabla N° 4:</b> Descripción de usuarios del Data Mart .....	49
<b>Tabla N° 5:</b> Consultas y medidas .....	50
<b>Tabla N° 6:</b> Objetivos de las tablas de Hecho .....	61
<b>Tabla N° 7:</b> Detalle de las claves de las dimensiones .....	63
<b>Tabla N° 8:</b> Nombres estándares para las Tablas Hechos y Dimensiones.....	69
<b>Tabla N° 9:</b> Nombres estándares para los atributos de las Dimensiones .....	70
<b>Tabla N° 10:</b> Nombres estándares para los atributos de Tablas de Hechos.....	70
<b>Tabla N° 11:</b> Determinación de las Agregaciones .....	75



## INTRODUCCIÓN

Actualmente el entorno empresarial está sometido a un estado de cambio permanente, acelerado e interdependiente, para enfrentarlo las organizaciones deben ser ágiles, capaces de enfrentar el ambiente competitivo y hacer las cosas de forma diferenciada, para ello el conocimiento es la clave para mantener la actitud abierta hacia el cambio y la mejora constante.

El valor fundamental de una empresa reside en su capacidad para superar a sus competidores en términos de costos y calidad de los productos y servicios que ofrece; es decir, de su competitividad. Esta última, a su vez, depende considerablemente de la habilidad de la empresa para adaptarse a las exigencias de su mercado y de su entorno. Una variable que considera cualquier compañía exitosa hoy en día es la innovación y la gestión de su conocimiento solo así se garantiza su presencia en el mercado.

Competitividad e innovación son nociones estrechamente ligadas a la Inteligencia de Negocios y, en este sentido, la empresa moderna debe estar consciente de la necesidad de crear un ambiente propicio para lograr que la efectiva transferencia de la información y que la aplicación del conocimiento obtenido se efectúe de una forma transparente en todas las actividades de sus trabajadores y, en todas las instancias de la estructura organizacional. Como se obtiene, manipula y usa la información junto a la capacidad de generar conocimiento será el factor que determine si la empresa gana o pierde.

Es aquí donde entra unas de las herramientas más mencionadas últimamente los sistemas de soporte de decisiones: Business Intelligence en la nube va permitir ayudar a las empresas a adquirir un mejor entendimiento de ellas mismas. Esto gracias a la capacidad de explotar su información, con la intención de poder manipularlos de una manera más sencilla y entender el porqué de nuestro desempeño o, mejor aún, plantear escenarios a futuro, lo cual nos ayudará a tomar mejores decisiones.

La Empresa Copy Ventas S.R.L. es una empresa que se dedica a la comercialización de Libros y útiles escolares para instituciones educativas y también para personas naturales. La empresa tiene ya 10 años en este rubro y se ha convertido en un espejo para otras muchas empresas por su sistema de gestión, cuidado, moderno y fácilmente adaptable a las

circunstancias del mercado, primando, antes, ahora y después, las necesidades del cliente, otorgándole a éste, la oportunidad de conseguir cultura al mejor precio.

La empresa necesita mantenerse a la vanguardia de la tecnología para ser líder en la región La Libertad brindando una mejor atención a sus clientes teniendo a la disposición la información de manera oportuna, permitiendo el mejor control y gestión de ellas. Es por ello que presentamos esta herramienta tecnológica; que permitirá brindara un soporte a las tomas de decisiones para la empresa con información confiable, precisa y oportuna.

La empresa ha decidido evaluar la implementación de un sistema de soporte de decisiones capaz de dar una solución a la gestión del flujo de información y toma de decisiones. Herramientas que dan apoyo a esta problemática son las Aplicaciones de Inteligencia de Negocios las cuales pueden ser utilizadas para mejorar la gestión de la Fuerza de Venta.

Actualmente se necesita profundizar las técnicas y herramientas que también soporten el desarrollo evolutivo para las bases de datos; y en nuestra opinión la más importante de estas técnicas es la inteligencia de negocios en la nube.

En la actualidad en la empresa se preparan reportes y/o informes, que son requeridas por parte de la Administración, para la solución de problemas y así poder tomar decisiones.

El principal problema de la empresas radica en que esta no utiliza directamente todo su potencial basado en el conocimiento para enfrentar día a día los cambios establecidos por la dinámica de mercado, debido a que generalmente no se encuentran organizados los procesos de generación y explotación del conocimiento, incluso la cultura organizacional instaurada no ayuda a su uso.

Es así, que la propuesta de Inteligencia de Negocios basado en los conocimientos adquiridos por la organización, apoyan a los procesos asociados a la Gestión del Conocimiento potenciando la generación de nuevas ventajas competitivas pudiendo incluso crear nuevos modelos de negocio.

La dispersión de información ocasiona que la creación y entrega de reportes tome un tiempo considerable. Algunos de los principales costos de realizar Reportes son:

- Tiempo Perdido: Realizar reportes incluye el tiempo que toma recopilar información, construir formulas, gráficos, verificar información y compartir el archivo.
- Estándares Definidos y Errores Humanos: Hay redundancia al hacer uso de diferentes fórmulas o conceptos para explicar una misma situación, al no haber un estándar definido los diseños de algunas planillas son difíciles de leer o contienen errores humanos.
- Talento Perdido: Cargos altos en cualquier departamento, pierde horas en crear reportes en vez de estar tomando decisiones sobre reportes que deberían estar terminados.
- Riesgos de Dependencia: Individuos específicos en la empresa, realizan tareas en hojas de cálculo que solo ellos saben usar.
- Duplicación de Información: Reportes con un mismo fin pueden ser creados más de una vez por diferentes usuarios.
- Toma de Decisiones erróneas: Decisiones apuradas son tomadas sin haber hecho los cálculos correctos.

Actualmente en la empresa los procesos no están definidos, estos cambian constantemente y en algunos casos no tienen un orden lógico. Realizar un levantamiento de procesos, además de optimizar la gestión y forma en que se hacen las cosas ayudará a que las herramientas de Inteligencia de Negocios sean implementadas de forma más eficaz.

Esta problemática se debe a que los sistemas con los que cuenta la empresa, no fueron desarrollados con el fin de brindar síntesis, análisis, consolidación, búsquedas de datos en línea y en cualquier ubicación en que se encuentre el tomador de decisiones.

Descrita la realidad problemática de la investigación se formula lo siguiente:

**¿Cómo se podría mejorar el soporte online en la toma de decisiones para el área de ventas de la empresa Copy Ventas S.R.L., utilizando tecnologías de información?**

La hipótesis es:

**La implementación de una Solución de Inteligencia de Negocios en la nube mejorará el proceso de Toma de Decisiones en el área de Ventas para la empresa Copy Ventas S.R.L.**

El objetivo general es **Generar y gestionar el conocimiento a través de una solución de Inteligencia de Negocios en la nube para el proceso de Ventas de la empresa Copy Ventas S.R.L.**

Los objetivos específicos son los siguientes:

- ✓ Elaborar un marco conceptual respecto al trabajo evolutivo sobre la base de datos en la nube de Microsoft basada en soluciones de inteligencia de negocios.
- ✓ Definir los Stakeholders y obtener los requerimientos a través de entrevistas.
- ✓ Realizar el análisis de requerimientos para la elaboración de un Modelado de Datos que de soporte a la toma de decisiones en la organización.
- ✓ Implementar el modelo dimensional en Microsoft Azure
- ✓ Integrar las distintas fuentes de datos de la empresa en una solución ETL usando SQL Server 2014 y Microsoft Azure.
- ✓ Implementar los cubos y reportes dinámicos de acuerdo a las necesidades del cliente.

El presente proyecto está organizado en tres capítulos que facilitarán el uso y entendimiento del mismo dando a continuación una breve descripción del mismo:

**Capítulo I:** Fundamento teórico y metodología, en este capítulo damos el conocimiento formal de los temas y el modelo de referencia a utilizar para la solución del problema planteado. En esta parte damos los conocimientos básicos de que es un DataMart, Data Warehouse, Soporte de tomas de decisiones y Cubos OLAP.

También en este capítulo referimos la metodología que empleamos para el desarrollo del proyecto, presentando el Enfoque, Esquema y Herramientas que intervendrán en la misma.

**Capítulo II:** Desarrollo del trabajo, en éste capítulo muestra el desarrollo de los pasos enunciados en el Esquema de la Metodología. Mostramos los resultados obtenidos con relación a los objetivos planteados al inicio de este proyecto.

**Capítulo III:** En este capítulo se verifica si la Hipótesis es aceptada a esto se le llama Contrastación de la Hipótesis.

**Conclusiones y Recomendaciones,** en éste último capítulo se encuentra las conclusiones a las que se llegó después de haber culminado el proyecto y las recomendaciones derivadas de la experiencia de la misma.

Además en la parte final del documento se encuentran los anexos con información adicional, referidos a los capítulos enunciados.

Esperamos que este trabajo contribuya a una mejor comprensión de un Data Mart y sirva como guía de consulta para otros trabajos similares que se realicen posteriormente.

## CAPITULO I. FUNDAMENTO TEÓRICO Y METODOLOGÍA

### 1.1. FUNDAMENTO TEÓRICO

#### 1.1.1. INTEGRACION DE DATOS

Podemos definir la integración de datos como un proceso de transformación y conciliación de datos que permita una mayor agilidad en la gestión, proporcionando datos conectados, seguros y de calidad.

Integrar significa combinar datos que se encuentran en diferentes fuentes para permitirle al usuario final tener una vista unificada de los mismos para una accesibilidad idónea, que sirva a las necesidades de negocio.

Actualmente, la aparición de nuevas tecnologías y la explosión de datos plantean un gran desafío en este aspecto, y ello se traduce en una mayor complejidad técnica a la hora de implementar un plan de integración de datos, si bien el objetivo sigue siendo el mismo: evitar su fragmentación mediante el desarrollo de soluciones ad hoc.

Las tecnologías de integración de datos permiten asumir proyectos que impliquen la transferencia de datos, transformaciones complejas de datos, el acceso a fuentes de datos múltiples, sistemas heterogéneos, con tiempos de latencia apropiados (batch, tiempo real) y minimizando los riesgos más frecuentes vinculados a este tipo de proyectos, entre otros:

- Tiempos de desarrollo excesivos
- Costes de mantenimiento altos
- Dificultades a la hora de responder a las necesidades empresariales en continuo cambio

Los proyectos de Data Integration comprenden desde el aprovisionamiento de datos para proyectos de Business Intelligence (BI) migraciones que requieren la transformación de los modelos de datos; hasta la sincronización de bases de datos o la consolidación de sistemas. En la actualidad, las empresas también se enfrentan al desafío de integrar datos desestructurados (Big Data) o datos en la nube.

Cuando los sistemas y/o fuentes de datos que operan en una misma organización y compañía superan un número razonable, por ejemplo de 1 a 3, esta heterogeneidad les empieza a acarrear un considerable problema. Y esto ocurre a menudo en compañías de un cierto tamaño.

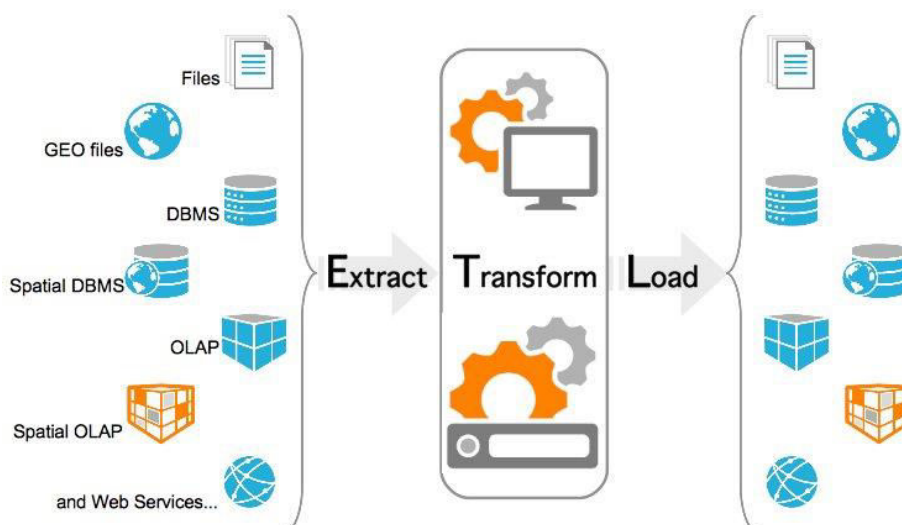
Llegados a este punto, muchas empresas empiezan a utilizar los **procesos ETL** con un objetivo distinto que el gerenciamiento o la optimización de datos. En estos casos, la utilidad de la ETL se diversifica hacia la integración y la migración de los datos.

### Más usos de los procesos ETL

- **Integración de datos.** Como hemos visto, los procesos ETL tienen la capacidad de leer y escribir en cualquier fuente o sistema. Por ello, no son pocas las empresas que recurren a este tipo de procesos para integrar datos de diversos sistemas o fuentes, muchos de ellos muy antiguos e incompatibles entre ellos. Si no existiera la posibilidad de utilizar los sistemas ETL como integradores de datos, no cabría otra opción de realizar esta tarea de forma manual y con multitud de errores.
- **Migración de datos.** Se trata de, únicamente, trasladar la información de una base de datos obsoleta a una nueva, realizando los cambios que sean precisos.

### Se trata de procesos menos comunes pero cada vez más utilizados

Aunque la aplicación más común para un proceso ETL es la construcción y carga de un *data warehouse*, cada vez más frecuentemente los procesos ETL son usados para operaciones tales como la integración y la migración de datos. **(PowerData, 2015).**



### 1.1.2. PROCESO DE TOMA DE DECISIONES

La Toma de Decisiones es un proceso por el cual se selecciona la mejor opción de entre muchas otras, este es un proceso que no solo se da en las empresas sino también en la vida cotidiana, o acaso al momento de seleccionar a tu proveedor ¿lo haces al azar?, o acaso al iniciar tu negocio ¿no tomaste alguna decisión que cambio tu vida?

Por lo tanto la toma de decisiones están en todo lugar, no solo en el mundo empresarial sino también en la vida cotidiana, para iniciar, cambiar o concluir algo, siempre tomamos antes una decisión, entonces la vida si es una Toma de Decisiones, pero:

¿Sabemos escoger bien nuestras decisiones?

¿Poseemos algún patrón para tomar nuestras decisiones?

¿El azar es parte de una decisión?

¿Es normal guiarse de la intuición para tomar una decisión?

Leamos entonces la importancia y el proceso de la Toma de Decisiones que suele un tema primordial en la administración de negocios.

#### ¿Qué es Toma de Decisiones?

La toma de decisiones es un proceso sistemático y racional a través del cual se selecciona una alternativa de entre varias, siendo la seleccionada la optimizadora (la mejor para nuestro propósito). **(Rodríguez-Gómez, 2011)**

Tomar una decisión es resolver diferentes situaciones de la vida en diferentes contextos: a nivel laboral, familiar, sentimental, empresarial, etc.

Tomar la correcta decisión en un negocio o empresa es parte fundamental del administrador ya que sus decisiones influirán en el funcionamiento de la organización, generando repercusiones positivas o negativas según su elección.



## ¿Qué debo saber antes, para tomar una buena decisión?

Antes de tomar una decisión debemos:

1. Definir las restricciones y limitaciones.
2. Saber la relación costo beneficio, rendimientos esperados u otros.
3. Saber cuándo se utilizan métodos cuantitativos y cuando los cualitativos.
4. Conocer los factores internos formales (cultura organizacional, políticas internas, estructura, etc.) y los factores internos informales (políticas implícitas, hábitos, experiencia, etc.)
5. Conocer los factores externos (políticos, económicos, sociales, internacionales, culturales)

Comprender los cinco puntos anteriores nos ayudara mucho al momento de tomar una buena decisión.

### **Etapas del Proceso de Toma de decisiones:**

- Identificar y analizar el problema: Un problema es la diferencia entre los resultados reales y los planeados, lo cual origina una disminución de rendimientos y productividad, impidiendo que se logren los objetivos.
- Investigación u obtención de información: Es la recopilación de toda la información necesaria para la adecuada toma de decisión; sin dicha información, el área de riesgo aumenta, porque la probabilidad de equivocarnos es mucho mayor.
- Determinación de parámetros: Se establecen suposiciones relativas al futuro y presente tales como: restricciones, efectos posibles, costos , variables, objetos por lograr, con el fin de definir las bases cualitativas y cuantitativas en relación con las cuales es posible aplicar un método y determinar diversas alternativas.
- Construcción de una alternativa: La solución de problemas puede lograrse mediante varias alternativas de solución; algunos autores consideran que este paso del proceso es la etapa de formulación de hipótesis; porque una alternativa de solución no es científica si se basa en la incertidumbre.

- Aplicación de la alternativa: De acuerdo con la importancia y el tipo de la decisión, la información y los recursos disponibles se eligen y aplican las técnicas, las herramientas o los métodos, ya sea cualitativo o cuantitativo, mas adecuados para plantear alternativas de decisión.
- Especificación y evaluación de las alternativas: Se desarrolla varias opciones o alternativas para resolver el problema, aplicando métodos ya sea cualitativos o cuantitativos. Una vez que se han identificado varias alternativas, se elige la optima con base en criterios de elección de acuerdo con el costo beneficio que resulte de cada opción. Los resultados de cada alternativa deben ser evaluados en relación con los resultados esperados y los efectos.
- Implantación: Una vez que se ha elegido la alternativa optima, se deberán planificarse todas las actividades para implantarla y efectuar un seguimiento de los resultados, lo cual requiere elaborar un plan con todos los elementos estudiados.

Lo presentado anteriormente fueron tan solo las etapas del proceso de toma de decisiones, desarrollarlas dependerá del tipo de problema que se quiera solucionar y del tipo de técnica que deba aplicar para solucionarlo.

En la toma de Decisiones existen también Técnicas Cuantitativas y Cualitativas para la selección de la mejor decisión.

- Técnicas Cualitativas: Cuando se basan en criterio de la experiencia, y habilidades
- Técnicas Cuantitativas: Cuando se utilizan métodos matemáticos, estadísticos, etc.

### **1.1.3. Tecnologías de Información**

Las Tecnologías de Información o TICs son aquellas herramientas computacionales e informáticas que procesan, almacenan, sintetizan, recuperan y presentan información representada de la más variada forma. Es un conjunto de herramientas, soportes y canales para el tratamiento y acceso a la información. Constituyen nuevos soportes y canales para dar forma, registrar, almacenar y difundir contenidos informacionales (Villalobos y Constenia, 2010).

#### 1.1.4. Sistema de Información

Es un conjunto formal de procesos que, operando sobre una colección de datos estructurada según las necesidades de la empresa, recopilan elaboran y distribuyen la información (o parte de ella) necesaria para las operaciones de dicha empresa y para las actividades de dirección y control correspondientes (decisiones) desempeñar su actividad de acuerdo a su estrategia de negocio.

Otra definición de sistemas de información es "Son procesar entradas, mantener archivos de datos relacionados con la organización y producir información, reportes y otras salidas".

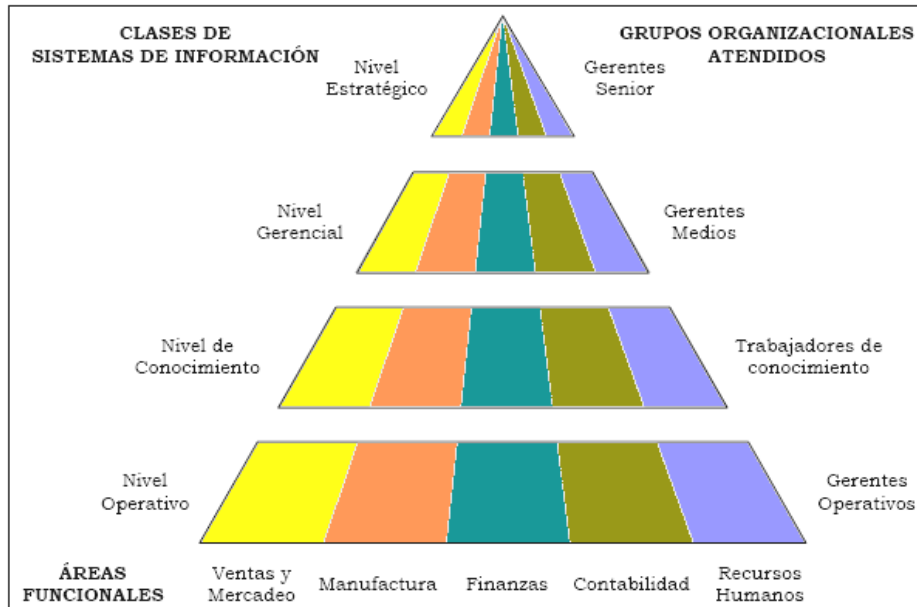
Por lo tanto podemos definir un sistema de información como un conjunto de subsistemas que incluyen hardware, software, medios de almacenamiento de datos ya sea primarios, secundarios y bases de datos relacionadas entre sí con el fin de procesar entradas para realizar transformaciones a esas entradas y convertirlas en salidas de información importantes en la toma de decisiones.

El objetivo de un sistema de información es ayudar al desempeño de las actividades que desarrolla la empresa, suministrando la información adecuada, con la calidad requerida, a la persona o departamento que lo solicita, en el momento y lugar especificados con el formato más útil para el receptor. **(Kendall y Kendall, 2014)**

##### **Tipos de Sistemas de Información**

En la Figura N° 1 muestra los tipos específicos de los sistemas de información que corresponden a cada uno de los niveles de la organización.

- Sistemas de procesamiento de transacciones (TPS)
- Sistemas de información administrativa (MIS)
- Sistemas de apoyo a las decisiones (DSS)
- Sistemas de apoyo a ejecutivos (ESS)



**Figura N°1: Tipos de Sistemas de Información**

**Fuente:** [Laudon y Laudon, 2013]

**a. Sistemas de Procesamiento de Transacciones**

Los sistemas de procesamiento de transacciones son sistemas de información encargados de procesar gran cantidad de transacciones rutinarias, es decir son todas aquellas que se realizan rutinariamente en la empresa entre estas tenemos el pago de nómina, facturación, entrega de mercancía y depósito de cheques. Estas transacciones varían de acuerdo al tipo de empresa.

Los sistemas de procesamiento de transacción o TPS (Transaction Procesation System) por sus siglas en inglés, eliminan el trabajo tedioso de las transacciones operacionales y como resultado reducen el tiempo que se empleaba en ejecutarlas actualmente, aunque los usuarios todavía deben alimentar de datos a los TPS.

"Los sistemas de procesamiento de transacciones son sistemas que traspasan sistemas y que permiten que la organización interactúe con ambientes externos. Debido a que los administradores consultan los datos generados por el TPS para información al minuto acerca de lo que está pasando en sus compañías, es esencial para las operaciones diarias que

estos sistemas funcionen lentamente y sin interrupción". (**Laudon y Laudon, 2013**)

#### **b. Sistemas de información de Gerencial**

Los sistemas de información gerencial (MIS por sus siglas en inglés) no reemplazan a los sistemas de procesamiento de transacciones ni tampoco son los mismos, sino que estos sistemas incluyen procesamiento de transacciones. Los sistemas de información gerencial son sistemas de información computarizada que trabajan con la interacción entre usuarios y computadoras. Requieren que los usuarios, el software (programas de computadora) y el hardware (computadoras, impresoras, etc.) trabajen a un mismo ritmo.

Los sistemas de información gerencial dan soporte a un aspecto más amplio de tareas organizacionales, a comparación de los sistemas de procesamiento de transacciones, los sistemas de información gerencial incluyen el análisis de decisiones y la toma de decisiones.

"Para poder ligar la información, los usuarios de un sistema de información gerencial comparten una base de datos común. La base de datos guarda modelos que ayudan a los usuarios a interpretar y aplicar esos mismos datos.

Los sistemas de información gerencial producen información que es usada en la toma de decisiones. Un sistema de información gerencial también puede llegar a unificar algunas de las funciones de información computarizada, aunque no exista como una estructura singular en ningún lugar del negocio". (**Laudon y Laudon, 2013**)

#### **c. Sistemas de Apoyo a Decisiones**

Los sistemas de apoyo a decisiones (DSS) ayudan a los gestores a tomar decisiones que son únicas, que cambia rápidamente, y no es fácil definirse de antemano. Ellos tratan de problemas que el procedimiento para llegar a una solución no puede ser plenamente predefinidos de antemano. Aunque el Departamento de Servicios Sociales uso de información interna de TPS y MIS, que a menudo traen consigo la información de fuentes externas, tales como los

precios de las acciones o los precios de los productos de los competidores. Es evidente que, de diseño, DSS tienen más poder analítico que otros sistemas. Utilizan una gran variedad de modelos para analizar los datos, o se condensan grandes cantidades de datos en un formulario en el que puedan ser analizados por los encargados de adoptar decisiones. DSS están diseñados para que los usuarios puedan trabajar con ellos directamente, estos sistemas incluyen explícitamente el software de uso fácil. DSS son interactivos, el usuario puede cambiar las hipótesis, pedir nuevas preguntas, e incluir nuevos datos. **(Laudon y Laudon, 2013)**

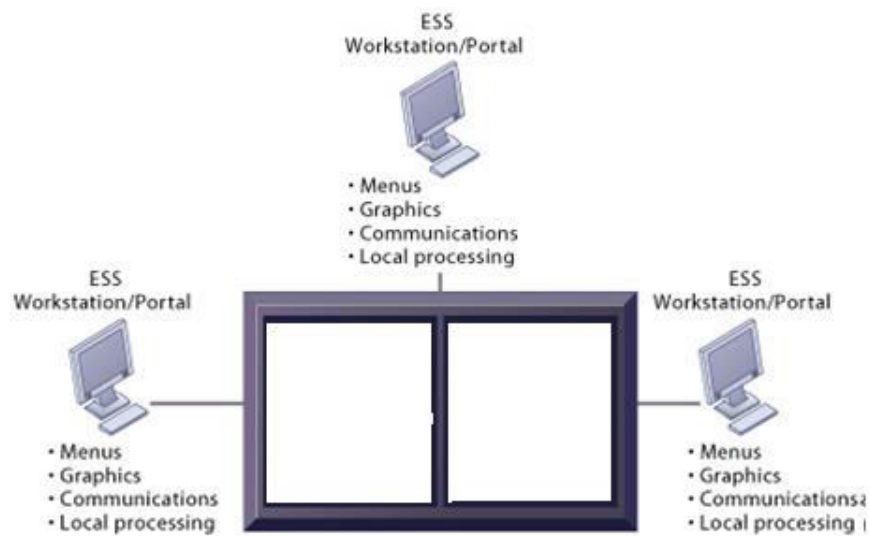
#### **d. Sistemas de Apoyo a Ejecutivos**

Los Sistemas de Apoyo a Ejecutivos (ESS) son utilizados en el nivel estratégico de la organización. Los ESS no solo están diseñados para incorporar información sobre eventos externos, como las nuevas leyes fiscales o de los competidores, sino que también sacar un resumen de la información interna de los sistemas MIS y DSS. Estos sistemas pueden filtrar, comprimir, y realizar un seguimiento de datos críticos, mostrando los datos de la mayor importancia para los altos ejecutivos. Por ejemplo, el CEO de Leiner Health Products, el mayor fabricante de etiqueta privada vitaminas y suplementos en los Estados Unidos, tiene un ESS que ofrece en su escritorio, minuto a minuto la vista de la empresa: estados financieros, medido por capital de trabajo, cuentas por cobrar, cuentas por pagar, flujo de caja, e inventario.

Los Sistemas de Apoyo a Ejecutivos, emplean los más avanzados software de gráficos y puede presentar gráficos y datos de muchas fuentes. A menudo la información se entrega a los altos ejecutivos a través de un portal, que utiliza una interfaz Web integrada para presentar contenido personalizado de negocio de una variedad de fuentes.

A diferencia de los otros tipos de sistemas de información, los ESS no están diseñados principalmente para resolver problemas específicos. Aunque muchos DSS están diseñados para ser muy analítico, los ESS tienden a hacer menos uso de modelos analíticos. Los Sistemas de Apoyo a Ejecutivos ayudan a responder a las siguientes preguntas: ¿En qué negocios deberíamos estar? ¿Cuáles son

nuestros competidores? ¿Qué nuevas adquisiciones deberíamos realizar? ¿Qué unidades debemos vender más para recaudar más utilidades?



**Figura N° 2: Modelo de un ESS**

**Fuente: (Laudon y Laudon, 2013)**

Figura N° 2 ilustra un modelo de un ESS. Se compone de los puestos de trabajo con menús, gráficos interactivos, y con la capacidad de comunicación que pueden utilizarse para el acceso histórico de los datos corporativos internos y los sistemas de bases de datos externas. Porque ESS están diseñados para ser utilizados por altos directivos que a menudo tienen poca información, en su caso, el contacto directo con su computadora y la experiencia, basado en los sistemas de información, que sea de fácil uso, con interfaces gráficas. **(Laudon y Laudon, 2013)**

### 1.1.5. Procesamiento Transaccional en Línea (OLTP)

Procesamiento Transaccional en Línea (OnLine Transactional Processing, en inglés), tiene como objetivo mantener la integridad de la información (relaciones entre los datos) necesaria para operar un negocio de la manera más eficiente. Sin embargo, este modelo no corresponde a la forma como el usuario percibe la operación de un negocio. El OLTP se basa en un modelo relacional.

### 1.1.6. Datawarehouse

El DataWarehouse (DW) convierte entonces los datos operacionales de una organización en una herramienta competitiva, por hacerlos disponibles a los empleados que lo necesiten para el análisis y toma de decisiones.

El objetivo del Datawarehouse es el de satisfacer los requerimientos de información interna de la empresa para una mejor gestión. El contenido de los datos, la organización y estructura son dirigidos a satisfacer las necesidades de información de los analistas y usuarios tomadores de decisiones. El DW es el lugar donde la gente puede acceder a sus datos.

El DW puede verse como una bodega donde están almacenados todos los datos necesarios para realizar las funciones de gestión de la empresa, de manera que puedan utilizarse fácilmente según se necesiten.

Los Datawarehouse (almacenes de datos) generan bases de datos tangibles con una perspectiva histórica, utilizando datos de múltiples fuentes que se fusionan en forma congruente. Estos datos se mantienen actualizados, pero no cambian al ritmo de los sistemas transaccionales. Muchos datawarehouses se diseñan para contener un nivel de detalle hasta el nivel de transacción, con la intención de hacer disponible todo tipo de datos y características, para reportar y analizar. Así un datawarehouse resulta ser un recipiente de datos transaccionales para proporcionar consultas operativas, y la información para poder llevar a cabo análisis multidimensional. De esta forma, dentro de un datawarehouse existen dos tecnologías que se pueden ver como complementarias, una relacional para consultas y una multidimensional para análisis (Kimball y Ross., 2013).

DW está basado en un procesamiento distinto al utilizado por los sistemas operacionales, es decir, este se basa en OLAP -Procesos de Análisis en Línea- (OnLine Analysis Process, en ingles), usado en el análisis de negocios y otras aplicaciones que requieren una visión flexible del negocio.

Para ampliar los conceptos anteriores, en la tabla N°1 se exponen las principales diferencias entre los sistemas Transaccionales (OLTP) y los basados en Datawarehouses.

**Tabla N° 1:** Diferencia entre Sistemas Transaccionales y Datawarehouse



Transaccionales	Basados en Datawarehouse
Admiten el acceso simultáneo de muchos usuarios -miles- que agregan y modifican datos.	Admiten el acceso simultáneo de muchos usuarios -cientos- que consultan y no modifican datos
Representan el estado, en cambio constante, de una organización, pero no guardan su historial.	Guardan el historial de una organización
Contienen grandes cantidades de datos, incluidos los datos extensivos utilizados para comprobar transacciones.	Contienen grandes cantidades de datos, resumizados, consolidados y transformados. También de detalle pero solo los necesarios para el análisis.
Tienen estructuras de base de datos complejas.	Tienen estructuras de Base de datos simples.
Se ajustan para dar respuesta a la actividad transaccional.	Se ajustan para dar respuesta a la actividad de consultas.
Proporcionan la infraestructura tecnológica necesaria para admitir las operaciones diarias de la empresa.	Proporcionan la infraestructura tecnológica necesaria para admitir análisis de los datos de la empresa.
Los analistas carecen de la experiencia técnica necesaria para crear consultas "ad hoc" contra la compleja estructura de datos.	Pueden combinar datos de orígenes heterogéneos en una única estructura homogénea y simple, facilitando la creación de informes y consultas.
Las consultas analíticas que resumen grandes volúmenes de datos afectan negativamente a la capacidad del sistema para responder a las transacciones en línea.	Organizan los datos en estructuras simplificadas buscando la eficiencia de las consultas analíticas más que del proceso de transacciones.
Los datos que se modifican con frecuencia interfieren en la coherencia de la información analítica.	Proporcionan datos estables que representan el historial de la empresa. Se actualizan periódicamente con datos adicionales, no con transacciones frecuentes.

**1.1.7. Introducción a Datamarts**

El acceso a los datos de toda la empresa a veces no es conveniente (o necesario) para determinados usuarios que solo necesitan un subconjunto de estos datos, en estos casos se utilizan los Datamarts. El concepto Datamart es una especialización del datawarehouse, y está enfocado a un departamento o área específica, como por ejemplo los departamentos de Finanzas o Marketing. Permitiendo así un mejor control de la información que se está abarcando **(Kimball y Ross., 2013)**

Los principales beneficios de utilizar Datamarts son:

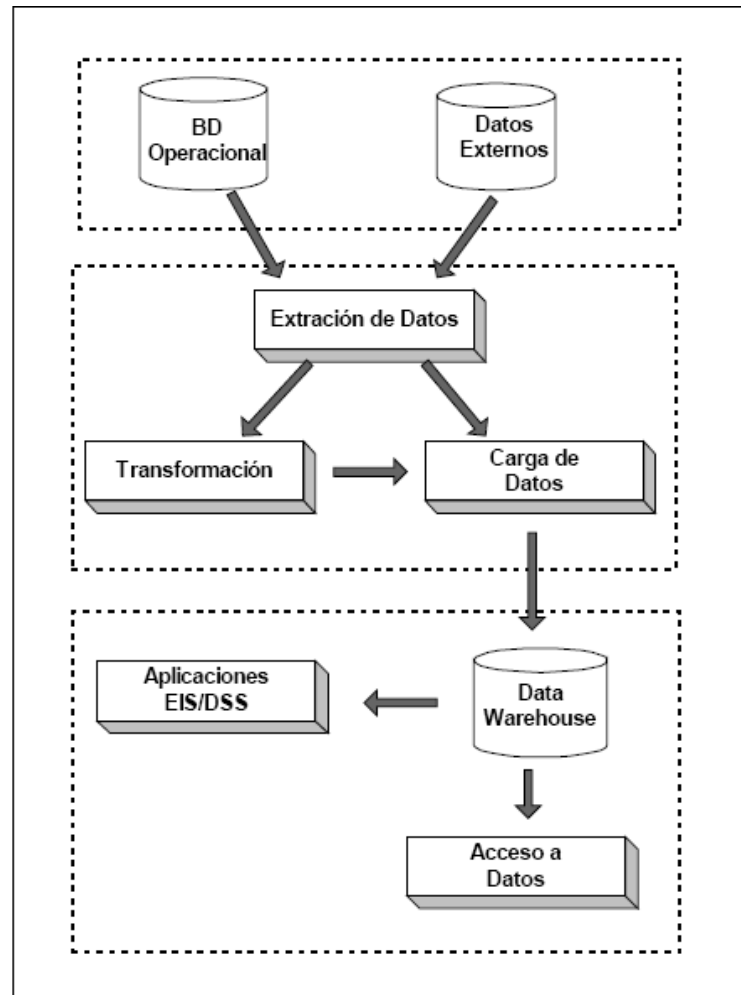
- Acelerar las consultas reduciendo el volumen de datos a recorrer
- Estructurar los datos para su adecuado acceso por una herramienta
- Dividir los datos para imponer estrategias de control de acceso
- Segmentar los datos en diferentes plataformas hardware
- Permite el acceso a los datos por medio de un gran número de herramientas del mercado, logrando independencia de estas.

#### **1.1.8. Arquitectura del Datawarehouse**

La estructura básica de la arquitectura Datawarehouse incluye:

- Datos operacionales: un origen o fuente de datos para poblar el componente de almacenamiento físico DW. El origen de los datos son los sistemas transaccionales internos de la organización como también datos externos a ésta.
- Extracción de Datos: selección sistemática de datos operacionales usados para poblar el componente de almacenamiento físico DW.
- Transformación de datos: procesos para sumarizar y realizar otros cambios en los datos operacionales para reunir los objetivos de orientación a temas e integración principalmente.
- Carga de Datos: inserción sistemática de datos en el componente de almacenamiento físico DW.
- Datawarehouse: almacenamiento físico de datos de la arquitectura DW.
- Herramientas de Acceso al componente de almacenamiento físico DW: herramientas que proveen acceso a los datos. Estas herramientas pueden ser herramientas específicas de mercado para visualización de bases

multidimensionales almacenadas en datawarehouses como también aplicaciones desarrolladas dentro de la organización del tipo EIS/DSS (Kimball y Ross., 2013).



**Figura N° 3:** Estructura básica Datawarehouse.

**Fuente:** (Kimball y Ross., 2013)

### 1.1.9. Introducción al Procesamiento Analítico en Línea

La tecnología de Procesamiento Analítico en Línea –OLAP- (Online Analytical Processing) permite un uso más eficaz de los datawarehouses para el análisis de datos en línea, lo que proporciona respuestas rápidas a consultas analíticas complejas e iterativas utilizada generalmente para sistemas de ayuda para la toma de decisiones. Primero y más importante, el OLAP presenta los datos a los usuarios a través de un modelo de datos intuitivo y natural. Con este estilo de navegación, los usuarios finales pueden ver y

entender más efectivamente la información de sus bases de datos, permitiendo así a las organizaciones reconocer mejor el valor de sus datos.

Las aplicaciones OLAP deberían proporcionar análisis rápidos de información multidimensional compartida. Las características principales del OLAP son:

- **Rápido:** proporciona la información al usuario a una velocidad constante. La mayoría de las peticiones se deben de responder al usuario en cinco segundos o menos.
- **Análisis:** realiza análisis estadísticos y numéricos básicos de los datos, predefinidos por el desarrollador de la aplicación o definido “ad hoc” por el usuario.
- **Compartida:** implementa los requerimientos de seguridad necesarios para compartir datos potencialmente confidenciales a través de una gran población de usuarios.
- **Multidimensional:** llena la característica esencial del OLAP, que es ver la información en determinadas vistas o dimensiones.
- **Información:** acceden a todos los datos y a la información necesaria y relevante para la aplicación, donde sea que ésta resida y no esté limitada por el volumen.

El OLAP es un componente clave en el proceso de almacenamiento de datos (data warehousing) y los servicios OLAP proporcionan la funcionalidad esencial para una gran variedad de aplicaciones que van desde reportes corporativos hasta soporte avanzado de decisiones.

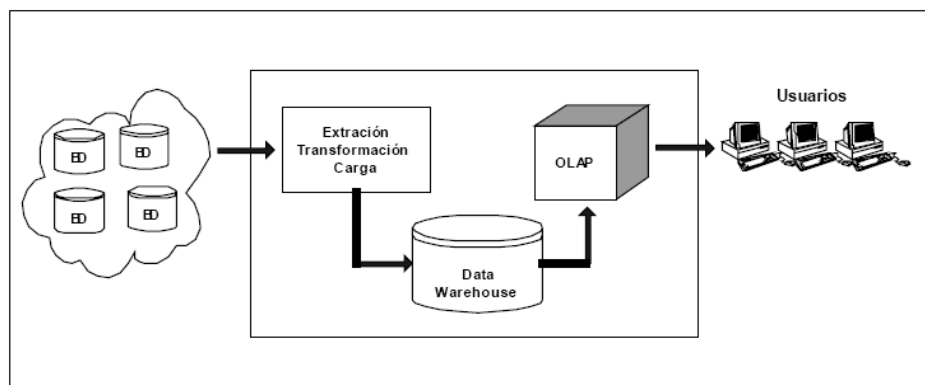
Dentro de cada dimensión de un modelo de datos OLAP, los datos se pueden organizar en una jerarquía que represente niveles de detalle de los datos. Por ejemplo, dentro de la dimensión de tiempo, se puede tener estos niveles: años, meses y días; de manera similar, dentro de la dimensión geografía, Se puede tener estos niveles: país, región, estado/provincia y ciudad. Una instancia particular del modelo de datos OLAP tendrá valores para cada nivel en la jerarquía. Un usuario que vea datos OLAP se moverá entre estos niveles para ver información con mayor o menor detalle. **(Kimball y Ross., 2013).**

#### **1.1.10. Arquitectura OLAP**

La tecnología OLAP permite un uso más eficaz de los almacenes de datos para el análisis en línea, lo que proporciona respuestas rápidas a consultas analíticas complejas e iterativas.

Los modelos de datos multidimensionales de OLAP y las técnicas de agregados de datos organizan y resumen grandes cantidades de datos para que puedan ser evaluados con rapidez mediante el análisis en línea y las herramientas gráficas. La respuesta a una consulta realizada sobre datos históricos a menudo suele conducir a consultas posteriores en las que el analista busca respuestas más concretas o explora posibilidades. Los sistemas OLAP proporcionan la velocidad y la flexibilidad necesarias para dar apoyo al analista en tiempo real.

La figura N° 5, muestra la integración del datawarehouse y los procesos OLAP, que generalmente se implementan por medio de una aplicación servidora que accede al datawarehouse y realiza los procesos de análisis. A través de este servicio OLAP, los usuarios acceden a la información residente en las bases de datos **(Kimball y Ross., 2013)**



**Figura N° 4:** Arquitectura básica para OLAP

**Fuente (Kimball y Ross., 2013)**

#### **1.1.11. El modelo de datos OLAP.**

En la mayoría de las implementaciones de OLAP, se asume que los datos han sido preparados para el análisis a través del almacenamiento de datos (data warehouse) y que la información se ha extraído de sistemas operacionales, limpiado, validado y resumido antes de incorporarse en una aplicación OLAP.

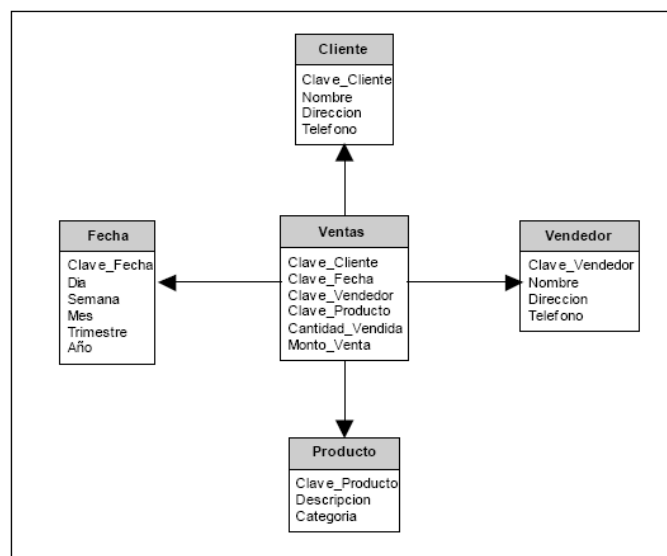
Este es un paso vital en el proceso, que asegura que los datos que son vistos por el usuario OLAP son correctos, consistentes y que llenan las definiciones organizacionales para los datos.

Cada vez más, la información en un datawarehouse se organiza en esquemas de estrella o de copo de nieve.

### a. Esquema Estrella

El esquema estrella se basa en una tabla de hechos central (las medidas) que se enlaza a las tablas de dimensiones relacionadas (las categorías descriptivas de las medidas), mientras que el esquema copo de nieve, una tabla de hechos central se enlaza a las tablas de dimensiones relacionadas, pero estas a su vez se enlaza a otras tablas dimensionales. Con este tipo de esquemas simplifica el entendimiento de los datos por parte del usuario, maximiza el desempeño de las peticiones de la base de datos para aplicaciones de soporte de decisiones y requiere menor espacio de almacenamiento para bases de datos grandes

La figura N° 6 muestra un ejemplo de esquema de estrella. En este tipo de base de datos, una tabla de hechos central se enlaza a las tablas de dimensiones relacionadas.



**Figura N°5:** Un esquema de estrella.

**Fuente (Kimball y Ross., 2013)**

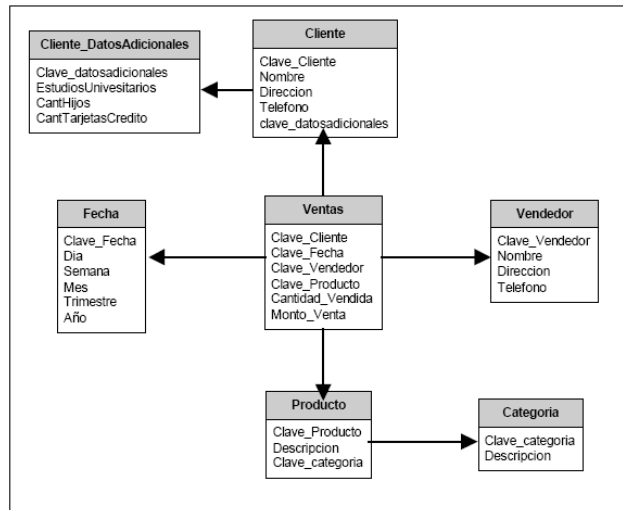
A continuación se enumeran algunas de las principales ventajas del esquema estrella.

- Crea una base de datos con tiempos de respuesta rápido.
- Diseño fácil de modificar.
- Simula como ven los datos los usuarios finales.
- Simplifica la navegación.
- Facilita la interacción con herramientas.

### **b. Esquema Copo de Nieve**

El esquema copo de nieve es una extensión del esquema estrella en donde cada uno de los puntos de la estrella se divide en más puntos. En esta forma de esquema, las tablas de dimensión del esquema estrella contienen más normas. Las ventajas que proporciona es esquema de copo de nieve son mejorar el desempeño de consultas debido aun mínimo almacenamiento en disco para los datos y mejorar el desempeño mediante la unión de tablas más pequeñas con normas. Así mismo el esquema copo de nieve incrementa la flexibilidad de las aplicaciones debido a la aplicación de normas y por lo tanto disminuye la granularidad de las dimensiones. **(Kimball y Ross., 2013)**

La figura N° 7, muestra un ejemplo de esquema copo de nieve. Este tipo de esquema se caracteriza por tener tablas dimensionales relacionadas con otras tablas dimensionales además de vincularse a la tabla de hechos



**Figura N° 6:** Un esquema de copo de nieve.

**Fuente (Kimball y Ross., 2013).**

Los esquemas de estrella y copo de nieve son aproximaciones relacionales del modelo de datos OLAP y son un punto de partida excelente para construir definiciones de cubo OLAP. Pocos productos OLAP han tomado ventaja de este hecho. Generalmente no han provisto herramientas sencillas para mapear un esquema de estrella a un modelo OLAP y como resultado mantienen el costo de construir el modelo OLAP extremadamente alto y el tiempo de desarrollo innecesariamente largo.

### 1.1.12. Implementación del OLAP

Los cubos, las dimensiones y las jerarquías son la esencia de la navegación multidimensional del OLAP. Al describir y representar la información en esta forma, los usuarios pueden navegar intuitivamente en un conjunto complejo de datos. Sin embargo, el solo describir el modelo de datos en una forma más intuitiva, hace muy poco para ayudar a entregar la información al usuario más rápidamente.

Un principio clave del OLAP es que los usuarios deberían de ver tiempos de respuesta consistentes para cada vista de datos que requieran. Dado que la información se colecta en el nivel de detalle solamente, el resumen de la



información es usualmente calculado por adelantado. Estos valores precalculados, son la base de las ganancias de desempeño del OLAP.

En los primeros días de la tecnología OLAP, la mayoría de las compañías asumía que la única solución para una aplicación OLAP era un modelo de almacenamiento no relacional. Después, otras compañías descubrieron que a través del uso de estructuras de base de datos (esquemas de estrella y de copo de nieve), índices y el almacenamiento de agregados, se podrían utilizar sistemas de administración de bases de datos relacionales (RDBMS) para el OLAP.

Estos vendedores llamaron a esta tecnología OLAP relacional (ROLAP). Las primeras compañías adoptaron entonces el término OLAP multidimensional (MOLAP), estos conceptos, MOLAP y ROLAP,

Un desarrollo reciente ha sido la solución OLAP híbrida (HOLAP), la cual combina las arquitecturas ROLAP y MOLAP para brindar una solución con las mejores características de ambas: desempeño superior y gran escalabilidad. Un tipo de HOLAP mantiene los registros de detalle (los volúmenes más grandes) en la base de datos relacional, mientras que mantiene las agregaciones en un almacén MOLAP separado.

#### **a. Sistemas MOLAP**

MOLAP (OLAP Multidimensional) una copia de los datos de origen del cubo, junto con sus agregaciones, es almacenada en una estructura multidimensional.

Debemos tener en cuenta que mientras los datos de origen cambian directamente con las operaciones, los objetos con almacenamiento MOLAP deben ser procesados para incorporar estos cambios.

El tiempo comprendido entre un procesamiento y el siguiente, crea un periodo de latencia durante el que puede que la información OLAP no coincida con los datos de origen actuales.

Como característica del almacenamiento MOLAP podemos destacar:

- Provee excelente rendimiento y compresión de datos.

- Tiene mejor tiempo de respuesta, dependiendo solo del porcentaje de las agregaciones del cubo.
- La estructura está muy optimizada para maximizar el rendimiento de las consultas.
- En general este método, es muy apropiado para cubos con uso frecuente por su rápida respuesta.

#### **b. Sistemas ROLAP**

En un modelo ROLAP (OLAP Relacional) toda la información del cubo, sus datos, su agregación, sumas, etc., son almacenados en una base de datos relacional.

A diferencia del modo de almacenamiento MOLAP, ROLAP no almacena copia de la base de datos, accede a las tablas originales cuando necesita responder a las consultas, generalmente es mucho más lenta que las otras estrategias de almacenamiento (MOLAP o HOLAP).

ROLAP se utiliza para ahorrar espacio de almacenamiento cuando se trabaja con grandes conjuntos de datos que se consultan con poca frecuencia; por ejemplo, datos exclusivamente históricos.

#### **c. Sistemas HOLAP**

HOLAP (OLAP híbrido) combina atributos de MOLAP y ROLAP.

Al igual que MOLAP, HOLAP hace que las agregaciones se almacenen en una estructura multidimensional, y los datos a nivel de detalle, en una base de datos relacional como lo hace el almacenamiento ROLAP.

Para procedimientos de búsqueda que accedan datos sumariados, HOLAP es equivalente a MOLAP. Por el contrario, si los procesos de consultas accedieran a los máximos niveles de detalle, deberían recuperar los datos de la base de datos relacional y esto no sería tan rápido comparado con una estructura MOLAP.

Los cubos almacenados como HOLAP, son más pequeños que los MOLAP y responden más rápidos que los ROLAP.

Usos comunes de HOLAP

- Cubos que requieren rápida respuesta
- Cuando existen sumalizaciones basadas en una gran cantidad de datos de origen.
- Solución de compromiso para bajar el espacio ocupado sin perjudicar totalmente el rendimiento de las consultas. **(Root y Mason , 2014)**

### 1.1.13. Cubos OLAP

Un cubo es una multidimensional estructura que almacena los datos de tu sistema OLAP. Multidimensional significa que los cubos te permiten mirar tu data en varios caminos o de varias maneras.

El cubo es un elemento clave en el proceso analítico en línea OLAP y consta de Medidas (o datos cuantitativos como ventas o costos) y dimensiones.

El Administrador de OLAP le permitirá convertir los datos almacenados en bases de datos relacionales en información empresarial significativa y fácil de explorar con sólo crear un cubo de datos. Los conceptos y terminología asociados con los cubos se describen en las siguientes pantallas. **(Root y Mason , 2014)**.

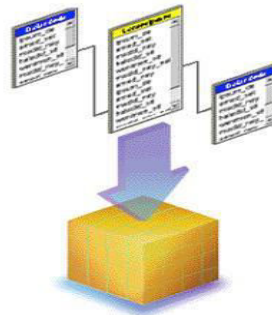


**Figura N° 7:** Cubos OLAP.

**Fuente:** (Root y Mason , 2014)

## ESQUEMAS RELACIONALES Y CUBOS

La manera más común de administrar datos relacionales para su empleo multidimensional es un esquema de estrella. Un esquema de estrella consiste en una única tabla de hechos que se combina con varias tablas de dimensiones. La tabla de hechos contiene los datos numéricos que se corresponden con las medidas de un cubo. Las columnas de la tabla de dimensiones, tal como implica su nombre, asignan los niveles jerárquicos de una dimensión. **(Root y Mason , 2014).**



**Figura N° 8: Esquemas Relacionales y cubo**

**Fuente:** (Root y Mason , 2014)

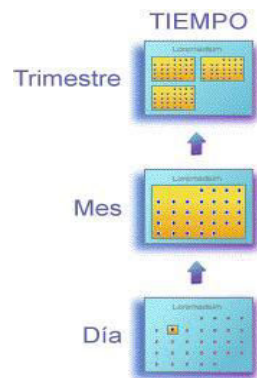
## **DIMENSIONES DE UN CUBO**

Las dimensiones de un cubo representan distintas categorías para analizar datos empresariales. Categorías tales como fecha, geografía o línea de productos son dimensiones típicas de cubos.

Nota: los cubos no están limitados a tres dimensiones. Pueden contener hasta 64 dimensiones. (Root y Mason , 2014)

## **DIMENSIONES Y JERARQUÍAS**

Las dimensiones se suelen organizar en jerarquías de información que se asignan a columnas en una base de datos relacional. Las jerarquías de dimensiones están agrupadas en niveles que constan de los miembros de una dimensión. Podrá unir los niveles de una dimensión para formar los valores de los que constará el siguiente nivel superior. Por ejemplo, en una dimensión temporal, los días se unen en meses y los meses forman trimestres.



**Figura N° 9:** Jerarquías de los cubos OLAP

**Fuente:** (Root y Mason , 2014)

### MEDIDAS DE UN CUBO

Las medidas son los valores cuantitativos contenidos en la base de datos que desea analizar. Las medidas típicas son ventas, costo y datos presupuestarios. Las medidas se analizan contra las distintas categorías de dimensiones de un cubo. Por ejemplo, tal vez desee analizar datos de ventas y de presupuestos (sus medidas) para un determinado producto (una dimensión) correspondientes a varios países (niveles específicos de una dimensión geográfica) durante dos años concretos (niveles de una dimensión temporal).



**Figura N° 11:** Jerarquías de los cubos OLAP

**Fuente:** (Root y Mason , 2014)

#### 1.1.14. Agregados

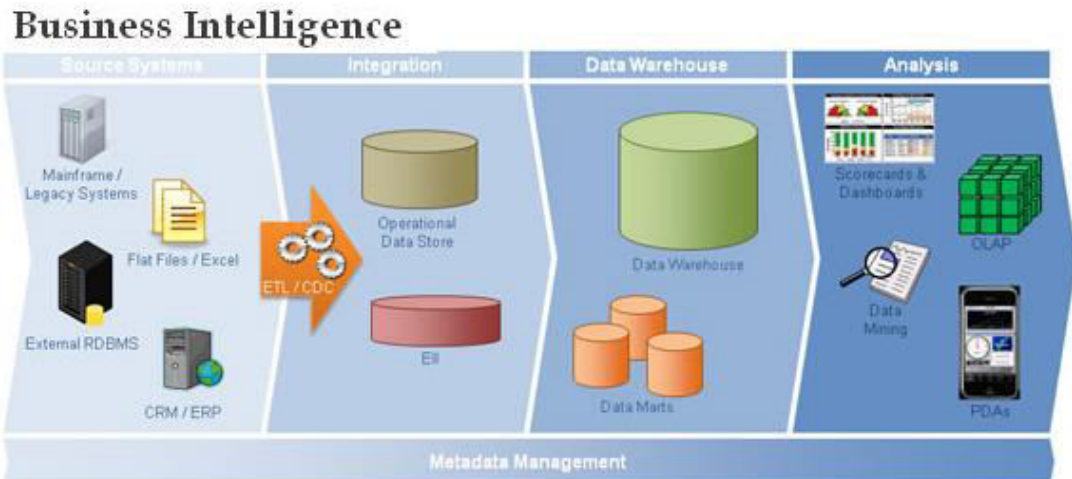
Los agregados son resúmenes de datos precalculados que mejoran el tiempo de respuesta a las consultas por el simple hecho de tener preparadas las respuestas antes de que se planteen las preguntas. Por ejemplo, la respuesta a una consulta que solicita el total de ventas semanales de una determinada línea de productos y que se realiza en una tabla de hechos de un almacén de datos que contiene cientos de miles de filas de información, puede llevar mucho tiempo si hay que explorar la tabla de hechos para calcular la respuesta. Por el contrario, la respuesta podría ser casi inmediata si los datos de resumen para la respuesta a esta consulta se han calculado previamente. El cálculo previo de los datos de resumen es la clave para obtener respuestas rápidas en la tecnología OLAP.

Los cubos son la forma en que la tecnología OLAP organiza los datos de resumen en estructuras multidimensionales. Las dimensiones y sus niveles jerárquicos reflejan las consultas que se pueden hacer al cubo. Los agregados se almacenan en la estructura multidimensional en celdas que tienen las coordenadas especificadas por las dimensiones. (Root y Mason , 2014)

### 1.1.15. **Business Intelligence (BI)**

Inteligencia de Negocios (Business Intelligence – BI) es una disciplina que, junto con sus correspondientes herramientas, hacen centro en el análisis de la información para la correcta toma de decisiones que le permita a la organización cumplir con los objetivos de negocio.

BI es un término "agrupador". El que sea considerado como un conjunto de conceptos le da un poder enorme, pues pueden integrarse funciones que tradicionalmente estaban separadas, tales como el acceso de datos, reportes, explotación, pronóstico y análisis. De ese modo, al menos en la actualidad en empresas grandes, BI se ha convertido en un apoyo indispensable para la Toma de Decisiones, en cualquier nivel de la organización y mucha gente está explotando el potencial estratégico de los datos operativos. Bien utilizada, BI puede ser un arma estratégica de la gente de negocios, sustentada en tecnología de sistemas (Kimball y Ross., 2013).



**Figura N° 12: Business Intelligence**

**Fuente: (Kimball y Ross., 2013)**

### 1.1.16. Microsoft SQL Server 2014

Con la aparición de **SQL Server 2014** el mundo de las Bases de datos está cambiando. Los desarrolladores ahora pueden ubicar su código apropiadamente en relación a su funcionalidad, acceder a datos nativos como XML, y construir sistemas complejos que sean manejados por el servidor de Bases de Datos. Estos puntos hacen que el desarrollo de Bases de Datos esté encaminado hacia una integración.

**SQL Server 2014** es más que un sistema gestor de Bases de Datos ya que incluye múltiples componentes y servicios que la convierten en una plataforma de aplicaciones corporativas.

#### **Aportes de SQL2014**

Entre todas las novedades que aporta la nueva versión de SQL Server 2014 hablaremos de las que encontramos en la vertiente de desarrolladores. En este apartado podemos destacar las siguientes:

- **Integración entre SQL Server y Common Language Runtime (CLR)**

Esta nueva característica hace que los desarrolladores de Bases de Datos puedan aprovechar las ventajas de la biblioteca de clases de Microsoft .NET Framework para implementar funcionalidades de servidor. Además se pueden codificar procedimientos almacenados, funciones y triggers en un lenguaje .NET Framework.

- **Notificaciones SQL (Service Brokers)**

Permite enviar notificaciones a los suscriptores cuando se produce un evento. Ésta funcionalidad es muy útil para invalidar cachés o vistas. Las notificaciones son generadas de forma eficiente i enviadas a múltiples tipos de dispositivos.

- **Múltiples resulsets y transacciones simultaneas para conexión**

Una novedad importante es la de permitir múltiples resulsets abiertos para conexión, el que permite consultar diversos grupos de resultados sin tener que a consultar al servidor y sin tener que abrir y cerrar conexiones.

Relacionado con este tema encontramos el nuevo modelo transaccional que ofrece, muy similar al de *COM+*, pero sin *COM+*. El funcionamiento es sencillo, cuando abres una *TransactionScope* toda la conexión que se abra dentro de el será asociada a la misma transacción.

- **Mejoras de seguridad**

Podemos agrupar las mejoras en 3 grandes áreas:

1. Restricciones de acceso de usuarios al servidor



2. Deshabilitar servicios y restricción de configuraciones de servicios.
3. Reducción de la superficie de ataques para las nuevas características.

Tres nuevos niveles de seguridad: *Safe*, *External* i *Unsafe*

- **Nuevo tratamiento de errores**

Nuevas funcionalidades como *ERROR\_NUMBER*, *ERROR\_MESSAGE*, *ETC*

- **Nuevo sistema de bloqueo (Snapshot Isolation)**

Nuevo nivel de aislamiento que hace que no haya bloqueos en las lecturas, ya que hace que los lectores lean una versión anterior de los registros.

Tiene la ventaja de una lectura más rápida y sin bloqueos, pero tiene el inconveniente de unas escrituras más lentas y un tamaño de Base de Datos más grande.

- **Soporte Nativo de XML (XQUERY)**

Da soporte por utilizar un tipo de datos XML para almacenar, validar y consultar información en forma XML.

### 1.1.17. Analysis Services

Analysis Services de Microsoft® SQL Server es un servidor de nivel intermedio para procesos analíticos en línea (OLAP) y minería de datos. El sistema Analysis Services incluye un servidor que administra cubos de datos multidimensionales para analizarlos y proporciona un rápido acceso cliente a la información de los cubos. Analysis Services organiza los datos de un almacenamiento de datos en cubos con datos de agregación precalculados para proporcionar respuestas

rápidas a consultas analíticas complejas. Analysis Services permite también crear modelos de minería de datos de orígenes de datos multidimensionales (OLAP) y relacionales. Puede aplicar modelos de minería de datos a ambos tipos de datos. Microsoft Excel y aplicaciones de otros fabricantes utilizan el servicio PivotTable, el proveedor compatible con OLE DB que se incluye, para recuperar datos del servidor y presentarlos al usuario o para crear cubos de datos locales para el análisis sin conexión.

Las características claves del Analysis Services son:

- **Facilidad de uso:** Extensa interfaz de usuario con asistentes
- **Modelo flexible de datos:** Un modelo flexible y eficaz de datos para la definición y almacenamiento de cubos
- **Escalabilidad:** Arquitectura escalable que proporciona una gran variedad de escenarios de almacenamiento y una solución automatizada para el síndrome de explosión de datos que existe en las tecnologías OLAP tradicionales.
- **Integración:** Integración de herramientas de administración, seguridad, orígenes de datos y caché de cliente-servidor.
- **API:** ampliamente compatibles y arquitectura abierta Compatibilidad con aplicaciones personalizadas

### Opciones del almacenamiento múltiple de datos

Analysis Services ofrece tres modos de almacenamiento para dimensiones, particiones y cubos:

- OLAP multidimensional (MOLAP)  
Los datos subyacentes de un cubo se almacenan con los datos de agregación en una estructura multidimensional de alto rendimiento. El sistema de almacenamiento MOLAP proporciona un rendimiento y compresión de datos excelentes.
- OLAP relacional (ROLAP)

Los datos subyacentes de un cubo se almacenan en una base de datos relacional con los datos de agregados. El sistema de almacenamiento ROLAP le permitirá sacar el máximo partido de la inversión que ha realizado en tecnología relacional y en herramientas de administración de datos empresariales.

- OLAP híbrido (HOLAP)

Los datos subyacentes de un cubo se almacenan en una base de datos relacional y los datos de agregación se almacenan en una estructura multidimensional de alto rendimiento. El sistema de almacenamiento HOLAP ofrece las ventajas de MOLAP para las agregaciones sin necesidad de duplicar los datos de detalle subyacentes.

Las herramientas que tenemos para trabajar con Analysis Services son:

- SQL Server Managment Studio, es la herramienta general de administración para Microsoft SQL Server. Desde la misma podremos administrar las bases de datos relacionales, las bases de datos de Analysis Services, los paquetes de Integration Services y también (como no) a Reporting Services.
- Business Intelligence Development Studio (BIDS) es la herramienta de desarrollo para Analysis Services integrada dentro de Visual Studio y es la que nos permite crear, editar, y administrar cubos, dimensiones, data source views, etc.

### **1.1.18. MICROSOFT AZURE**

Microsoft Azure (anteriormente Windows Azure y Azure Services Platform) es una plataforma ofrecida como servicio y alojada en los Data Centers de Microsoft. Anunciada en el Professional Developers Conference de Microsoft

(PDC) del 2008 en su versión beta, pasó a ser un producto comercial el 1 de enero de 2010.

Windows Azure es una plataforma general que tiene diferentes servicios para aplicaciones, desde servicios que alojan aplicaciones en alguno de los centros de procesamiento de datos de Microsoft para que se ejecute sobre su infraestructura (Cloud Computing) hasta servicios de comunicación segura y federación entre aplicaciones. En el reporte de Gartner "Magic Quadrant" más reciente, Azure fue uno de solo dos vendedores (el otro siendo Amazon Web Services) otorgado el título de "Líderes."

Windows Azure utiliza un sistema operativo especializado, llamado de la misma forma, para correr sus "capas" (en inglés "fabric layer") — un cluster localizado en los servidores de datos de Microsoft que se encargan de manejar los recursos almacenados y procesamiento para proveer los recursos (o una parte de ellos) para las aplicaciones que se ejecutan sobre Windows Azure.

Windows Azure se describe como una "capa en la nube" (en inglés "cloud layer") funcionando sobre un número de sistemas que utilizan Windows Server, estos funcionan bajo la versión 2008 de Windows Server y una versión personalizada de Hyper-V, conocido como el Hipervisor de Windows Azure que provee la virtualización de los servicios. La capa controladora de Windows Azure se encarga de escalar y de manejar la confiabilidad del sistema evitando así que los servicios se detengan si alguno de los servidores de datos de Microsoft tiene problemas y a su vez maneja la información de la aplicación web del usuario dando como ejemplo los recursos de la memoria o el balanceo del uso de esta.

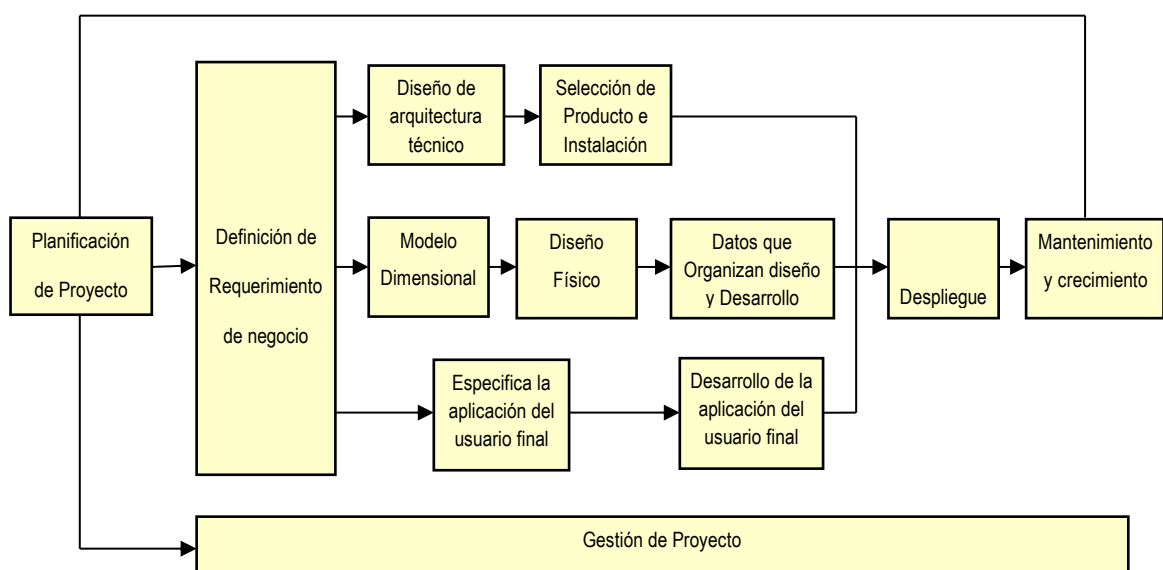
## **1.2. METODOLOGÍA**

### **1.2.1. METODOLOGÍA DE RALPH KIMBALL**

La metodología de Ralph Kimball se enfoca principalmente en el diseño de la base de datos que almacenará la información para la toma de decisiones.

El diseño se basa en la creación de tablas de hechos, es decir, tablas que contengan la información numérica de los indicadores a analizar, o sea la parte cuantitativa de la información para la toma de decisiones.

Las tablas anteriores se relacionan con tablas de dimensiones, las cuales contienen la información cualitativa, de los indicadores, es decir, toda aquella información que clasifique la información requerida.



**Figura N° 13:** Metodología de implementación

**Fuente:** (Kimball y Ross., 2013)

**a. Planificación del Proyecto**

La planificación busca identificar la definición y el alcance del proyecto de data warehouse, incluyendo justificaciones del negocio y evaluaciones de factibilidad. La planificación del proyecto se focaliza sobre recursos, perfiles, tareas, duraciones y secuencialidad.

El plan de proyecto resultante identifica todas las tareas asociadas con el ciclo de vida del datawarehouse e identifica las partes involucradas.

**b. Definición de Requerimientos del negocio**

Un factor determinante en el éxito de un proceso de Data Warehouse es la interpretación correcta de los diferentes niveles de requerimientos expresados por los diferentes niveles de usuarios. Aquí se identificará la información que requiere el usuario para desempeñar sus tareas. En esta etapa se especifica las funciones específicas que se obtendrán del Data Mart describiendo con claridad los requerimientos tales como:

- Definir los requerimientos del propietario.
- Definir los requerimientos del usuario final.

Estos requerimientos permitirán tener el ambiente operativo en el que se entregará el Data Mart.

**c. Diseño Técnico de la Arquitectura**

Los ambientes de data warehouse requieren la integración de numerosas tecnologías. Se debe tener en cuenta tres factores: los requerimientos del negocio, los actuales ambientes técnicos y las directrices técnicas estratégicas futuras planificadas para de esta forma poder establecer el diseño de la arquitectura técnica del ambiente de data warehouse.

**d. Modelo Dimensional**

La definición de los requerimientos del negocio determina los datos necesarios para cumplir los requerimientos analíticos de los usuarios. Diseñar los modelos de datos para soportar estos análisis requieren un enfoque diferente al usado en los sistemas operacionales. Básicamente se comienza con una matriz donde se determina la dimensionalidad de cada indicador y luego se especifican los diferentes grados de detalle (atributos) dentro de cada concepto del negocio (dimensión), como así también la granularidad de cada indicador (variable o métrica) y las

diferentes jerarquías que dan forma al modelo dimensional del negocio (BDM) o mapa dimensional.

**e. Diseño Físico**

El diseño físico de las base de datos se focaliza sobre la selección de las estructuras necesarias para soportar el diseño lógico. Algunos de los elementos principales de este proceso son la definición de convenciones estándares de nombres y seteos específicos del ambiente de la base de datos. La indexación y las estrategias de particionamiento son también determinadas en esta etapa.

**f. Diseño y Desarrollo de Presentación de Datos**

Las principales sub-etapas de esta zona del ciclo de vida son: la extracción, la transformación y la carga (ETL process). Se definen como procesos de extracción a aquellos requeridos para obtener los datos que permitirán efectuar la carga del Modelo Físico acordado. Así mismo, se definen como procesos de transformación los procesos para convertir o recodificar los datos fuente a fin poder efectuar la carga efectiva del Modelo Físico. Por otra parte, los procesos de carga de datos son los procesos requeridos para poblar el Data Warehouse.

Todas estas tareas son altamente críticas pues tienen que ver con la materia prima del data warehouse: los datos. La desconfianza y pérdida de credibilidad del data warehouse serán resultados inmediatos e inevitables si el usuario choca con información inconsistente. Es por ello que la calidad de los datos es un factor determinante en el éxito de un proyecto de data warehouse. Es en esta etapa donde deben sanearse todos los inconvenientes relacionados con la calidad de los datos fuente.

**g. Selección de Productos e Instalación**

Utilizando el diseño de arquitectura técnica como marco, es necesario evaluar y seleccionar componentes específicos de la arquitectura como

ser la plataforma de hardware, el motor de base de datos, la herramienta de ETL o el desarrollo pertinente, herramientas de acceso, etc. Una vez evaluados y seleccionados los componentes determinados se procede con la instalación y prueba de los mismos en un ambiente integrado de data warehouse.

**h. Especificación de Aplicaciones para Usuarios Finales**

No todos los usuarios del data warehouse necesitan el mismo nivel de análisis. Es por ello que en esta etapa se identifican los diferentes roles o perfiles de usuarios para determinar los diferentes tipos de aplicaciones necesarias en base al alcance de los diferentes perfiles (gerencial, analista del negocio, vendedor, etc.)

**i. Desarrollo de Aplicaciones para Usuarios Finales**

Siguiendo a la especificación de las aplicaciones para usuarios finales, el desarrollo de las aplicaciones de los usuarios finales involucra configuraciones del metadata y construcción de reportes específicos. Una vez que se ha cumplido con todos los pasos de la especificación y se tiene la posibilidad de trabajar con algunos datos de prueba, comienza el desarrollo de la aplicación.

**j. Implementación**

La implementación representa la convergencia de la tecnología, los datos y las aplicaciones de usuarios finales accesible desde el escritorio del usuario del negocio. Hay varios factores extras que aseguran el correcto funcionamiento de todas estas piezas, entre ellos se encuentran la capacitación, el soporte técnico, la comunicación. Todas estas tareas deben ser tenidas en cuenta antes de que cualquier usuario pueda tener acceso al data warehouse.



**k. Mantenimiento y crecimiento**

Data Warehouse es un proceso, de etapas bien definidas con comienzo y fin, pero de naturaleza espiral, pues acompaña a la evolución de la organización durante toda su historia. Se necesita continuar con los relevamientos de forma constante para poder seguir la evolución de las metas por conseguir. Según afirma Kimball, “si se ha utilizado el Ciclo de Vida, el data warehouse esta preparado para evolucionar y crecer”. Al contrario de los sistemas tradicionales, los cambios en el desarrollo deben ser vistos como signos de éxito y no de falla. Es importante establecer las prioridades para poder manejar los nuevos requerimientos de los usuarios y de esa forma poder evolucionar y crecer.

**l. Gerenciamiento del Proyecto**

El gerenciamiento del proyecto asegura que las actividades del ciclo del datawarehouse se lleven en forma y sincronizadas. Como lo indica el diagrama, el gerenciamiento acompaña todo el ciclo de vida. Entre sus actividades principales se encuentra el monitoreo del estado del proyecto y la comunicación entre los requerimientos del negocio y las restricciones de información para poder manejar correctamente las expectativas en ambos sentidos.

FASES	DESCRIPCION	ENTREGABLES
<b>Planeación</b>	<ul style="list-style-type: none"> <li>- Seleccionar la estrategia de implementación</li> <li>- Seleccionar la metodología de desarrollo</li> <li>- Seleccionar el ámbito de implementación</li> <li>- Seleccionar el enfoque arquitectónico</li> <li>- Desarrollar un programa y presupuesto del proyecto.</li> </ul>	<ul style="list-style-type: none"> <li>- Docum. de la estrategia de implementación</li> <li>- Docum. de la metodología de desarrollo.</li> <li>- Docum. del ámbito de implementación</li> <li>- Docum. del enfoque arquitectónico</li> <li>- Docum. del programa y presupuesto del proyecto</li> </ul>

<p><b>Determinación de Requerimientos</b></p>	<ul style="list-style-type: none"> <li>- Definir los requerimientos del propietario.</li> <li>- Definir los requerimientos del usuario final.</li> </ul>	<ul style="list-style-type: none"> <li>- Documento requerimientos del propietario.</li> <li>- Documento requerimiento del usuario</li> </ul>
<p><b>Análisis</b></p>	<ul style="list-style-type: none"> <li>- Análisis de cada Requerimiento encontrado mediante el Modelo Starnet.</li> </ul>	<ul style="list-style-type: none"> <li>- Modelo Starnet</li> </ul>
<p><b>Diseño y Modelización</b></p>	<ul style="list-style-type: none"> <li>- Diseño de la Base de Datos (Modelo Estrella).</li> <li>- Diseño del Sistema de Extracción de Datos.</li> <li>- Preparación de los datos.</li> <li>- Diseño de la administración de datos</li> </ul>	<ul style="list-style-type: none"> <li>- Documento Identificación de los Componentes del Modelo</li> <li>- Esquema estrella de cada Tabla</li> <li>- Documento Diseño Físico de las Tablas</li> <li>- Especificación de las claves foráneas</li> <li>- Diagrama del Diseño Físico del Data Mart</li> </ul>
<p><b>Implementación</b></p>	<ul style="list-style-type: none"> <li>- Determinación de Agregaciones.</li> <li>- Extracción de los datos y transformación.</li> <li>- Carga de los datos válidos.</li> <li>- Administración del Data Mart.</li> </ul>	<ul style="list-style-type: none"> <li>- Construcción de las Tablas y la Base de base de datos Estratégica en SQL AZURE</li> <li>- Documento Extracción de los Datos Y Transformaciones</li> <li>- Paquete de Servicio de Transformación de Datos (DTS)</li> <li>- Documento Sentencia de SQL para cada paso del DTS</li> </ul>
<p><b>Prueba</b></p>	<ul style="list-style-type: none"> <li>- Prueba de Performance</li> <li>- Prueba de calidad de datos</li> <li>- Prueba de los requerimientos</li> </ul>	<ul style="list-style-type: none"> <li>- Cuadro de la comparación de tiempo de demora en la ejecución de las consultas.</li> </ul>



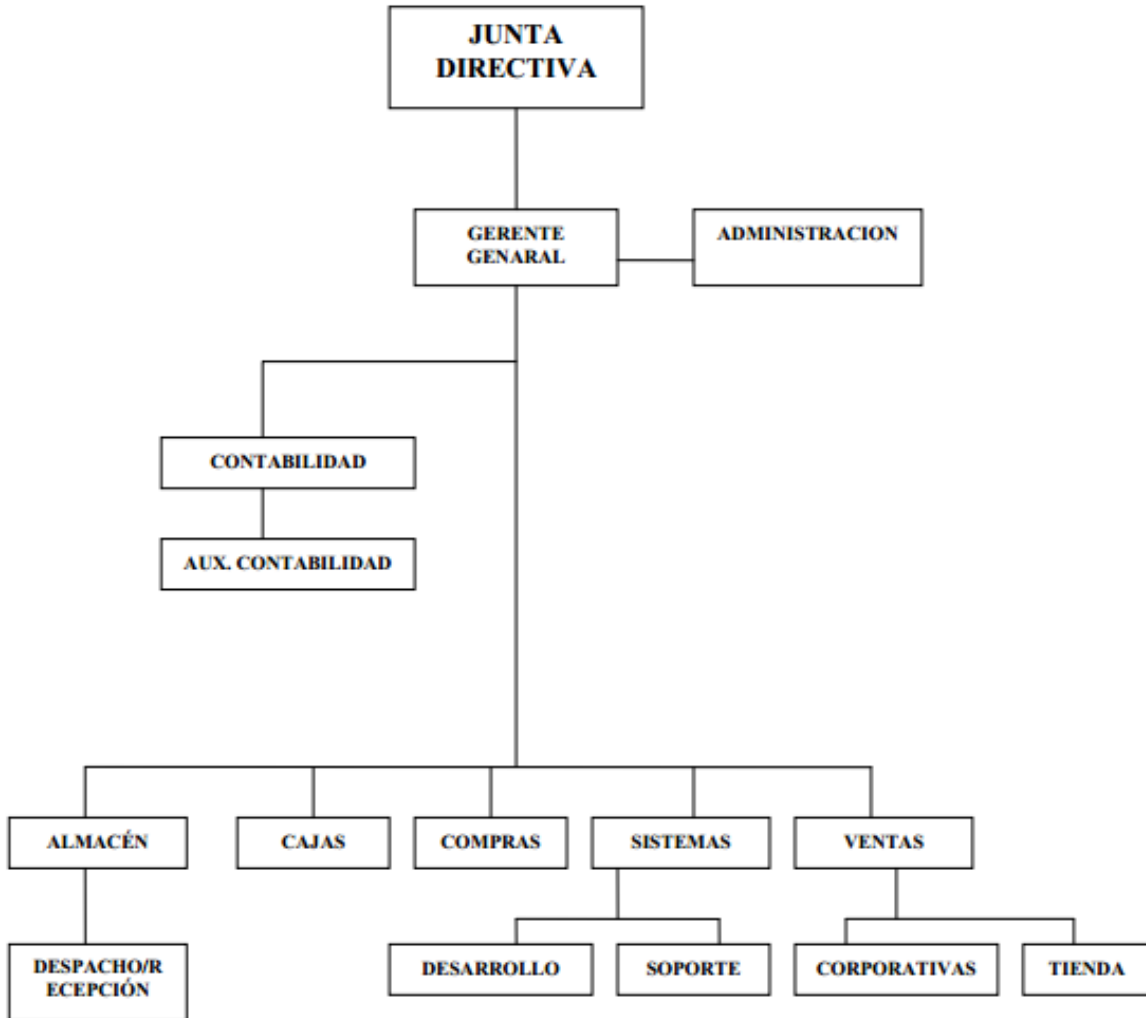
## CAPITULO II : DESARROLLO DEL TRABAJO DE TESIS

### I. PLANEACIÓN Y ADMINISTRACIÓN DEL PROYECTO

#### 1. El negocio

- **DESCRIPCIÓN** :
  - ✓ Razón Social: COPY VENTAS S.R.L.
  - ✓ Tipo de Sociedad: Sociedad Anónima Cerrada
  - ✓ RUC: 20440272593
  - ✓ Ubicación: Jr. Ayacucho No 538, Cercado, Distrito de Trujillo, Provincia de Trujillo, Departamento de La Libertad.
  - ✓ Rubro económico: Librerías en venta mayorista de otros productos
  
- **VISIÓN:** Ser la organización N° 1 en el norte el país, en el rubro de librería, con la finalidad de fomentar la cultura; manteniendo un crecimiento sostenido y garantizando la calidad de nuestros productos y servicios a través de la capacitación y especialización de nuestros colaboradores.
  
- **MISIÓN:** Mantener un alto estándar de calidad que supere las expectativas de nuestros clientes y/o consumidores en cuanto a los productos y servicios, invirtiendo en tecnología, logística y capacitación.

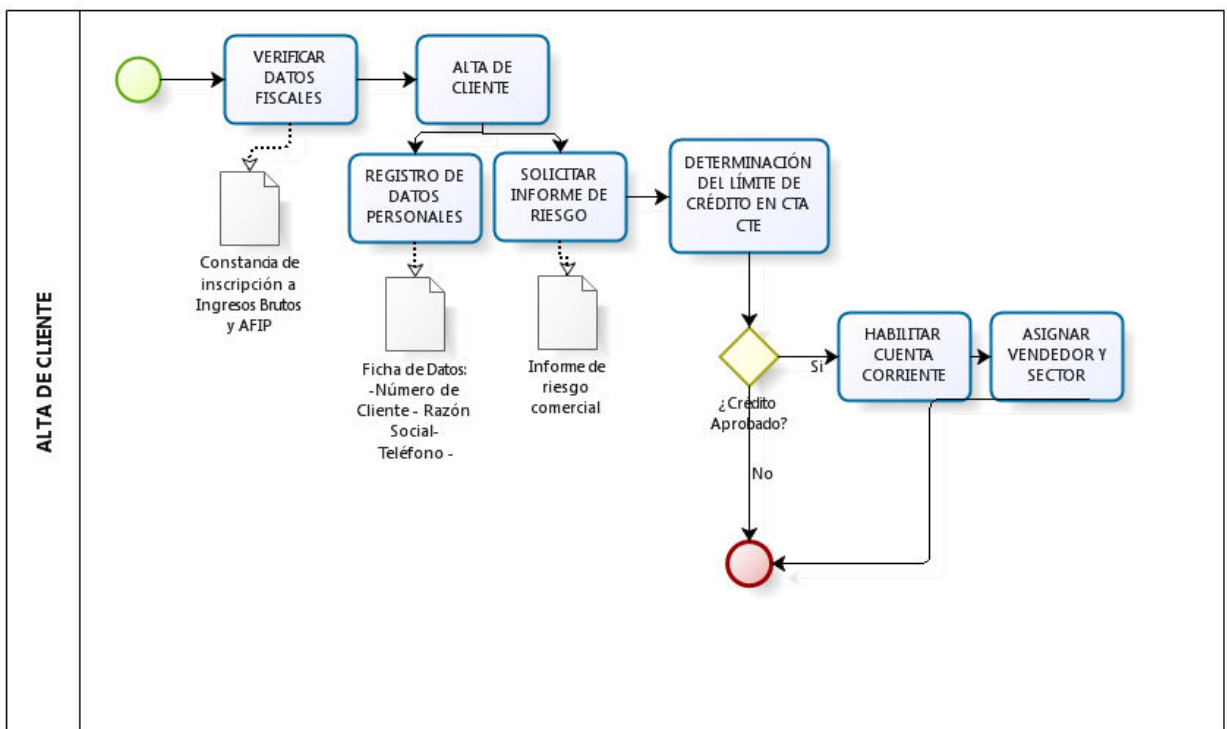
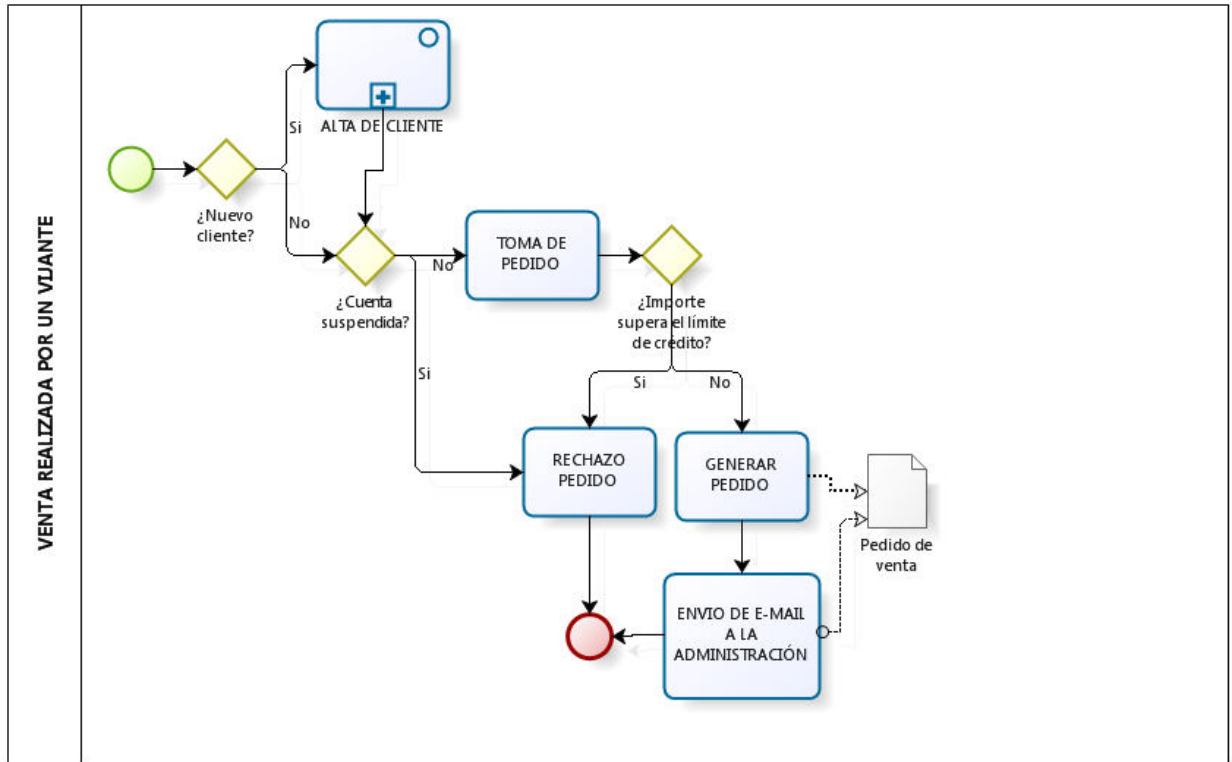
- **ORGANIGRAMA:**



**PROCESO A ESCOGER Y MODELAR EL PROCESO ( BIZAGI)**

Hemos considerado:

**a) Ventas**



## 2. Selección de la Estrategia de Implementación

Contiene las siguientes tareas:

- a) Conocer las necesidades de la empresa y entender los requerimientos que se definen en las entrevistas.
- b) Planificación y asignación de recursos de la empresa.
- c) Analizar las políticas de área de ventas de la empresa
- d) Recolectar los requerimientos empresariales.
- e) Analizar los requerimientos.
- f) Realizar el diseño y la arquitectura para el Data Mart.
- g) Implementar el Data Mart.

Resaltamos algunas pautas que recomendamos deben seguirse a la hora de abordar un proyecto de este tipo:

*"La Base de Datos de Riesgos debe estar separada de las Bases de Datos Operacionales"* con objeto de no interferir en la actividad del día a día, disponiendo de la información necesaria para Riesgos (interna y externa) y en un entorno orientado hacia la consulta y el análisis (Data Warehouse).

“La planificación de recursos entraña dos niveles de consideración”. Primero, están las cuestiones más amplias de cómo deben repartirse los recursos entre las diversas funciones, departamentos. Segundo, la cuestión más detallada de cómo se deben desplegar los recursos, dentro de cualquier parte de la organización para que apoye mejor las estrategias, está relacionado con los aspectos operacionales de la planificación de recursos y se apoya en la valoración detallada de la capacidad estratégica.

*"Concepción del sistema como un conjunto de herramientas de análisis"*, debido a que las actividades de Análisis de Riesgos no se pueden automatizar completamente, puesto que requieren análisis y decisiones del usuario.

*"Diseño del sistema no orientado a procesos"*; se debe disponer de un conjunto abierto de herramientas que se utilizan con propósitos determinados no relacionados con las necesidades operativas.

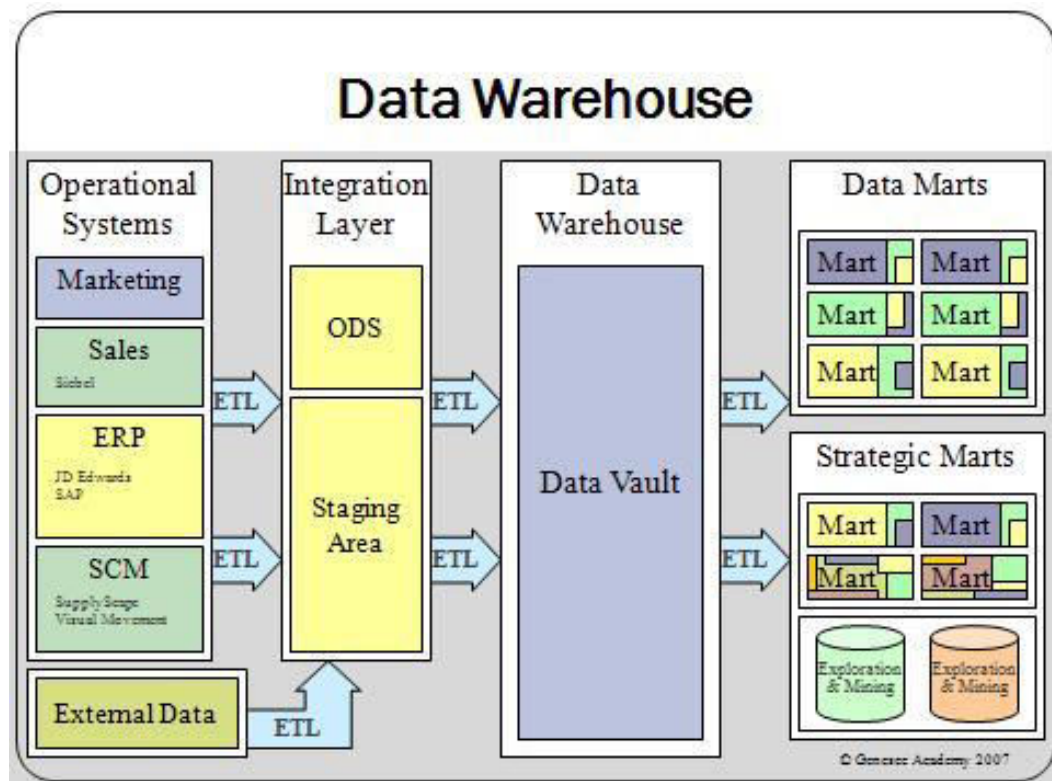
*"Abordar el sistema con un enfoque de desarrollo gradual"*, se debe comenzar con un esqueleto básico de funcionalidad y datos que produzcan resultados a corto

plazo y permita aprender en la práctica, y a continuación ir configurando progresivamente nuevas funcionalidades conforme la experiencia lo vaya requiriendo.

### 3. Selección de la Metodología de Desarrollo

Para el desarrollo del Datamart es necesario contar con una base teórica en lo referente a que es la inteligencia de negocios y para que se usa un Datamart.

Inteligencia de Negocios es el conjunto de estrategias y herramientas enfocadas a la administración y creación de conocimiento mediante el análisis de datos existentes en una organización o empresa. Una de ellas es el método propuesto por Ralph Kimball, el cual se utilizará para desarrollar la metodología de trabajo, adaptándola al caso específico de este proyecto



### 4. Selección del Ámbito de Implementación

El ámbito de la implementación de los Data Mart fue determinado basándose en los requerimientos de información del personal de nivel ventas y almacén de la Empresa COPY VENTAS S.R.L. Para ello se formuló las siguientes preguntas:



- ✓ **¿Cuál es el área más recomendable estratégicamente para aplicar el Data Mart?**

El área de ventas es la más importante, en el cual debe ser aplicado el Data Mart, por ser el área donde se desarrolla la actividad principal de la Empresa.

- ✓ **¿Cuál es periodo de carga de que tendrá el Data Mart?**

El Data Mart tendrá que procesar la información diaria de todas las sucursales de la empresa,

- ✓ **¿Cuál es el rango de consultas empresariales a los que se debe responder inicialmente el Data Mart?**

El rango de consultas empresariales que debe responder inicialmente el Data Mart es de 10 consultas.

## 5. Selección del enfoque arquitectónico

Estarán compuestos de los siguientes niveles:

- **BACK ROOM**

Compuesto por:

- ✓ 1 Servidor HP ProLiant DL380 Generación 5 DL380 G6 Intel Xeon de 3 GHz, cuya estructura se encuentra en SQL Server.
- ✓ 1 servidor de Base de Datos SQL Server 2014, que servirá como repositorio de los data marts de la Empresa.

- **FRONT ROOM**

PC's clientes, que serán cada una de las computadoras desde las que se acceda a la información que brindaran los data marts, a través de documentos en Excel.

## 6. Desarrollo de un Programa y del Presupuesto del Proyecto

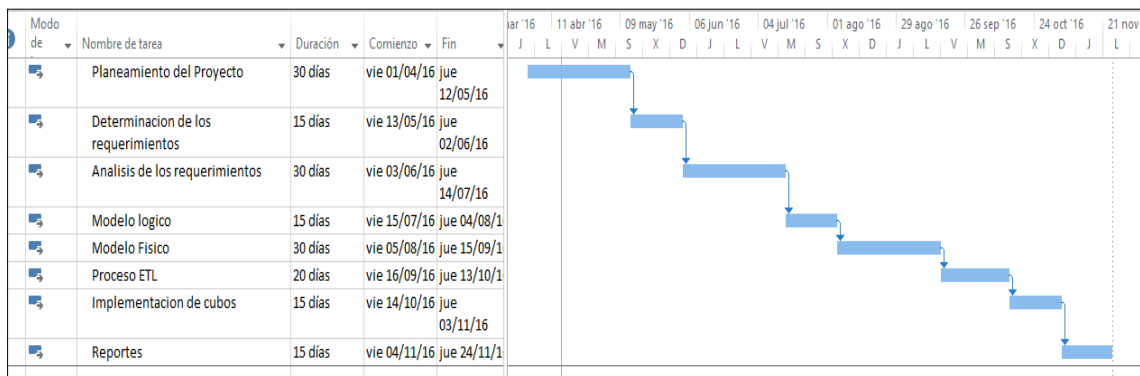
### a) Presupuesto

<b>RECURSOS HUMANOS</b>	<b>COSTO S/.</b>
01 Investigadores	3,000.00
01 Asesor	3,000.00
<b>TOTAL</b>	<b>6,000.00</b>

<b>BIENES</b>	<b>COSTO S/.</b>	
MATERIALES	01 Millar de papel bond A-4	42.00
	05 Lapiceros	10.00
	02 Correctores	3.50
	01 Lapiz	1.00
	02 Cartuchos de tinta Negra	110.00
	02 Cartuchos de tinta Color	120.00
	01 Toner 35 <sup>a</sup>	210.00
	10 Folder manila A-4	10.00
	<b>TOTAL</b>	<b>506.5</b>
EQUIPOS	01 Computadora Intel Pentium – 4GB RAM	1,200.00
	01 Impresora HP deskjet D2500	199.00
	01 Grabador de DVD	120.00
	01 Memoria USB 8GB	30.00
	<b>TOTAL</b>	<b>1,949.00</b>
SOFTWARE	01 SQL Server 2014	230.00
	01 Licencia Excel (Office 2013)	400.00
	<b>TOTAL</b>	<b>630.00</b>
SERVICIOS	Servicio de transporte	90.00
	Servicio de tipos	150.00
	Servicio de Fotocopiados	10.00
	Servicios de Internet	50.00
	<b>TOTAL</b>	<b>300.00</b>

<b>PRESUPUESTO TOTAL GENERAL</b>	
<b>CATEGORIA</b>	<b>COSTO TOTAL S/.</b>
Recursos Humanos	6,000.00
Materiales	506.50
Equipo	1,949.00
Software	630.00
Servicios	300.00
<b>COSTO TOTAL :</b>	<b>9,385.50</b>

**b) Cronograma de actividades**



## 7. Desarrollo del escenario del uso empresarial

Para este proyecto de tesis el escenario de uso empresarial es el que se muestra a continuación con sus respectivos componentes:

### A. Descripción de los Stakeholder

- **Personal Involucrado en el Proyecto**

Nombre	Representa	Rol
<b>GERENTE GENERAL</b>	Directivo de la empresa que fija estrategias y directrices para el buen funcionamiento de la empresa	<ul style="list-style-type: none"> <li>• Establecer objetivos para toda la organización, planificando el trabajo mediante planes estratégicos para que se logren los objetivos controlando para comparar los resultados reales con los esperados.</li> </ul>
<b>JEFE DE ALMACÉN</b>	Personal encargado de Controlar y registrar todo tipo de formatos y documentos que determine la empresa, asimismo, tramitara los que sean necesarios para las actividades de la especialidad, a fin de contar con el reabastecimiento oportuno	<ul style="list-style-type: none"> <li>• Planificar y establecer una adecuada distribución de los espacios físicos para el almacenaje de la mercadería</li> <li>• Mantener al día el sistema de inventarios en lo que corresponde al registro adecuado y oportuno de ingresos y egresos de mercadería</li> </ul>
<b>JEFE DE VENTAS</b>	Personal encargado de Realizar las compras de mercadería para la institución en coordinación con cada área de la empresa, para mantener un adecuado control de gastos	<ul style="list-style-type: none"> <li>• Recopilación de requerimiento por área, verificación y validación de los requerimientos tanto en sistema como físicamente para posteriormente preparar de la orden de compras.</li> </ul>

- **Usuarios del Data Mart**

<b>Nombre</b>	<b>Representa</b>	<b>Stakeholder</b>
<b>GERENTE GENERAL</b>	Es aquel que se encarga de plantear las estrategias para el buen funcionamiento de la empresa	<i>Acceso total de las dimensiones del Data Mart</i>
<b>JEFE DE ALMACÉN</b>	Encargado de Controlar y registrar todo tipo de formatos y documentos que determine la empresa, así como de ver el reabastecimiento de productos	<b>Jefe de Almacén</b>
<b>JEFE DE VENTAS</b>	Es el encargado de Realizar las compras de mercadería para la institución en coordinación con cada área de la empresa	<b>Jefe de ventas</b>

## II. DETERMINACIÓN DE LOS REQUERIMIENTOS

### II.1. Requerimientos del Propietario

Al hacer uso de las técnicas mencionadas anteriormente, respondemos a las preguntas que proponen los propietarios de Copy Ventas.

✓ **¿Por qué construir un Data Mart?**

Para poder consultar la información sin tener que depender de personal técnico que genere los reportes o consultas ad hoc a través de un lenguaje y/o herramienta de programación, lo cual además conlleva en disminuir el tiempo de espera en la generación de reportes por parte del personal técnico. beneficiando de esta manera un mejor desempeño del área de producción de la Empresa.

✓ **¿Cuál será el impacto sobre la organización?**

El impacto del Data Mart sobre el área de ventas de la Empresa será beneficioso a corto y largo plazo de acuerdo a la rapidez con que se tomen las decisiones y a la creatividad que se emplee, permitiendo la reducción en la obtención de informes solicitadas por esta, que servirán para tomar decisiones en el momento necesario, lo cual le permitirán tomar ventajas sobre el mercado.

✓ **¿Cuáles son los riesgos?**

Puede presentar problemas de consistencia, el tiempo que se emplea es su desarrollo es largo. Al crecer el Data Mart, el rendimiento de las consultas decae y deja de ser óptimo.

## II.2. Requerimientos del Usuario Final

A continuación se muestran las consultas expresadas en la terminología del usuario final:

Nº	CONSULTAS	UNIDAD
1	¿Cuál es el monto de un determinado cliente en compras en un mes?	Monto en soles
2	¿Cuál es el Monto de ventas por tipo de producto?	Monto en soles
3	¿Cuál es el monto de un tipo de producto por proveedor en 3 meses?	Monto en soles,
4	¿Cuál es el monto de ventas clientes en un semestre?	Monto en soles,
5	¿Cuál es el monto que generan las productos en un año?	Monto en soles
6	¿Cuál es la cantidad de productos vendidos por Tipo, vendedor, mes y semestre?	Cantidad de Productos

7	¿Cuál es la cantidad de Productos que se han pagado al crédito y al contado?	Cantidad de Productos
8	¿Cuál es el monto que generan las promociones que da un proveedor en un semestre?	Monto en soles
9	Listar a los Vendedores y el monto total de cada una de las ventas que han generado	Monto en soles,
10	¿Cuál es el porcentaje de las ventas de los productos según su tipo por mes?	Monto en soles

### II.3. Requerimientos no Funcionales

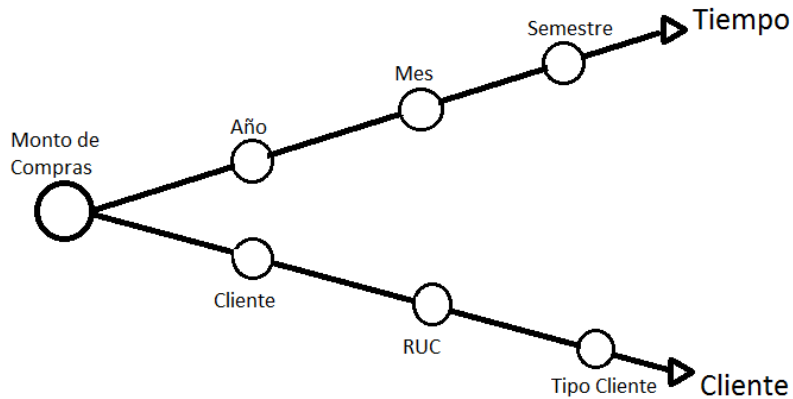
- ✓ Establecer seguridad adecuada para la administración de los Data Marts.
- ✓ Brindar mejor información a los usuarios, en el menor tiempo.
- ✓ Los Data Marts deben funcionar, con el manejador de base de datos Microsoft SQL Server 2014 y utilizando el Microsoft SQL Server 2012 Análisis Services como herramienta de diseño y construcción de los cubos.
- ✓ Utilizar MS Excel 2013 para mostrar los reportes de los cubos.
- ✓ El sistema debe poseer un tiempo de respuesta breve ya que es utilizado en un solo puesto de trabajo
- ✓ El rendimiento del Data Mart debe ser superior a las herramientas utilizadas para la consulta en los sistemas transaccionales.

### II.4. Análisis de los requerimientos

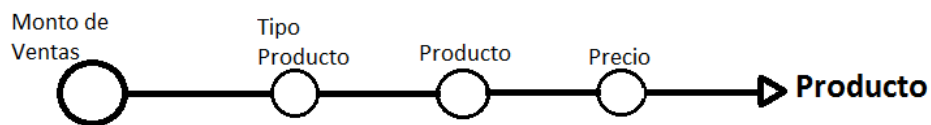
Se hará un análisis de cada requerimiento del usuario final y definiendo algunas tablas de la Base de Datos Operacional de “Copy Ventas”. Para la gestión de reportes y apoyo a la toma de decisiones, manejar la información, examinarla desde diferentes puntos de vista para un mejor análisis estratégicos según Ralph Kimball.

En equipo con los usuarios finales Gerente, Jefe de almacén y Ventas así como los analistas se podrán definir las mediciones del Data Mart.

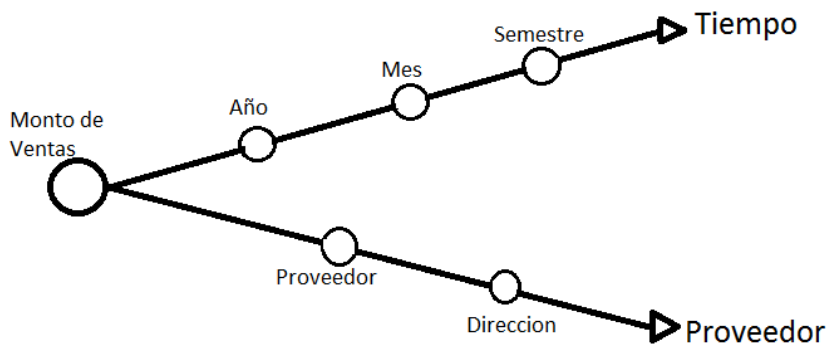
**R1: ¿Cuál es el monto de un determinado cliente en compras en un mes?**



**R2: ¿Cuál es el Monto de ventas por tipo de producto?**

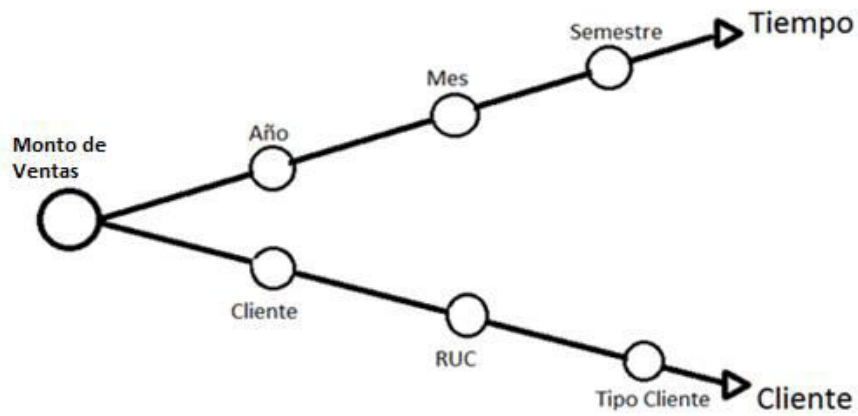


**R3: ¿Cuál es el monto de un producto para un determinado proveedor en 3 meses?**

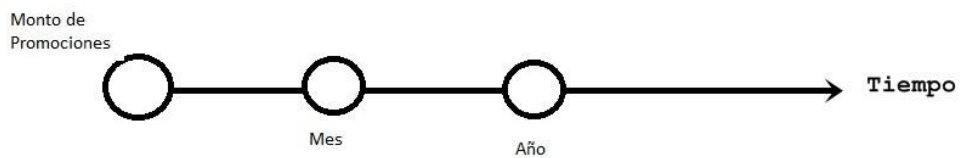




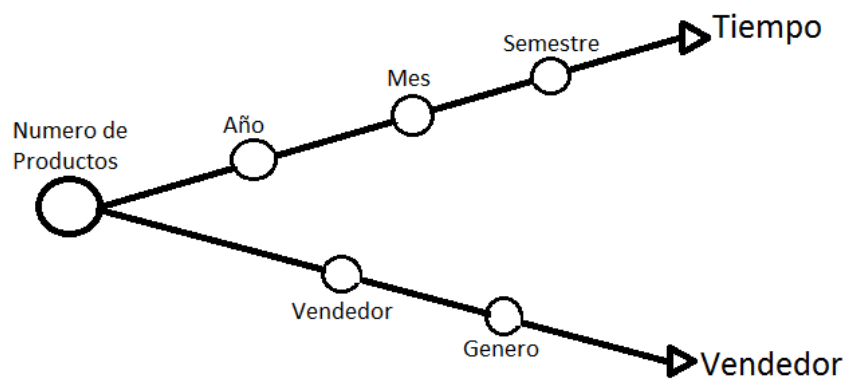
R4: ¿Cuál es el monto de ventas de los clientes en un semestre?



R5: ¿Cuál es el monto que generan las promociones en un año?



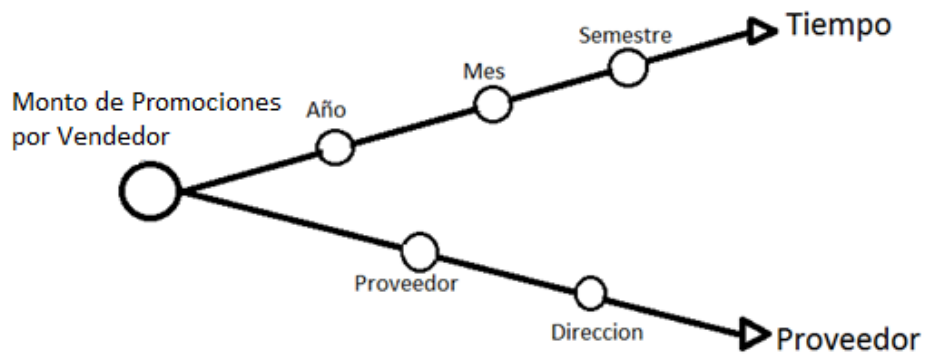
R6: ¿Cuál es la cantidad de productos vendidos por Tipo, vendedor, mes y año?



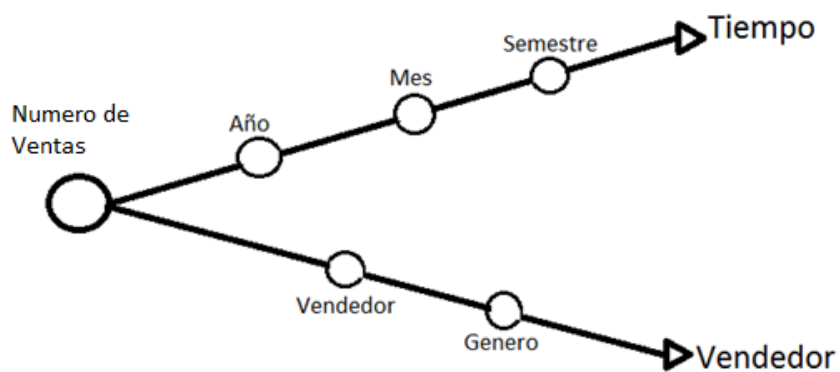
R7: ¿Cuál es la cantidad de Ventas que se han pagado al crédito y al contado?



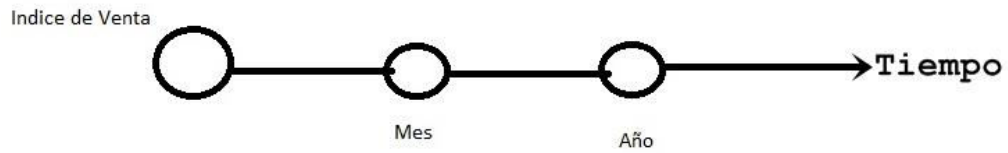
R8: ¿Cuál es el monto que generan las promociones que da un proveedor en un año?



R9: Listar a los Vendedores y el monto total de cada una de las ventas que han generado

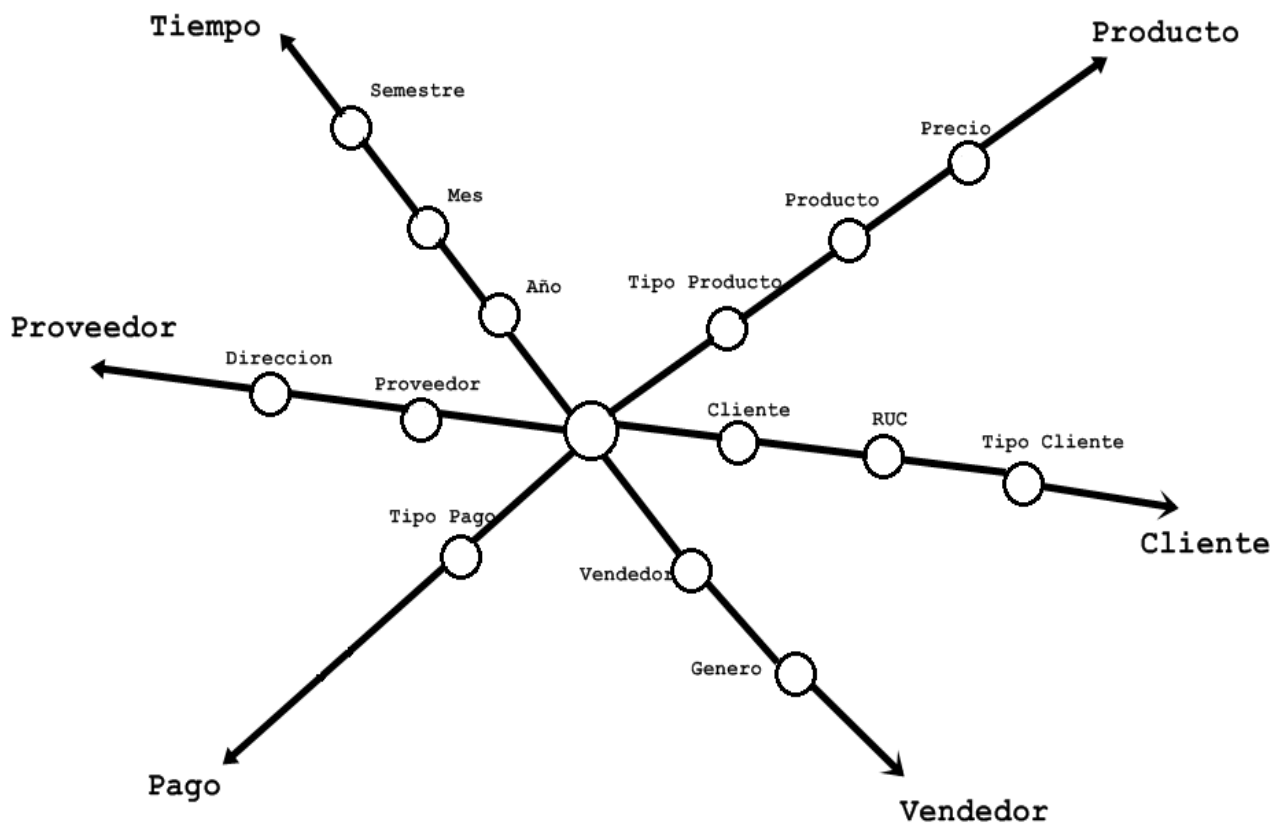


R10: ¿Cuál es el índice de incremento de las ventas de los productos según su tipo por mes?

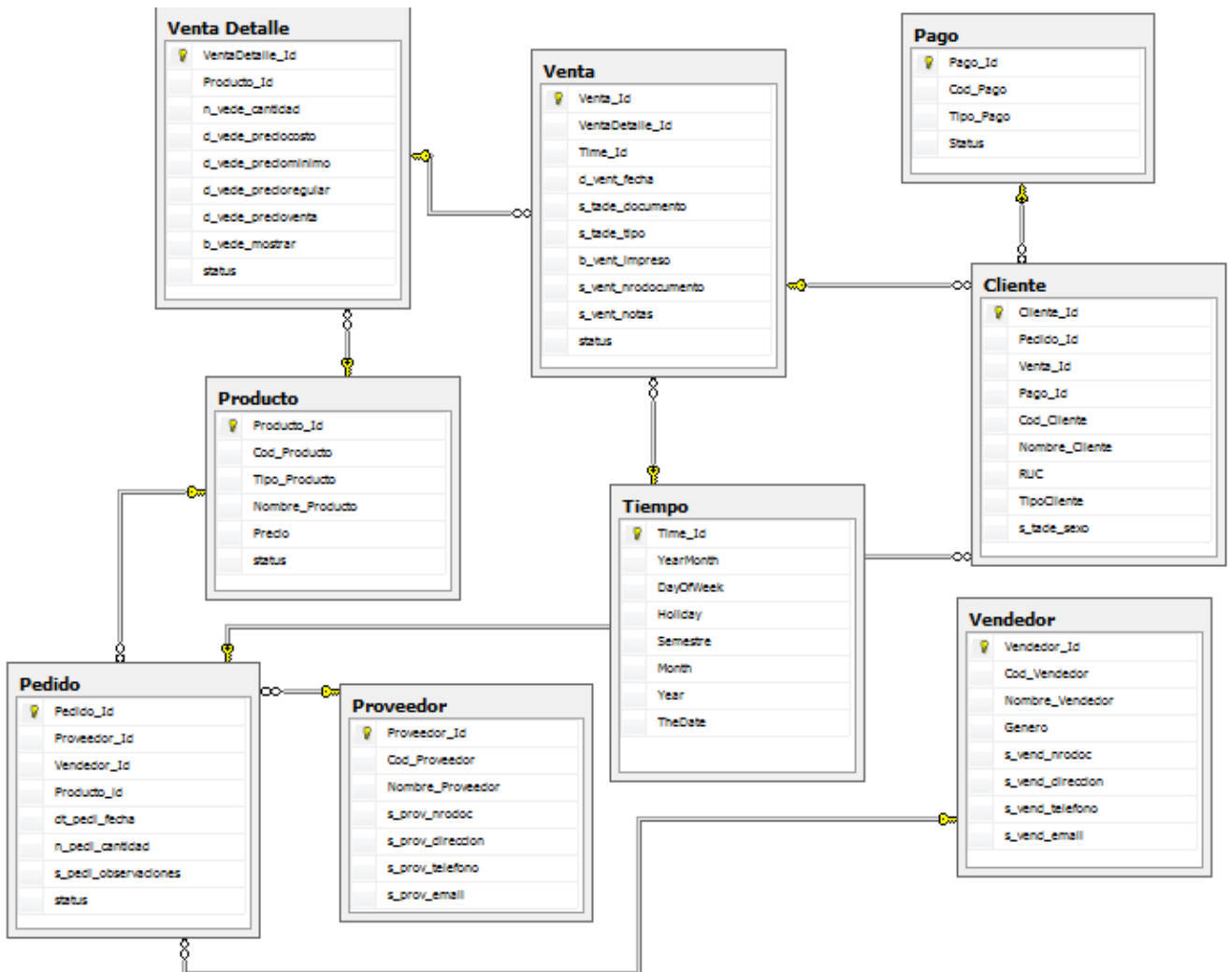


### III. DISEÑO TÉCNICO DE LA ARQUITECTURA

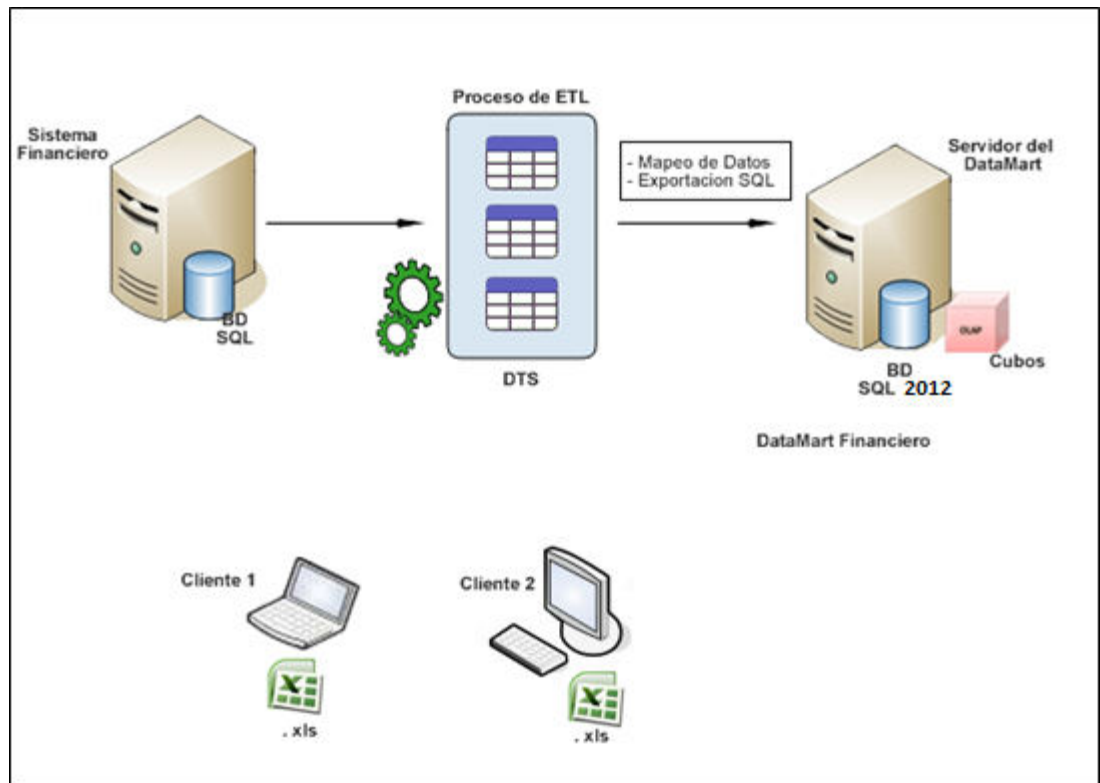
#### III.1. Modelo Starnet



### III.2. Nivel de Datos



### III.3. Nivel Técnico



Enfoque Arquitectónico del DataMart

Estarán compuestos de los siguientes niveles:

- **BACK ROOM**

Compuesto por:

- ✓ 1 Servidor HP ProLiant DL380 Generación 5 DL380 G6 Intel Xeon de 3 GHz, cuya estructura se encuentra en SQL Server.
- ✓ 1 servidor de Base de Datos SQL Server 2014, que servirá como repositorio de los data marts de la Empresa.

- **FRONT ROOM**

PC's clientes, que serán cada una de las computadoras desde las que se acceda a la información que brindaran los data marts, a través de documentos en Excel.

## **IV. MODELADO DIMENSIONAL**

### **IV.1. Identificación de los Componentes del Modelo**

Se busca presentar los datos en un marco de trabajo estándar.

Identificaremos los componentes del modelo desarrollando el método de los cuatro pasos. Luego, se construirá el primer diagrama para el modelo y se detallará cada componente del diagrama.

Para identificar los componentes utilizaremos el método propuesto por Ralph Kimball.

#### **El Método de los cuatro pasos**

Consiste en cuatro elecciones:

- Elección del Data Mart
- Elección del Objetivo de la Tabla de Hechos
- Elección de las Dimensiones
- Elección de los Hechos

Cada paso está desarrollado a continuación:

#### **A. Elección del Data Mart**

En este paso identificaremos los posibles Data Mart que se pueden construir en la Librería Copy Ventas S.R.L.. En nuestro caso identificaremos el área que está en estudio por este proyecto, el cual ya tenemos conocimiento sobre su Modelo General de Datos.

Para identificar nuestro Data Mart en estudio recurrimos al siguiente método:

##### **a) Método de la Matriz**

Menciona los posibles hechos relacionados y que son usados juntos.  
Listamos las entidades que apoyan a estos hechos.

**b) Listado de los Data Marts**

Según la situación general de la Empresa, se identificó las posibles fuentes de Data Marts que corresponderán a las filas de la matriz:

- ✓ Área de Ventas.

**c) Listado de las Dimensiones**

Para el Data Mart mencionados anteriormente listamos las siguientes posibles dimensiones:

- ✓ Pago
- ✓ Tiempo
- ✓ Vendedor
- ✓ Proveedor
- ✓ Cliente
- ✓ Producto

**d) Listado de Medidas**

- ✓ Monto en soles
- ✓ Cantidad de Clientes
- ✓ Cantidad de Productos
- ✓ Cantidad de Pagos
- ✓ Cantidad de Proveedores
- ✓ Numero de Vendedores

**e) Marcado de las Intersecciones**

Ordenamos las filas y las columnas en una tabla, dando forma a la matriz, y marcamos las intersecciones donde exista una dimensión relacionada a un Data Mart.

DIMENSIONES	AREA DE VENTAS
CLIENTE	X
PRODUCTO	X
VENDEDOR	X
PROVEEDOR	X
TIEMPO	X
PAGO	X

### B. Elección del Objetivo de la Tabla de Hechos

Este paso consiste en declarar como es el Registro del Hecho en las Tablas de Hechos, es decir hay que definir claramente y exactamente que registros de cada Tabla de Hechos figurarán en el diseño del modelado del Data Mart.

Tomamos la siguiente definición para cada Tabla de Hechos:

Tabla de hechos	Objetivo
Ventas	“Administrar los recursos económicos para tener un control diario de las Ventas de la Librería Copy Ventas S.R.L.”.



### C. Elección De Dimensiones

Al establecer el objetivo para cada tabla de hechos, podemos escoger cuales serán nuestras dimensiones para cada una de estas tablas. El objetivo por si mismo establece cuales serán nuestras dimensiones.

A continuación por cada objetivo escogeremos las dimensiones:

Tabla de Hechos	Objetivos	Dimensiones
Ventas	“Administrar los recursos económicos para tener un control diario de las Ventas de la Librería Copy Ventas S.R.L.”.	Cliente Producto Vendedor Proveedor Tiempo Pago

### D. Elección de Hechos

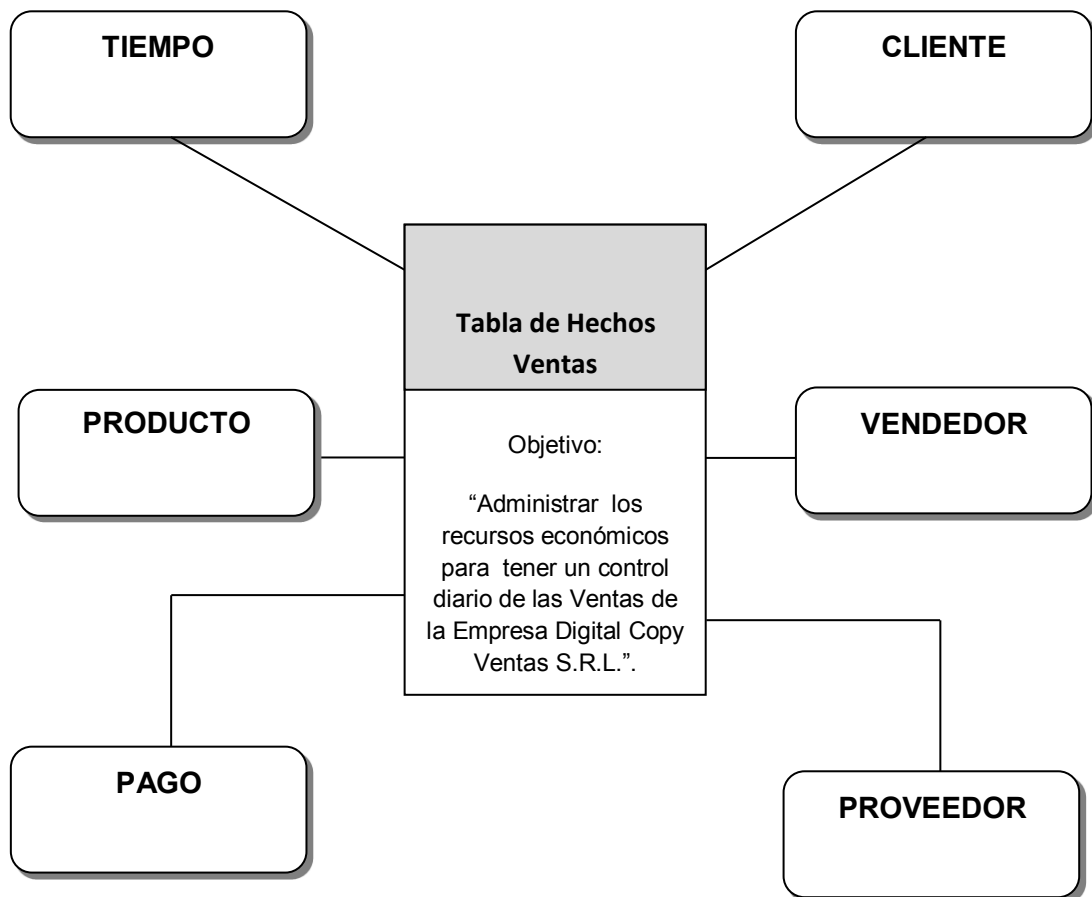
El Objetivo de la tabla de hechos también permite definir las medidas y hacer claro el alcance que debe tener

Tabla de Hechos	Hechos
	Monto en Soles
	Numero de Vendedores
	Cantidad de Ventas
	Cantidad de Productos
	Cantidad de Proveedores
	% Descuento

## IV.2. Diagrama de la Tabla de Hechos

Para esta etapa se prepara un diagrama lógico para la Tabla de Hechos completada. Se nombra a la tabla de Hechos, esto establece su objetivo y muestra todas las dimensiones conectadas a la tabla de hechos.

- **Diagrama de la Tabla de Hechos Operaciones**



✓ **Detalle de las Tablas de Hechos**

Nombre de la Tabla	Nombre de la Columna	Descripción de la Columna
FACT_VENTAS	Cliente_Id	Llave primaria única para la dimensión Cliente.
	Producto_Id	Llave primaria única para la dimensión Producto.
	Vendedor_Id	Llave primaria única para la dimensión Vendedor.
	Proveedor_Id	Llave primaria única para la dimensión Proveedor.
	Tiempo_Id	Llave primaria única para la dimensión Tiempo.
	Pago_Id	Llave primaria única para la dimensión Pago.

✓ **Tabla Detalle de las claves de las dimensiones**

Nombre de la Tabla	Nombre de la Columna	Descripción de la Columna
FACT_VENTAS	Monto en Soles	Monto Costos en soles
	Cantidad de Proveedores	Cantidad de Proveedores
	Numero de Vendedores	Cantidad de Vendedores con los que cuenta la Empresa
	Cantidad de Ventas	Cantidad de Ventas realizadas por la Empresa
	Cantidad de Productos	Cantidad de Productos con los que cuenta la Empresa
	% Descuento	Descuento a Clientes

Tabla N° 20 Detalle de las Medidas de la Tabla de Hechos.

Por consiguiente de las Tablas completadas anteriormente se obtiene la siguiente tabla de Hechos:

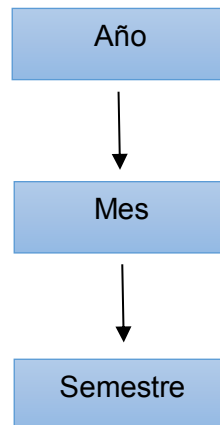
<b>COLUMNAS</b>	<b>TIPO DE DATO</b>	<b>VALORES NULL</b>	<b>LONGITUD</b>	<b>DESCRIPCION</b>
<b>Cliente_Id</b>	<b>Int</b>	<b>No</b>	<b>4</b>	<b>Tabla de hechos que está relacionada con todas las dimensiones y que contiene datos cuantitativos de vital importancia para el análisis de Ventas de la Empresa.</b>
<b>Producto_Id</b>	<b>Int</b>	<b>No</b>	<b>4</b>	
<b>Vendedor_Id</b>	<b>Int</b>	<b>No</b>	<b>4</b>	
<b>Proveedor_Id</b>	<b>int</b>	<b>No</b>	<b>4</b>	
<b>Tiempo_Id</b>	<b>Int</b>	<b>No</b>	<b>4</b>	
<b>Pago_Id</b>	<b>Int</b>	<b>No</b>	<b>4</b>	
<b>Monto</b>	<b>money</b>	<b>No</b>	<b>6</b>	
<b>Numero de Vendedores</b>	<b>Int</b>	<b>No</b>	<b>100</b>	
<b>Numero de Clientes</b>	<b>Int</b>	<b>No</b>	<b>100</b>	
<b>Cantidad de Proveedores</b>	<b>int</b>	<b>No</b>	<b>100</b>	
<b>Cantidad de Productos</b>	<b>Int</b>	<b>No</b>	<b>100</b>	
<b>Cantidad de Pagos</b>	<b>int</b>	<b>No</b>	<b>100</b>	

**Tabla de Hechos del Data Mart**

## Dimensión

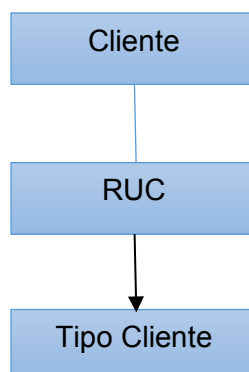
De acuerdo a nuestras Dimensiones elegidas anteriormente construimos los detalles y sus jerarquías de cada una de ellas.

### Dimensión Tiempo:



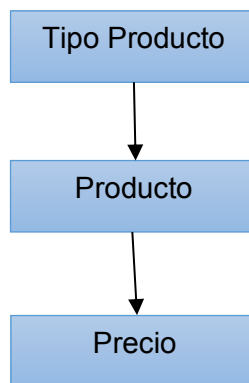
Nombre del Atributo	Descripción del Atributo	Valores de Muestra
<b>Año</b>	Representa el año.	2012, 2013,...
<b>Mes</b>	Representa los meses que posee un año.	Abril, Mayo,...
<b>Semestre</b>	Representa el semestre	Enero-Marzo, Marzo-Junio

### Dimensión Cliente:



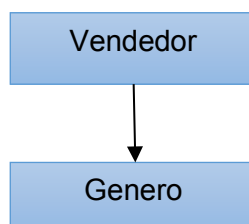
Nombre del Atributo	Descripción del Atributo	Valores de Muestra
<b>Cliente</b>	Representa el cliente.	Carlos, Juan,...
<b>RUC</b>	Representa el Ruc	0175434643,..
<b>Tipo Cliente</b>	Representa el tipo cliente.	Natural, Jurídico

### Dimensión Producto:

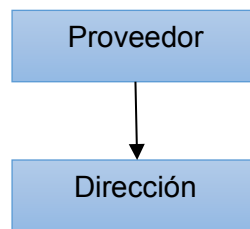


Nombre del Atributo	Descripción del Atributo	Valores de Muestra
<b>Tipo Producto</b>	Representa al tipo de producto.	Útiles de escritorio, libros...
<b>Producto</b>	Representa al nombre del producto	Lapicero, Lápiz, etc.
<b>Precio</b>	Representa el precio de cada producto	12,14,15

### Dimensión Vendedor:



Nombre del Atributo	Descripción del Atributo	Valores de Muestra
<b>Vendedor</b>	Representa al nombre del vendedor	Juan, Carlos,..
<b>Genero</b>	Representa el sexo de cada vendedor	Hombre, Mujer, Otro

**Dimensión Proveedor**

Nombre del Atributo	Descripción del Atributo	Valores de Muestra
<b>Proveedor</b>	Representa el nombre del proveedor.	José, Juan,..
<b>Dirección</b>	Representa la dirección del proveedor	Jr. Puno 234,..

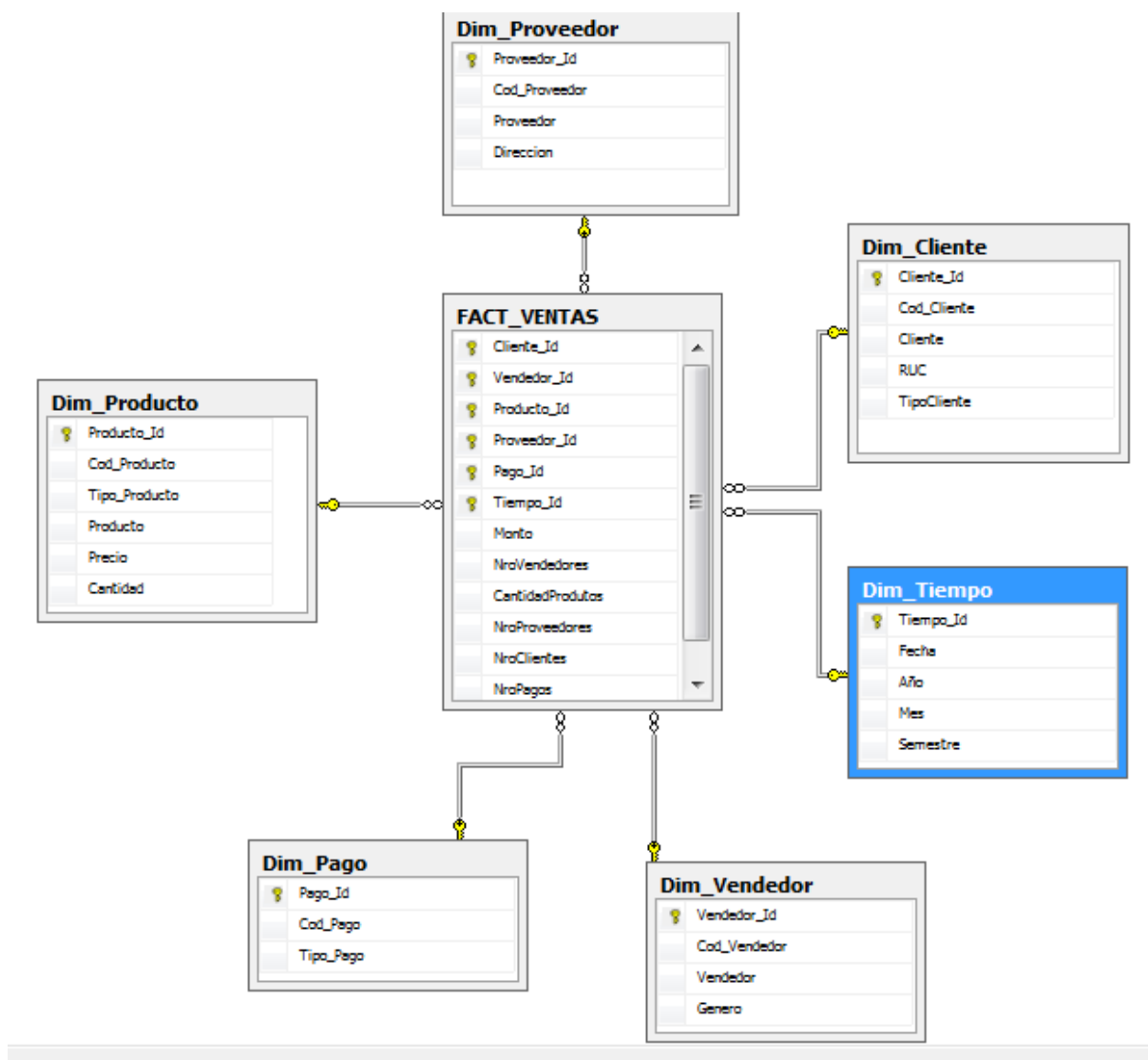
**Dimensión Pago:**

Nombre del Atributo	Descripción del Atributo	Valores de Muestra
<b>Tipo Pago</b>	Representa a la forma de Pago que realiza el Cliente	Crédito, Contado

**IV.3. Esquema Estrella**

Sabiendo el número de Tablas de Hechos y las dimensiones asociadas a estas, orientamos nuestro Data Mart al Esquema Estrella para una mejor visualización del Modelo para el Diseño Físico.

- Componente: TABLA DE HECHOS VENTAS
- Componente: TABLA DE DIMENSION CLIENTE
- Componente: TABLA DE DIMENSION PRODUCTO
- Componente: TABLA DE DIMENSION VENDEDOR
- Componente: TABLA DE DIMENSION PROVEEDOR
- Componente: TABLA DE DIMENSIÓN TIEMPO
- Componente: TABLA DE DIMENSIÓN PAGO



Por consiguiente, con los datos del diseño lógico y los cambios en la estructura física, se obtiene el siguiente **Modelo de Base de Datos Física** para el DataMart.



## V. DISEÑO FÍSICO

Se tiene que tener en cuenta lo siguiente:

Modificar nombres a nombres estándar, si fuera necesario. Para nuestro Data Mart se realizaron los siguientes cambios:

DISEÑO LÓGICO	DISEÑO FÍSICO
Tabla de Hechos Ventas	FACT_VENTAS
Dimensión Cliente	DIM_CLIENTE
Dimensión Producto	DIM_PRODUCTO
Dimensión Vendedor	DIM_VENDEDOR
Dimensión Proveedor	DIM_PROVEEDOR
Dimensión Tiempo	DIM_TIEMPO
Dimensión Pago	DIM_PAGO

**Tabla N° 1:** Nombres estándares para la Tabla de Hechos y sus Dimensiones

NOMBRE DE LA TABLA	DISEÑO LÓGICO	DISEÑO FISICO
Dimensión Tiempo	Año	Año
	Mes	Mes
	Semestre	Semestre
Dimensión Cliente	Cliente	Cliente
	RUC	RUC
	Tipo Cliente	TipoCliente

Dimensión Producto	Tipo de Producto	TipoProducto
	Producto	Producto
	Precio	Precio
Dimensión Proveedor	Proveedor	Proveedor
	Dirección	Dirección
Dimensión Vendedor	Vendedor	Vendedor
	Genero	Genero
Dimensión Pago	Tipo de Pago	Tipo_Pago

**Tabla N° 2:** Nombres estándares para los atributos de las dimensiones

<b>NOMBRE DE LA TABLA</b>	<b>DISEÑO LÓGICO</b>	<b>DISEÑO FISICO</b>
Tabla de Hecho FACT_VENTAS	Monto en Soles	Monto
	Numero de Vendedores	NroVendedores
	Cantidad de Productos	CantidadProductos
	Numero de Proveedores	NroProveedores
	Cantidad De Pagos	NroPagos
	Numero de Clientes	NroClientes

**Tabla N° 3:** Nombres estándares para los atributos de las Tablas de Hechos

Determinar el Tipo de Dato para cada Tabla que intervendrá en el Data Mart. Estos cuadros muestran el detalle de cada tabla:

### 1. Dim\_Cliente

ATRIBUTOS	TIPOS DE VALORES		LONGITUD	DESCRIPCION
	DATOS	NULOS		
Cliente_Id	int	no	12	Dimensión que contiene toda la información detallada de los Clientes de Copy Ventas S.R.L.
Cod_Cliente	nchar	no	12	
Cliente	varchar	no	50	
RUC	char	no	11	
TipoCliente	varchar	no	50	

### 2. Dim\_Vendedor

ATRIBUTOS	TIPOS DE VALORES		LONGITUD	DESCRIPCION
	DATOS	NULOS		
Vendedor_Id	int	no	12	Dimensión que contiene toda la información detallada de los Vendedores de Copy Ventas S.R.L.
Cod_Vendedor	char	no	10	
Vendedor	varchar	no	50	
Genero	char	no	10	

### 3. Dim\_Producto

ATRIBUTOS	TIPOS DE VALORES		LONGITUD	DESCRIPCION
	DATOS	NULOS		
Producto_Id	int	no	12	Dimensión que contiene toda la
Cod_Producto	nchar	no	10	

Tipo_Producto	nchar	no	10	información detallada de los Productos de Copy Ventas S.R.L.
Producto	varchar	no	50	
Precio	nchar	no	10	
Cantidad	int	no	4	

#### 4. Dim\_Proveedor

ATRIBUTOS	TIPOS DE VALORES		LONGITUD	DESCRIPCION
	DATOS	NULOS		
Proveedor_Id	int	no	12	Dimensión que contiene toda la información detallada de los Proveedores de Copy Ventas S.R.L.
Cod_Proveedor	char	no	10	
Proveedor	varchar	no	50	
Direccion	varchar	no	50	

#### 5. Dim\_Pago

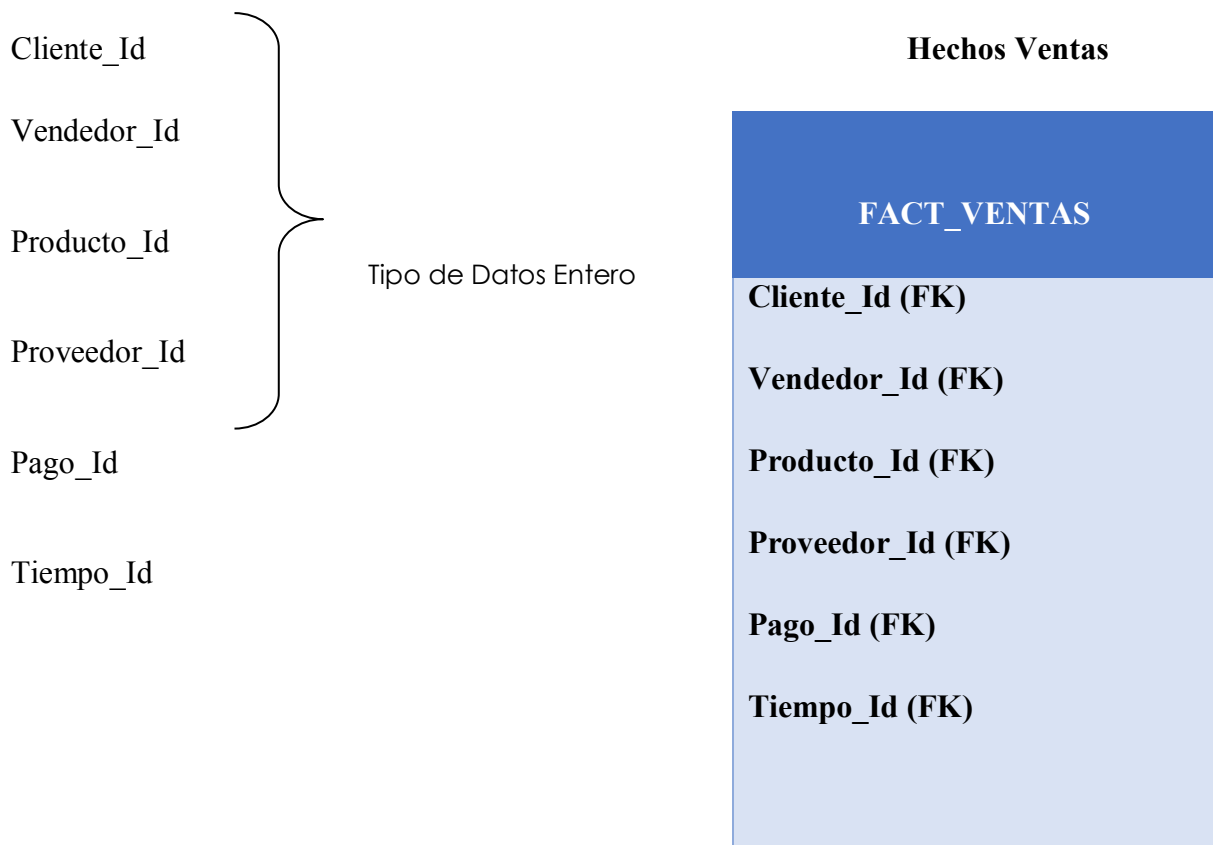
ATRIBUTOS	TIPOS DE VALORES		LONGITUD	DESCRIPCION
	DATOS	NULOS		
Pago_Id	int	no	12	Dimensión que contiene la información de los Pagos de Copy Ventas S.R.L..
Cod_Pago	char	no	10	
Tipo_Pago	char	no	10	

#### 6. Dim\_Tiempo

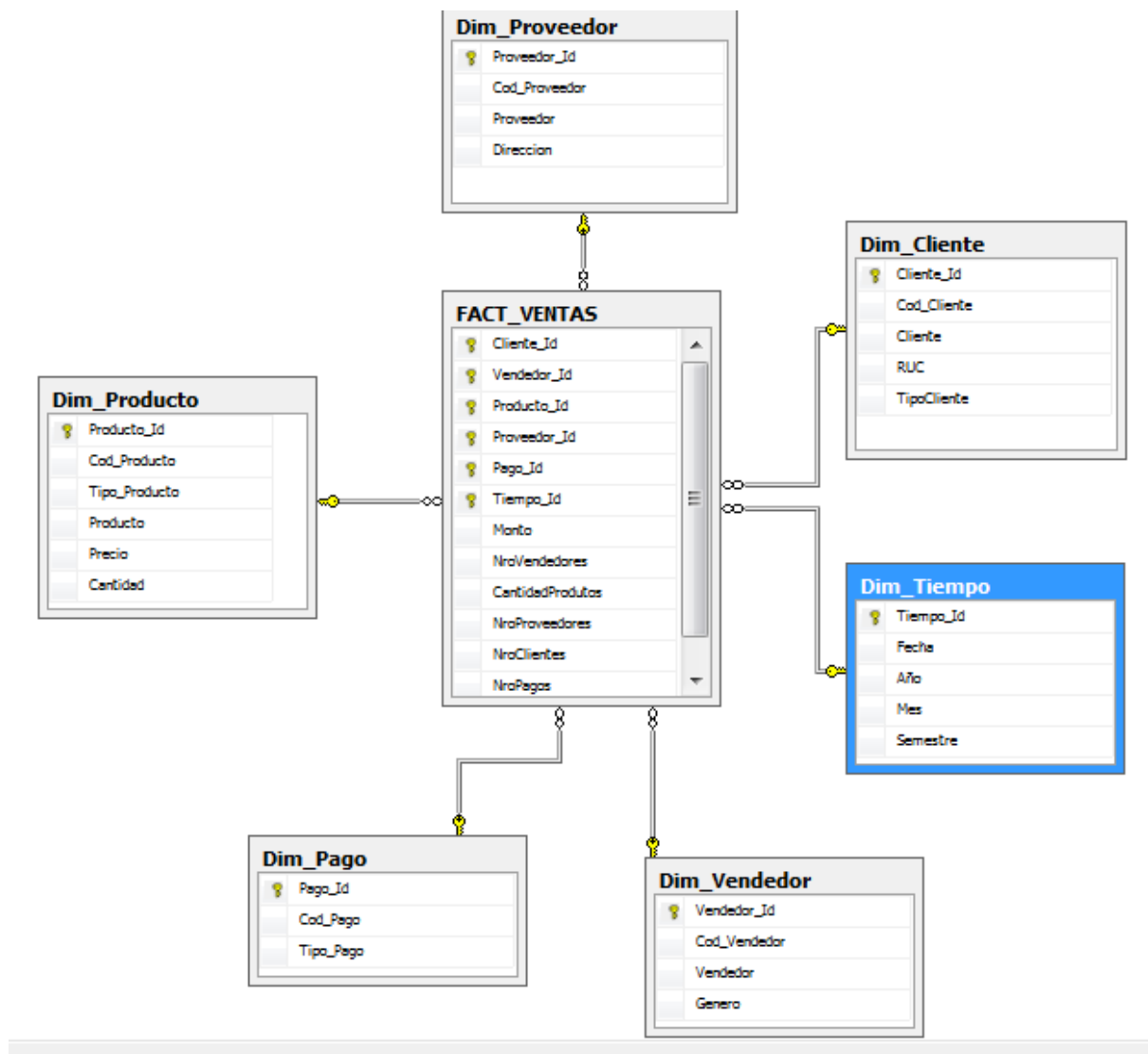
ATRIBUTOS	TIPOS DE VALORES		LONGITUD	DESCRIPCION
	DATOS	NULOS		
Tiempo_Id	int	no	12	

Fecha	date	no	10	Dimensión que contiene detalle del tiempo de las ventas en Copy Ventas S.R.L..
Año	int	no	4	
Mes	varchar	no	50	
Semestre	varchar	no	50	

Determinar el tipo de Datos de las claves Primarias. Para nuestro Data Mart se generaran en forma automática y ordenada:



**Tabla:**Claves foraneas de la tabla de Hechos

**MODELO ESTRELLA: DATA MART VENTAS**

## V.1. Determinación de las agregaciones

TABLA DE HECHOS	MEDIDAS	REGLA DE LA AGREGACION	FORMULA MDX
FACT_VENTAS	MONTO	La suma del precio por la cantidad de productos	SUM(PRECIO*CANTIDAD)
	CANTIDAD DE PROVEEDORES	La de los proveedores que colaboran en la empresa	SUM(CANTIDAD)
	CANTIDAD DE VENDEDORES	La de los vendedores de laboran en la empresa	SUM(CANTIDAD)
	CANTIDAD DE PAGOS	La suma de pagos que se han realizado	SUM(CANTIDAD)
	CANTIDAD DE PRODUCTOS	La suma de los productos con los que cuenta la empresa	SUM(CANTIDAD)
	CANTIDAD DE cLIENTES	La suma de los clientes con los que cuenta la empresa	SUM(CANTIDAD)

## V.2.Construcción de las Tablas y la Base de Datos en SQL

Después de haber realizado el Diseño Físico de las tablas pasamos a la etapa de construcción de dichas tablas utilizando el SQL Server.

**Tabla 1 : Dim\_Cliente**

MUP-G807-40397....dbo.Dim_Cliente		MUP-G807-40397...	
	Nombre de columna	Tipo de datos	Permitir v...
▶	Cliente_Id	int	<input type="checkbox"/>
	Cod_Cliente	char(12)	<input type="checkbox"/>
	Cliente	nvarchar(50)	<input type="checkbox"/>
	RUC	char(10)	<input type="checkbox"/>
	TipoCliente	char(10)	<input type="checkbox"/>
			<input type="checkbox"/>

Figura: Tabla de la Dimensión Cliente

**Tabla 2 : Dim\_Pago**

MUP-G807-40397....C - dbo.Dim_Pago		MUP-G807-40397...	
	Nombre de columna	Tipo de datos	Permitir v...
▶	Pago_Id	int	<input type="checkbox"/>
	Cod_Pago	char(10)	<input type="checkbox"/>
	Tipo_Pago	char(10)	<input type="checkbox"/>
			<input type="checkbox"/>

Figura: Tabla de la Dimensión Pago



**Tabla 3 : Dim\_Producto**

JOSECARLOS.Nort...dbo.Dim_Producto		
Nombre de columna	Tipo de datos	Permitir v...
Producto_Id	int	<input type="checkbox"/>
Cod_Producto	nchar(10)	<input type="checkbox"/>
Tipo_Producto	nchar(10)	<input type="checkbox"/>
Producto	varchar(50)	<input type="checkbox"/>
Precio	nchar(10)	<input type="checkbox"/>
Cantidad	int	<input checked="" type="checkbox"/>

Figura : Tabla de la Dimensión Producto

**Tabla 4 : Dim\_Proveedor**

MUP-G807-40397....o.Dim_Proveedor		
Nombre de columna	Tipo de datos	Permitir v...
Proveedor_Id	int	<input type="checkbox"/>
Cod_Proveedor	char(10)	<input type="checkbox"/>
Proveedor	varchar(50)	<input type="checkbox"/>
Direccion	varchar(50)	<input type="checkbox"/>
		<input type="checkbox"/>

Figura N°: Tabla de la Dimensión Proveedor

**Tabla 5 : Dim\_Tiempo**

MUP-G807-40397... dbo.Dim_Tiempo		
Nombre de columna	Tipo de datos	Permitir v...
Tiempo_Id	int	<input type="checkbox"/>
Fecha	datetime	<input type="checkbox"/>
Año	int	<input type="checkbox"/>
Mes	varchar(50)	<input type="checkbox"/>
Semestre	varchar(50)	<input type="checkbox"/>
		<input type="checkbox"/>

Figura: Tabla de la Dimensión Tiempo

Tabla 6 : Dim\_Vendedor


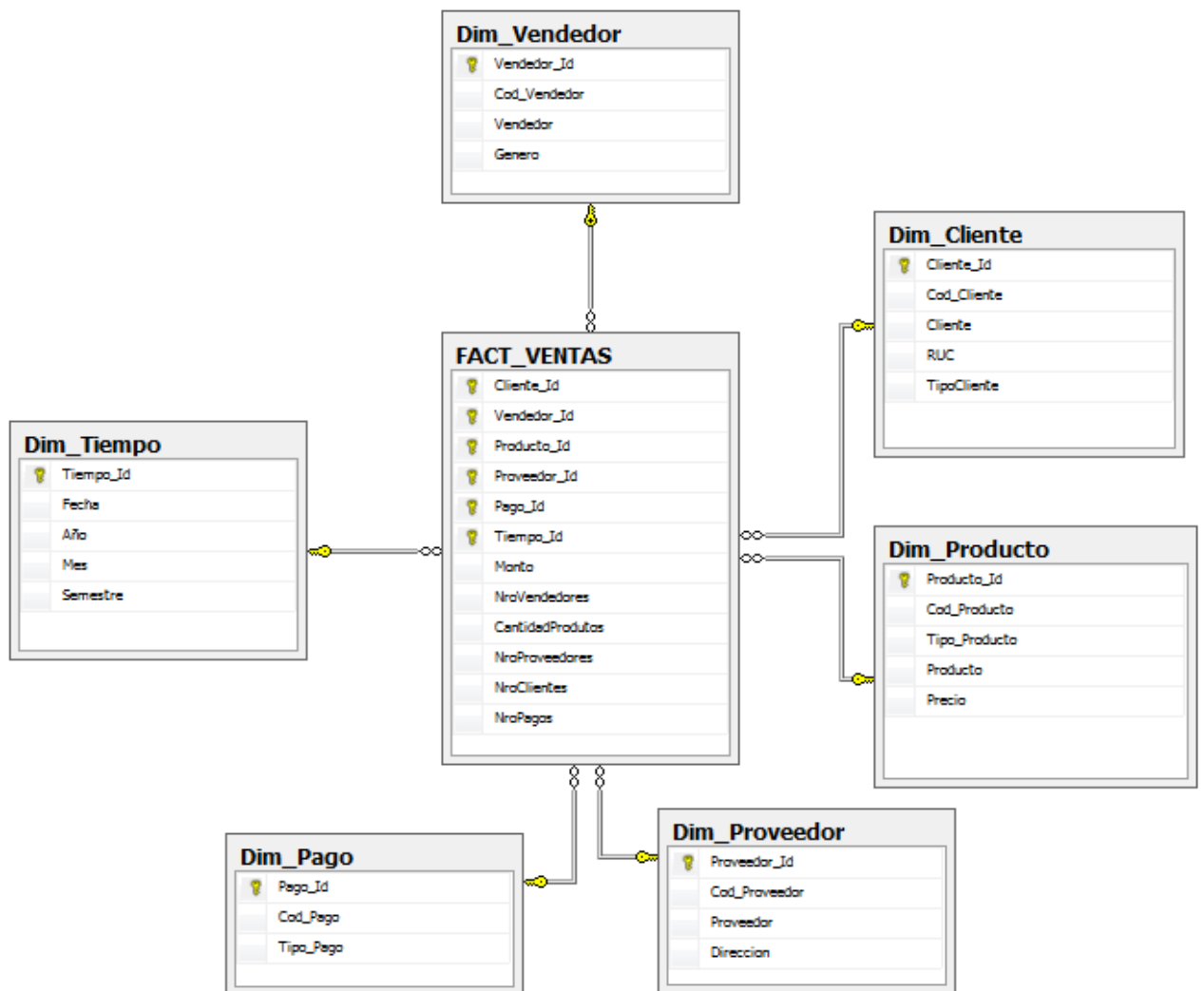
MUP-G807-40397...bo.Dim_Vendedor		MUP-G807-40397	
	Nombre de columna	Tipo de datos	Permitir v...
	Vendedor_Id	int	<input type="checkbox"/>
	Cod_Vendedor	nchar(10)	<input type="checkbox"/>
	Vendedor	varchar(50)	<input type="checkbox"/>
	Genero	char(10)	<input type="checkbox"/>
			<input type="checkbox"/>

Figura N° 44: Tabla de la Dimensión Vendedor

Una vez construido todas las tablas para el Data Mart para la Libreria, continuamos con la construcción del Diagrama de la Base de Datos:



## VI. PROCESO DE EXTRACCIÓN, TRANSFORMACIÓN Y CARGA DE DATOS

Para completar la construcción del Data Mart debemos poblar cada Tabla de la Base de Datos. Como ya conocemos la estructura de cada tabla, ahora debemos seguir los siguientes pasos para el Poblamiento:

- Definición de los Pasos de Transformación.
- Definición de los Workflows.
- Creación de los Paquetes de Servicio de Transformación de Datos (DTS).

### A. Definición de los Pasos de Transformación

Para un correcto Poblamiento de los datos debemos definir los Pasos de Transformación, con la unidad de trabajo como parte de un proceso de transformación. Para poblar el Data Mart se tiene que realizar los siguientes pasos:

**Paso 1:** Limpiando Dimensiones, consiste en eliminar los datos de todas las Dimensiones y Tablas de Hechos, paso que nos permite asegurar de que no pueda existir algún dato que se pudiera duplicar.

**Paso 2:** Poblamiento de la Dimensión Pago, consiste en ejecutar una sentencia SQL, para transferir datos a la Dimensión Pago.

**Paso 3:** Poblamiento de la Dimensión Cliente, consiste en mover los datos de la Tabla Agencia para poblar la Dimensión Cliente.

**Paso 4:** Poblamiento de la Dimensión Tiempo, consiste en ejecutar una sentencia SQL, para transferir datos a la Dimensión Tiempo

**Paso 5:** Poblamiento de la Dimensión Producto, consiste en ejecutar una sentencia SQL, para transferir datos a la Dimensión Producto.

**Paso 6:** Poblamiento de la Dimensión Vendedor, consiste en ejecutar una sentencia SQL, para transferir datos a la Dimensión Vendedor.

**Paso 7:** Poblamiento de la Dimensión Proveedor, consiste en ejecutar una sentencia SQL, para transferir datos a la Dimensión Proveedor.

**Paso 8:** Poblamiento de las Tabla de Hechos.

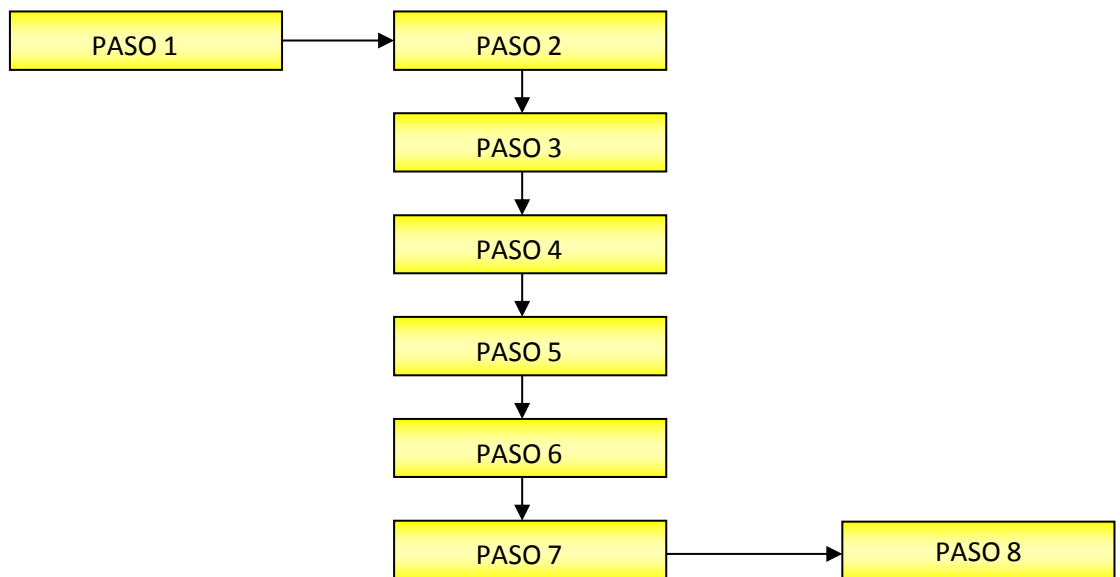
**B. Definición de los Workflows**

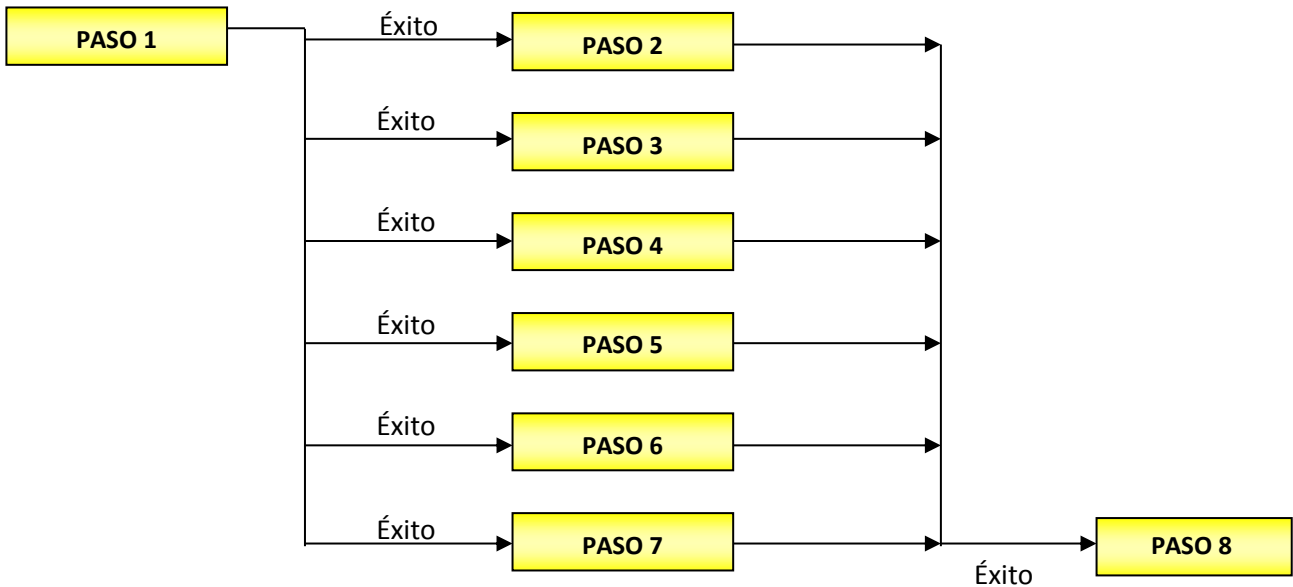
Figura N° 48: Workflow de los Pasos de Transformación

**Restricciones de Precedencia:**

- a) La Limpieza de las Dimensiones (Paso 1) debe realizarse al inicio del proceso.
- b) El Poblamiento de la Dimensión Pago (Paso 2) debe realizarse sólo cuando se tenga la seguridad de que el paso 1 se ha ejecutado con éxito.
- c) El Poblamiento de la Dimensión Cliente (Paso 3) debe realizarse sólo cuando se tenga la seguridad de que el paso 1 se ha ejecutado con éxito.
- d) El Poblamiento de la Dimensión Tiempo (Paso 4) debe realizarse sólo cuando se tenga la seguridad de que el paso 1 se ha ejecutado con éxito.
- e) El Poblamiento de la Dimensión Producto (Paso 5) debe realizarse sólo cuando se tenga la seguridad de que el paso 1 se ha ejecutado con éxito.
- f) El Poblamiento de la Dimensión Vendedor (Paso 6) debe realizarse sólo cuando se tenga la seguridad de que el paso 1 se ha ejecutado con éxito.
- g) El Poblamiento de la Dimensión Proveedor (Paso 7) debe realizarse sólo cuando se tenga la seguridad de que el paso 1 se ha ejecutado con éxito.

- h) El Poblamiento de las Tablas de Hechos (Ventas) (Paso 7) debe realizarse sólo cuando se tenga la seguridad de que el paso 2,3,4,5,6,7 se hallan ejecutado con éxito.

Una vez conocido las relaciones de precedencias diagramaremos el workflows que se necesitará realizar al construir el paquete de poblamiento del Data Mart, para la Librería Copy Ventas S.R.L..



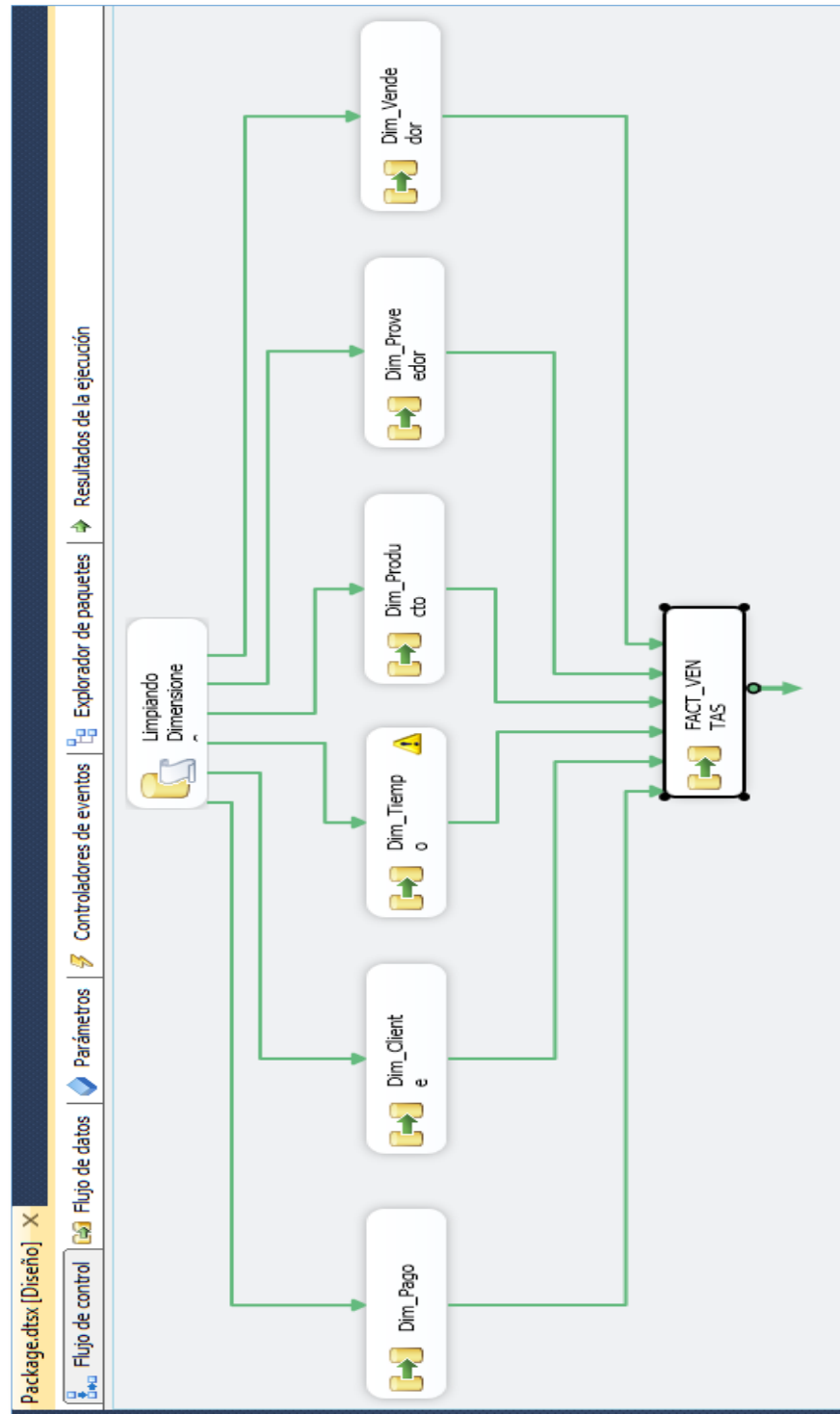
**Figura N° 49:** Diagrama Workflow con Restricciones de Precedencia

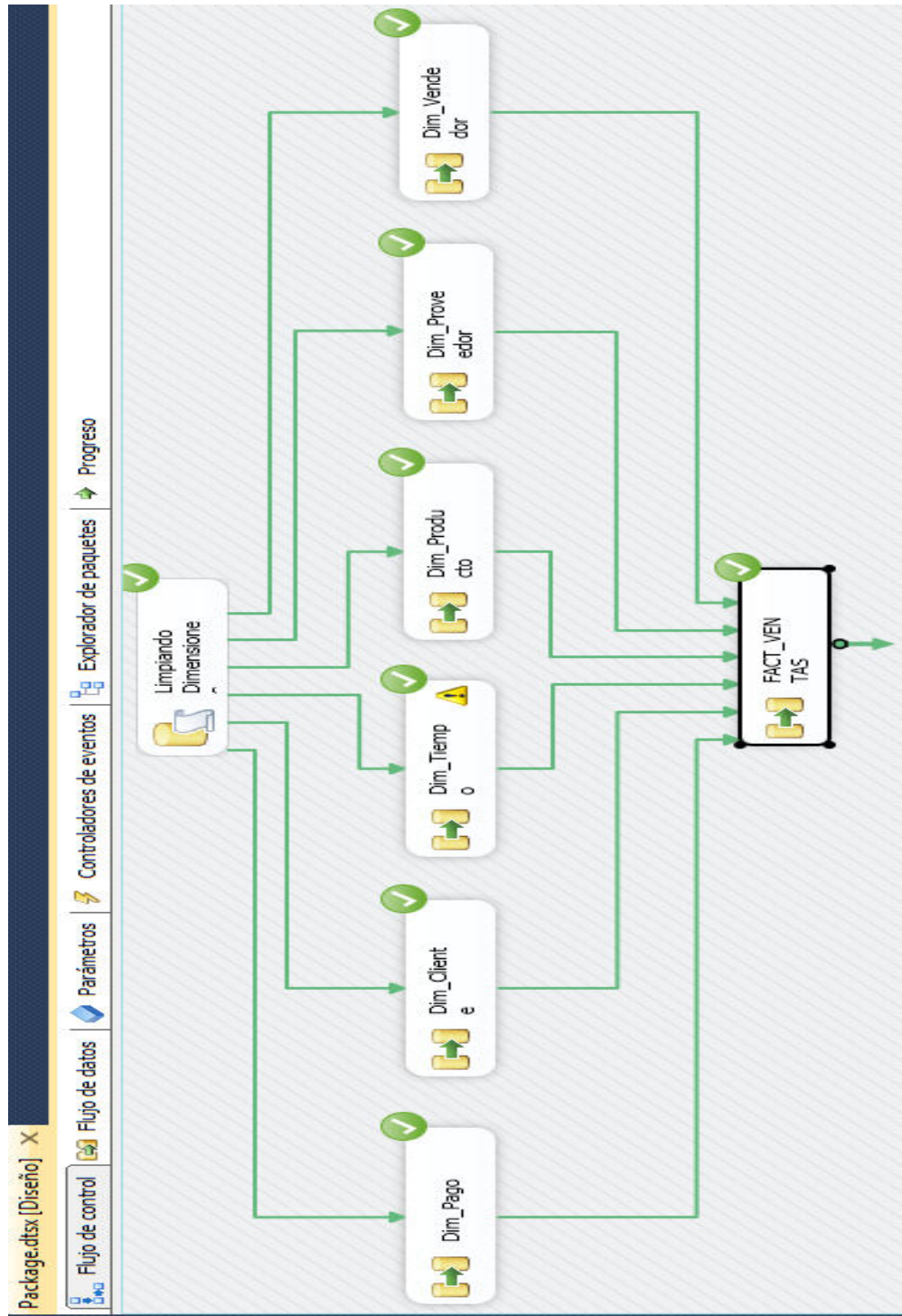
### C. Creación de los Paquetes de Servicio de Transformación de Datos

Definidos los pasos de la transformación de datos y las restricciones de precedencia, podemos crear el paquete Servicio de Transformación de Datos (DTS); recordaremos que un DTS tiene como objetivos importar, exportar y realizar cambios en el formato de datos.

Para su construcción utilizaremos el Servicio de Transformación de Datos SQL Server 2008 R2, aquí los datos pueden ser almacenado en varios formatos y en muchos lugares diferentes lo cual no es ningún problema.

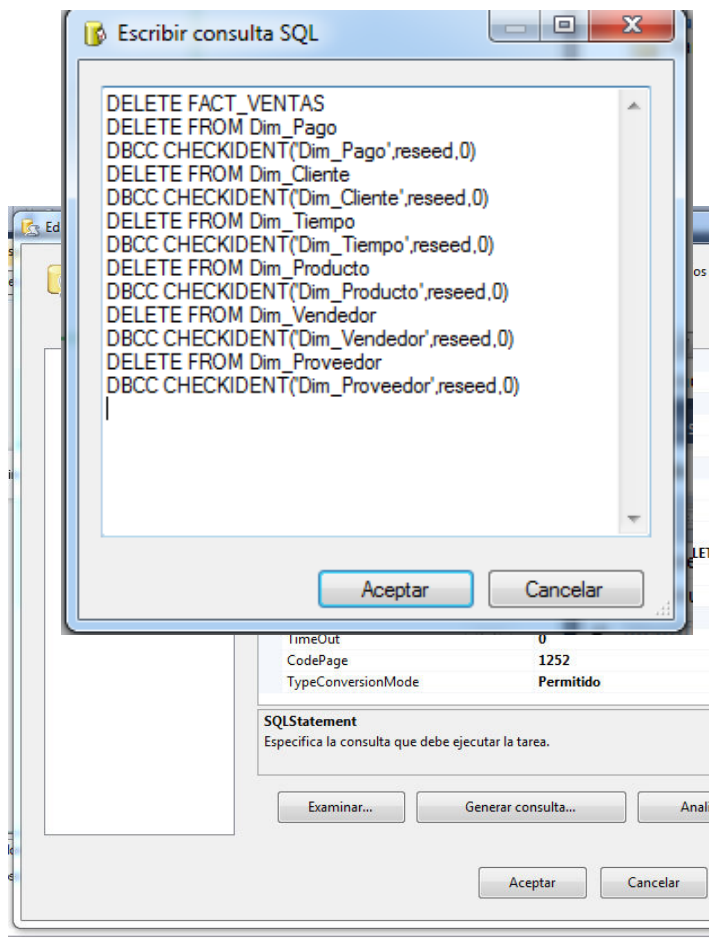
Basándonos en Workflow con restricciones de la figura anterior, creamos el paquete DTS al cual llamamos “Poblamiento Copy Ventas”, con los pasos previos definidos así como todas sus restricciones.





## Paso 1: Limpieza de todas las dimensiones

Esta sentencia nos permite limpiar los datos de todas las tablas de nuestro DataMart. Esta limpieza asegura que no se dupliquen los datos.



```

DELETE FACT_VENTAS
DELETE FROM Dim_Pago
DBCC CHECKIDENT('Dim_Pago',reseed,0)
DELETE FROM Dim_Cliente
DBCC CHECKIDENT('Dim_Cliente',reseed,0)
DELETE FROM Dim_Tiempo
DBCC CHECKIDENT('Dim_Tiempo',reseed,0)
DELETE FROM Dim_Producto
DBCC CHECKIDENT('Dim_Producto',reseed,0)
DELETE FROM Dim_Vendedor
DBCC CHECKIDENT('Dim_Vendedor',reseed,0)
DELETE FROM Dim_Proveedor
DBCC CHECKIDENT('Dim_Proveedor',reseed,0)

```

DELETE FACT\_VENTAS

DELETE FROM Dim\_Pago

DBCC CHECKIDENT('Dim\_Pago',reseed,0)

DELETE FROM Dim\_Cliente

DBCC CHECKIDENT('Dim\_Cliente',reseed,0)

DELETE FROM Dim\_Tiempo

DBCC CHECKIDENT('Dim\_Tiempo',reseed,0)

DELETE FROM Dim\_Producto

DBCC CHECKIDENT('Dim\_Producto',reseed,0)

DELETE FROM Dim\_Vendedor

DBCC CHECKIDENT('Dim\_Vendedor',reseed,0)

DELETE FROM Dim\_Proveedor

DBCC CHECKIDENT('Dim\_Proveedor',reseed,0)

## Paso 2: Poblamiento de la Dimensión Pago

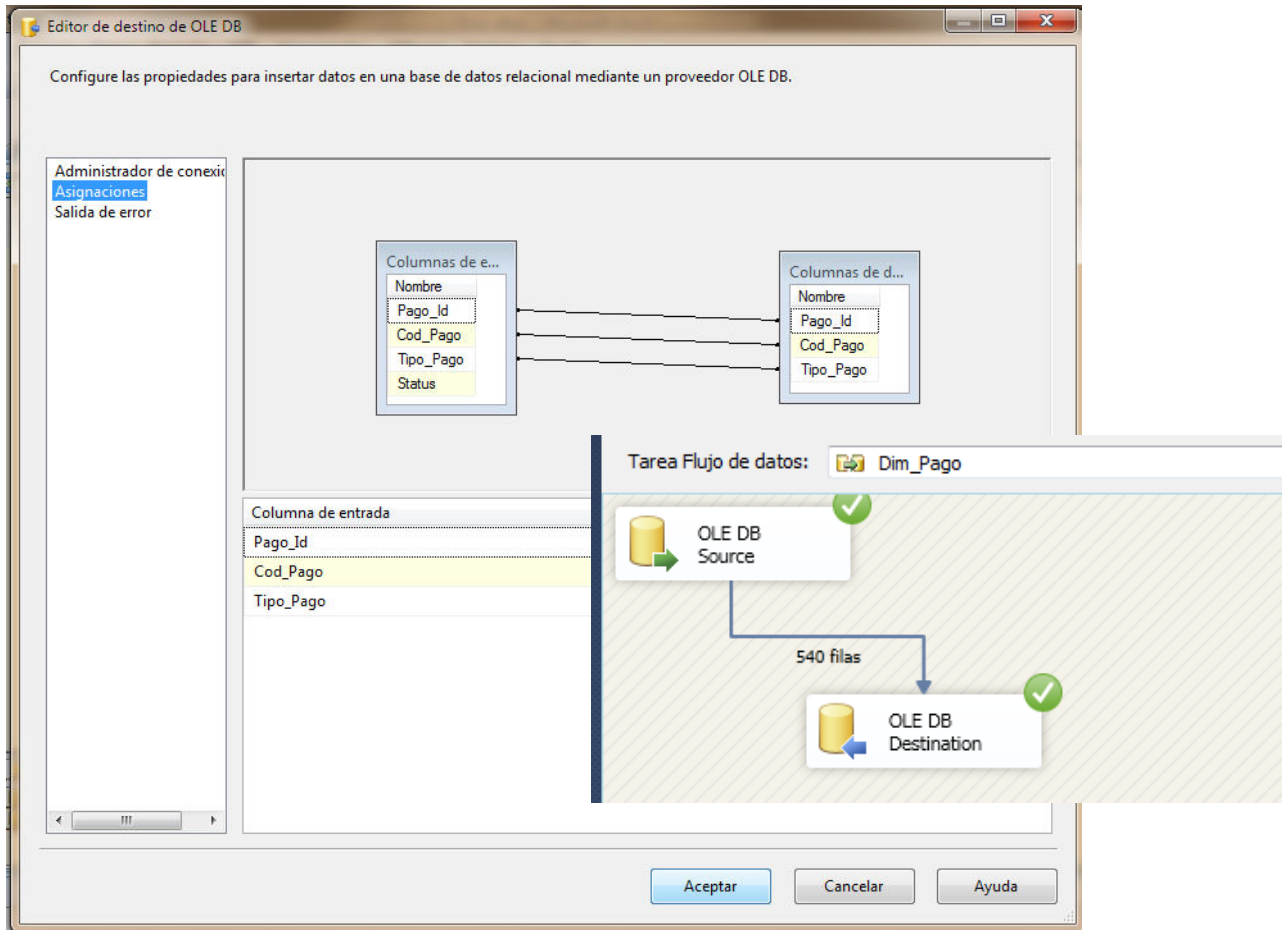
Para transferir los datos a la dimensión Pago, se efectúan los siguientes pasos:

- Se define la tabla Pago como origen de datos de la Base de Datos Operacional de la Empresa.



- Determinamos la Dimensión Pago como destino de datos.

En las transformaciones se determina que columnas del origen corresponderán con las del destino.



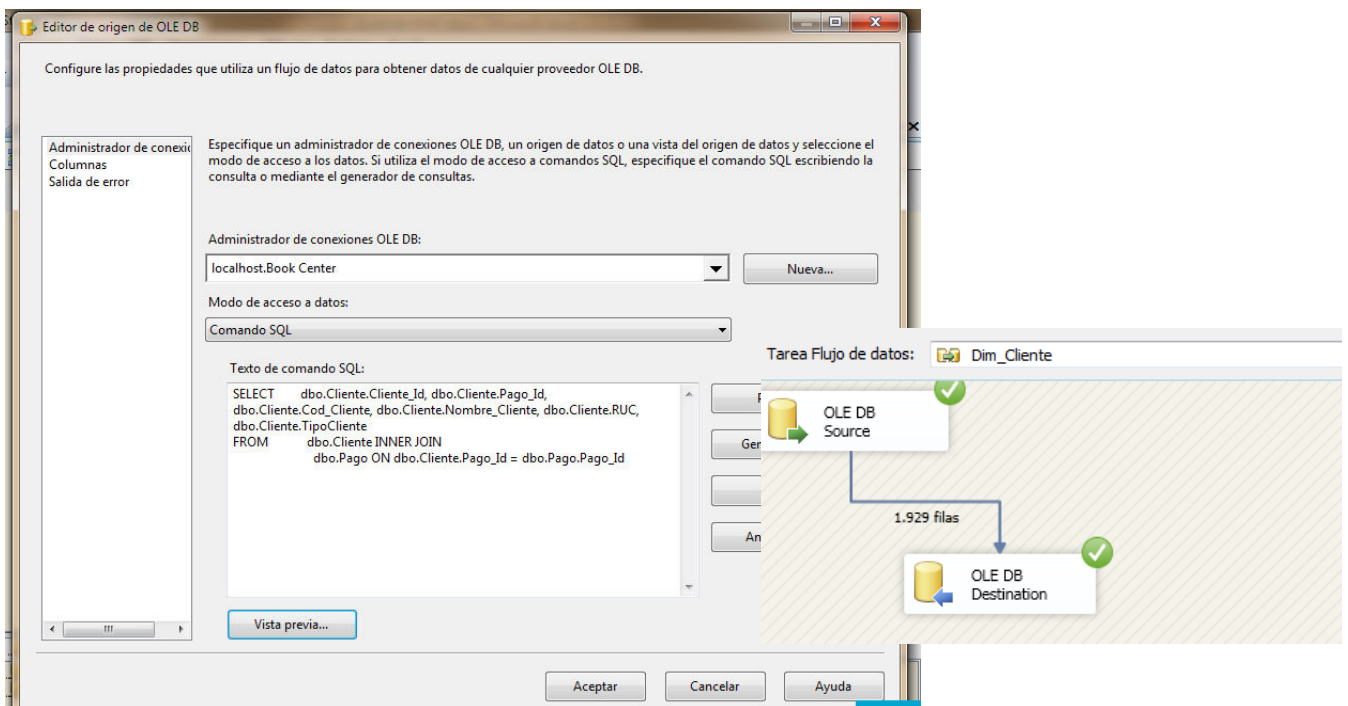
### Paso 3: Poblamiento de la Dimensión Cliente

Para transferir los datos a la dimensión Cliente, se efectúan los siguientes pasos:

- Definimos una sentencia SQL como origen de datos.

```
SELECT dbo.Cliente.Cliente_Id, dbo.Cliente.Pago_Id,
dbo.Cliente.Cod_Cliente, dbo.Cliente.Nombre_Cliente,
dbo.Cliente.RUC, dbo.Cliente.TipoCliente
FROM   dbo.Cliente INNER JOIN
        dbo.Pago ON dbo.Cliente.Pago_Id = dbo.Pago.Pago_Id
```

- Determinamos la Dimensión Cliente como destino de datos.
- En las transformaciones se determina que columnas del origen corresponderán con las de la Dimensión Cliente



#### Paso 4: Poblamiento de la Dimensión Tiempo

- Para transferir los datos a la dimensión Tiempo, se efectúan los siguientes pasos:
- Se define la tabla Tiempo como origen de datos de la Base de Datos Operacional de la Empresa.

```

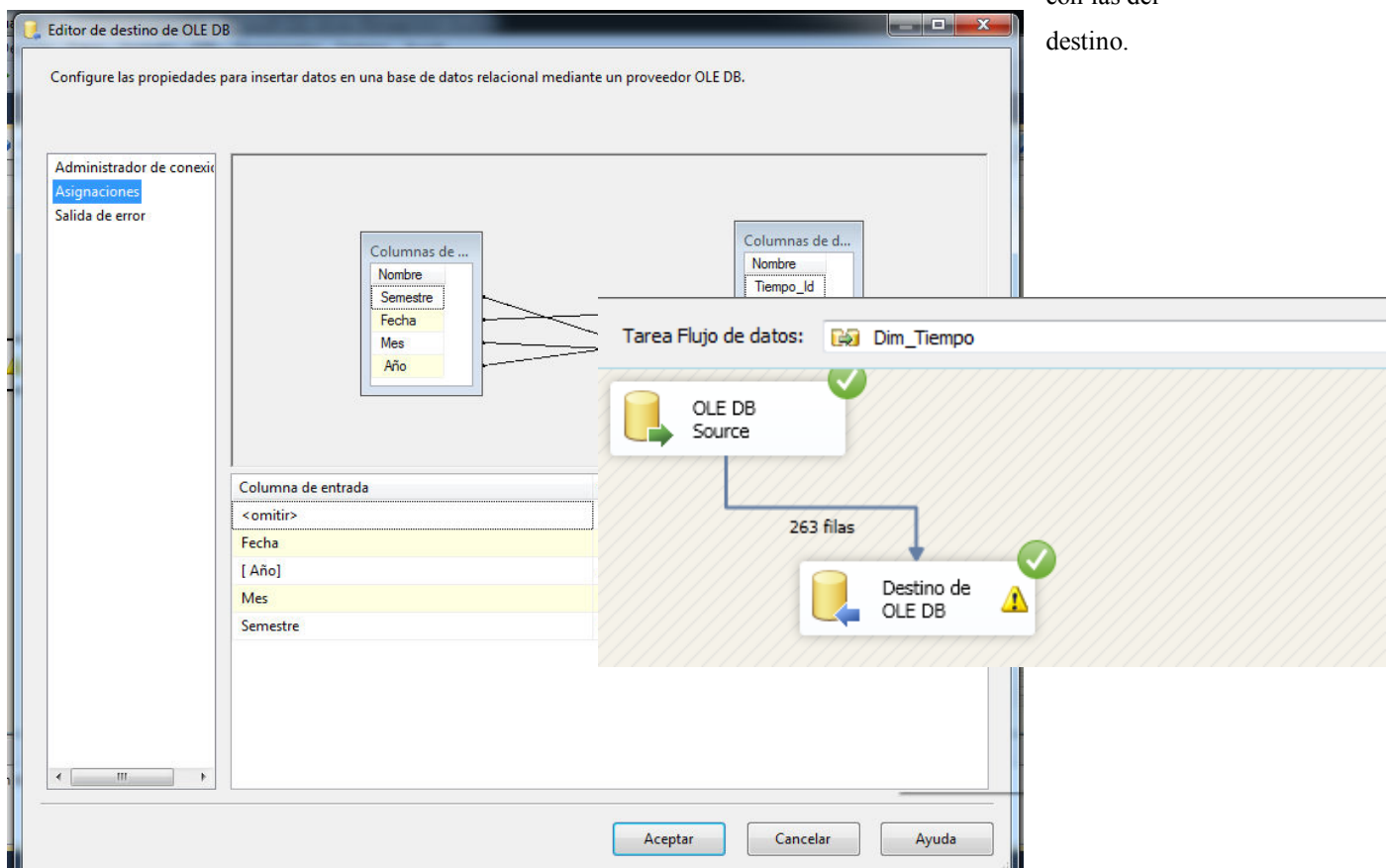
SELECT DISTINCT
T.TheDate AS Fecha,
DatePart(mm, T.TheDate) AS [Mes], DATENAME(MM,T.TheDate) AS
[Nombre Mes],
floor ((DatePart(mm, T.TheDate) -(1))/(6))+1 AS [Semestre],
DatePart(yy,T.TheDate) AS [ Año]
FROM Tiempo T

```

- Determinamos la Dimensión Tiempo como destino de datos.

- En las transformaciones se determina que columnas del origen corresponderán

con las del destino.

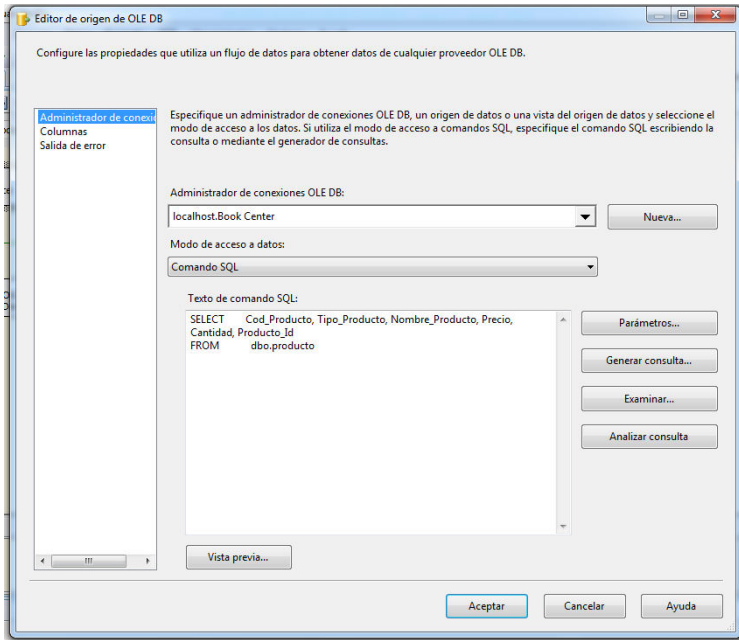


### Paso 5: Poblamiento de la Dimensión Producto

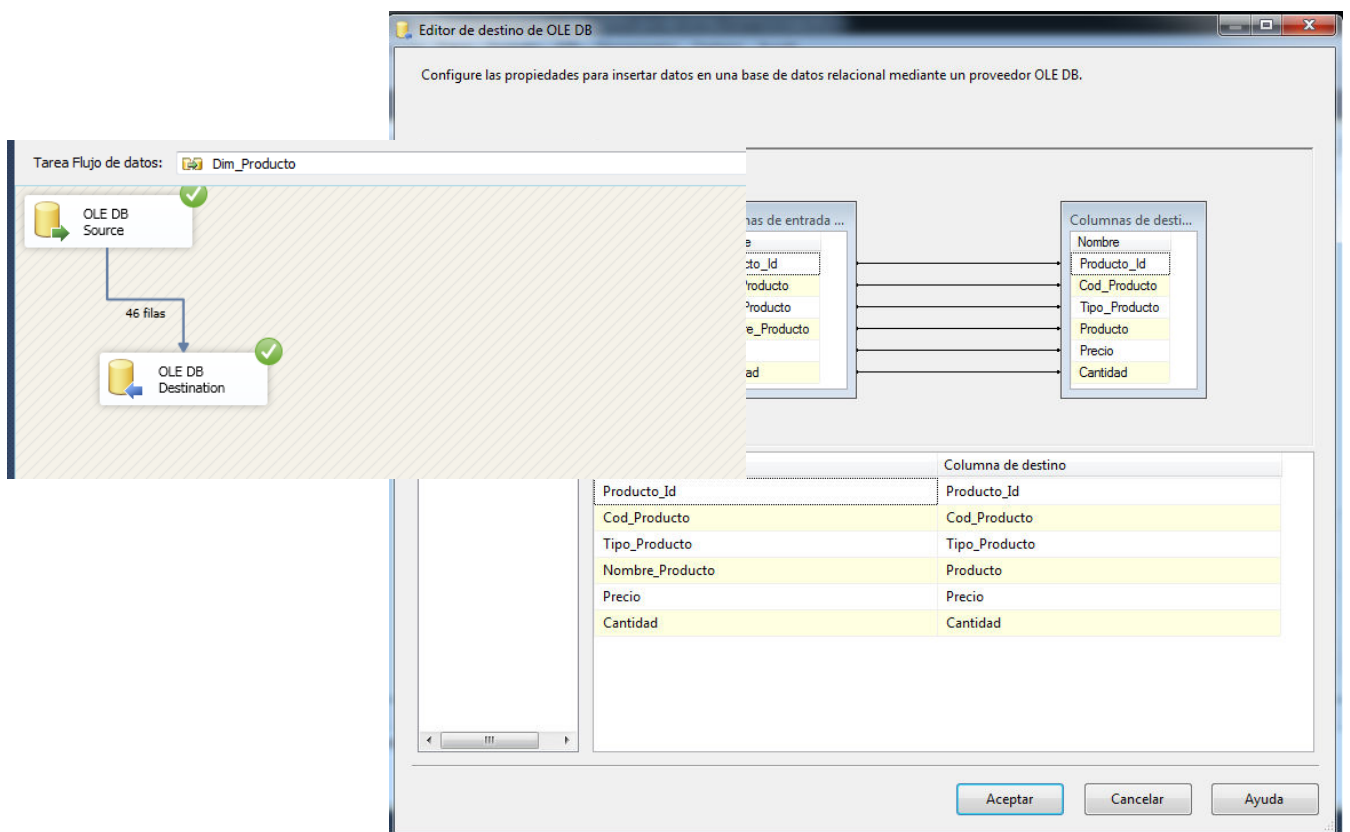
Para transferir los datos a la dimensión Producto, se efectúan los siguientes pasos:

- Definimos una sentencia SQL como origen de datos.  

```
SELECT    Cod_Producto, Tipo_Producto,
          Nombre_Producto, Precio, Cantidad
FROM      dbo.producto
```
- Determinamos la Dimensión Producto como destino de datos.
- En las transformaciones se determina que columnas del origen corresponderán con las de la Dimensión Producto



```
SELECT    Cod_Producto,
Tipo_Producto,
Nombre_Producto, Precio,
Cantidad
FROM      dbo.producto
```

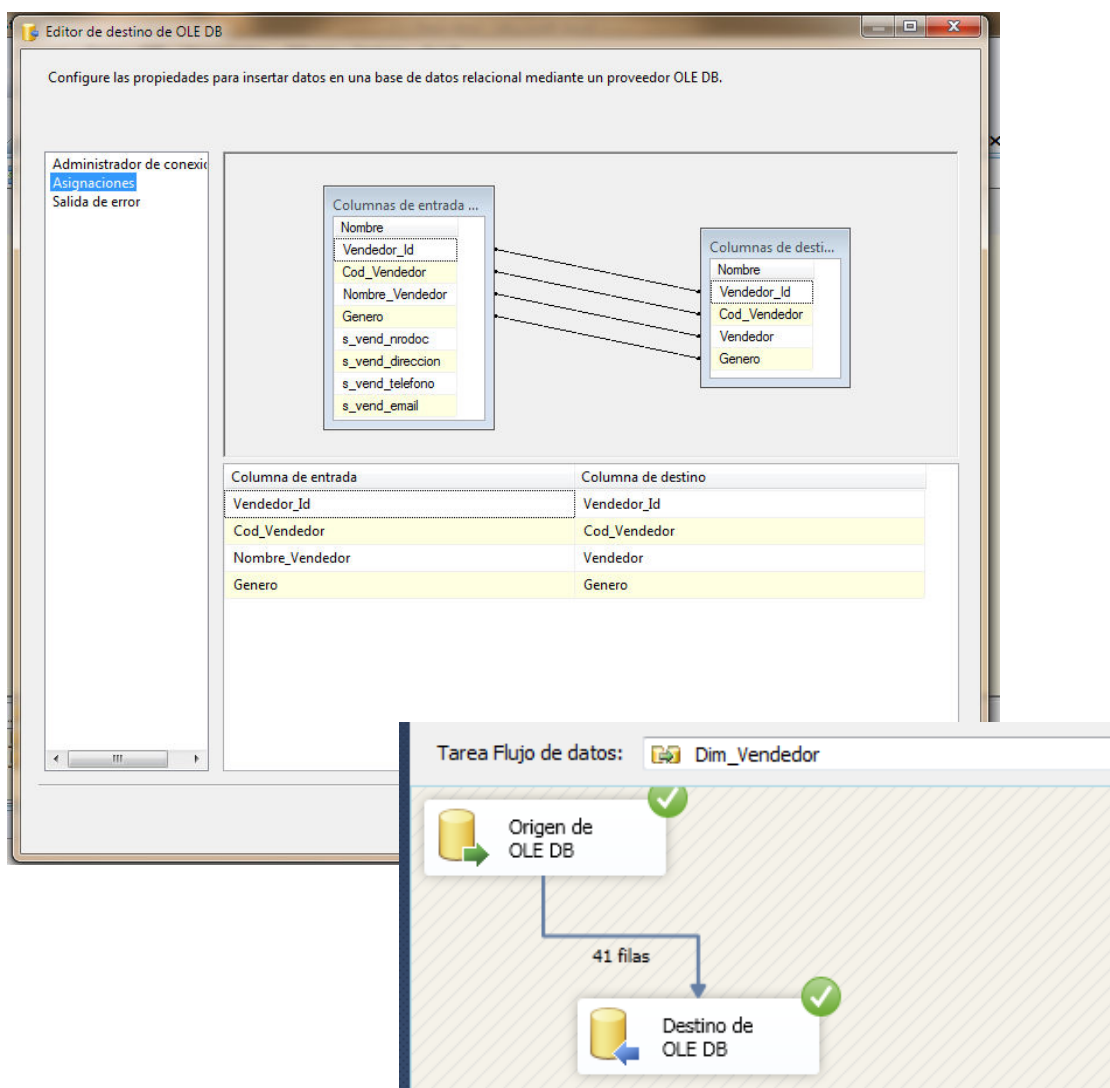


## Paso 6: Poblamiento de la Dimensión Vendedor

Para transferir los datos a la dimensión Vendedor, se efectúan los siguientes pasos:

- Se define la tabla Vendedor como origen de datos de la Base de Datos Operacional de la Empresa.
- Determinamos la Dimensión Vendedor como destino de datos.

En las transformaciones se determina que columnas del origen corresponderán con las del destino.

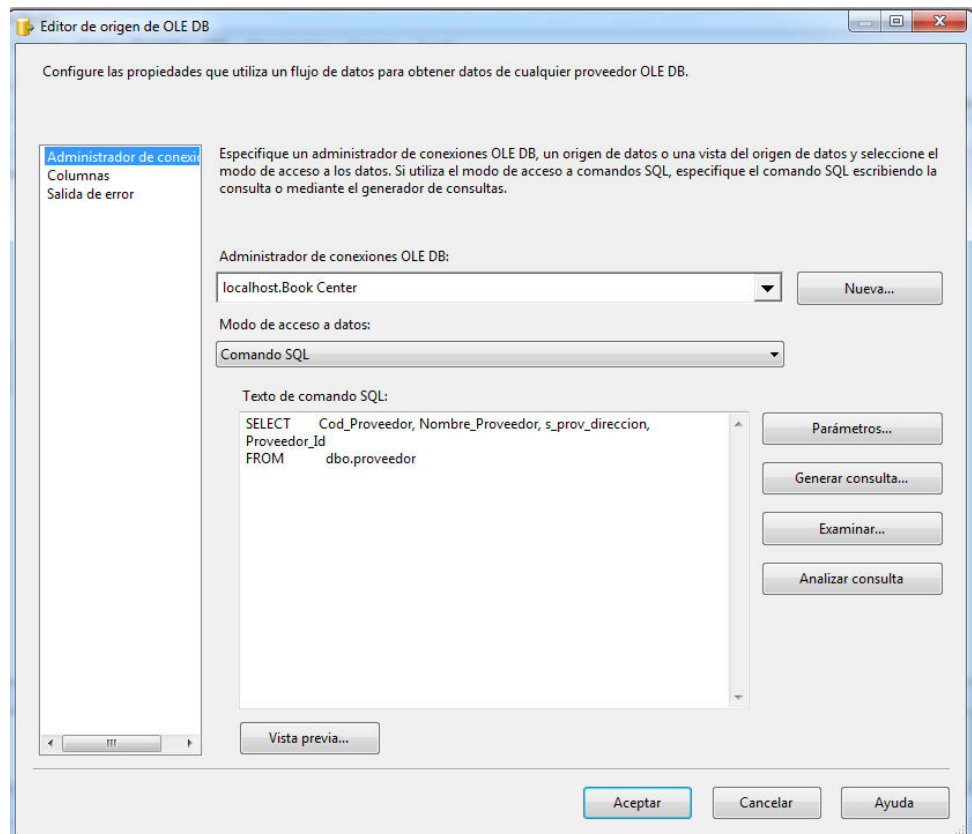


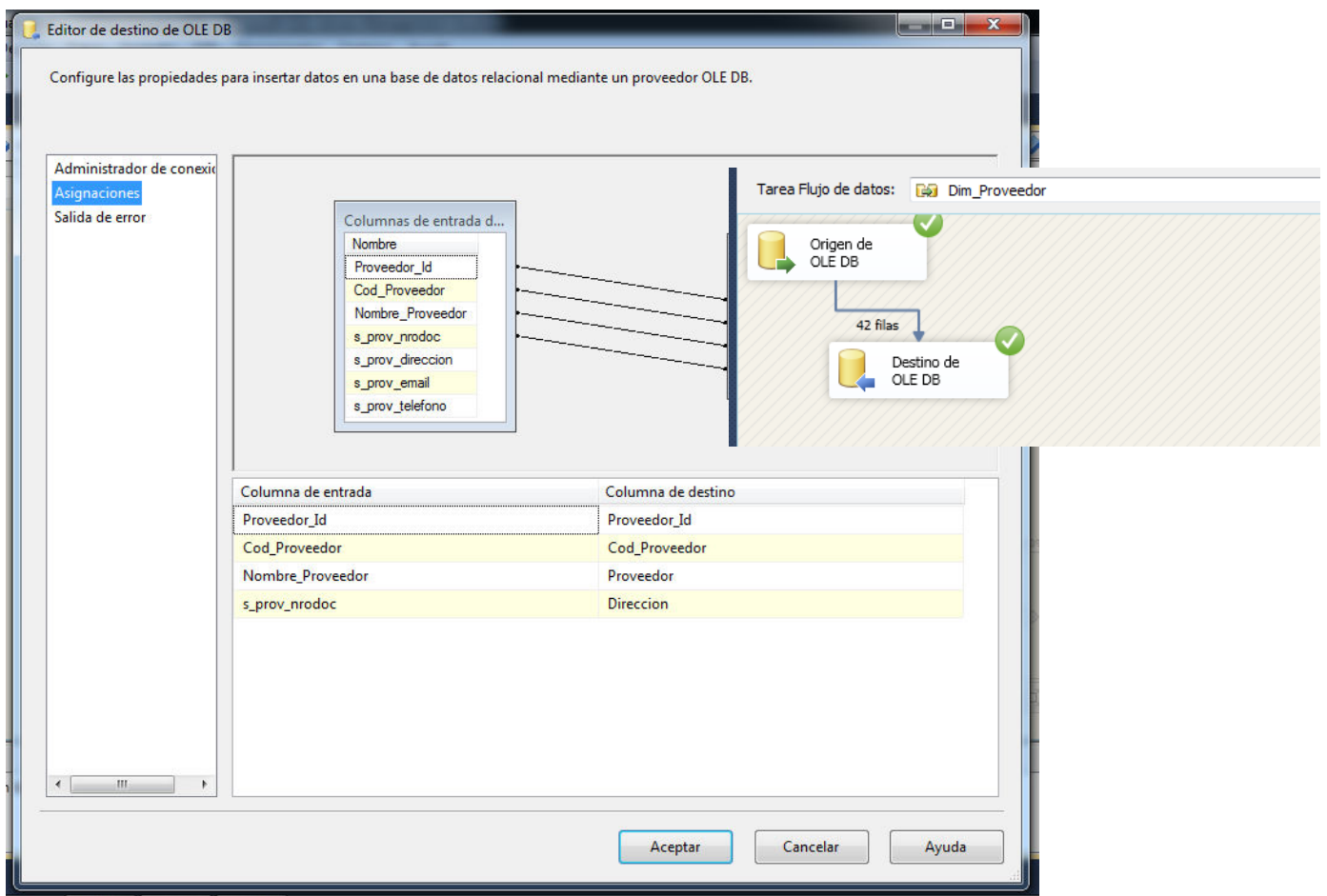
### Paso 7: Poblamiento de la Dimensión Proveedor

Para transferir los datos a la dimensión Proveedor, se efectúan los siguientes pasos:

- Definimos una sentencia SQL como origen de datos.  

```
SELECT    Cod_Proveedor, Nombre_Proveedor, s_prov_direccion,  
Proveedor_Id  
FROM      dbo.proveedor
```
- Determinamos la Dimensión Proveedor como destino de datos.
- En las transformaciones se determina que columnas del origen corresponderán con las de la Dimensión Proveedor





### Paso 8: Poblamiento de FACT\_VENTAS

SELECT C.Cliente\_Id,P.Producto\_Id,

V.Vendedor\_Id,Pro.Proveedor\_Id,

T.Tiempo\_Id,Pa.Pago\_Id,

Convert(int,sum(P.precio\*P.Cantidad)) as Monto,

Count (Pro.Proveedor\_Id) as NroProveedores,

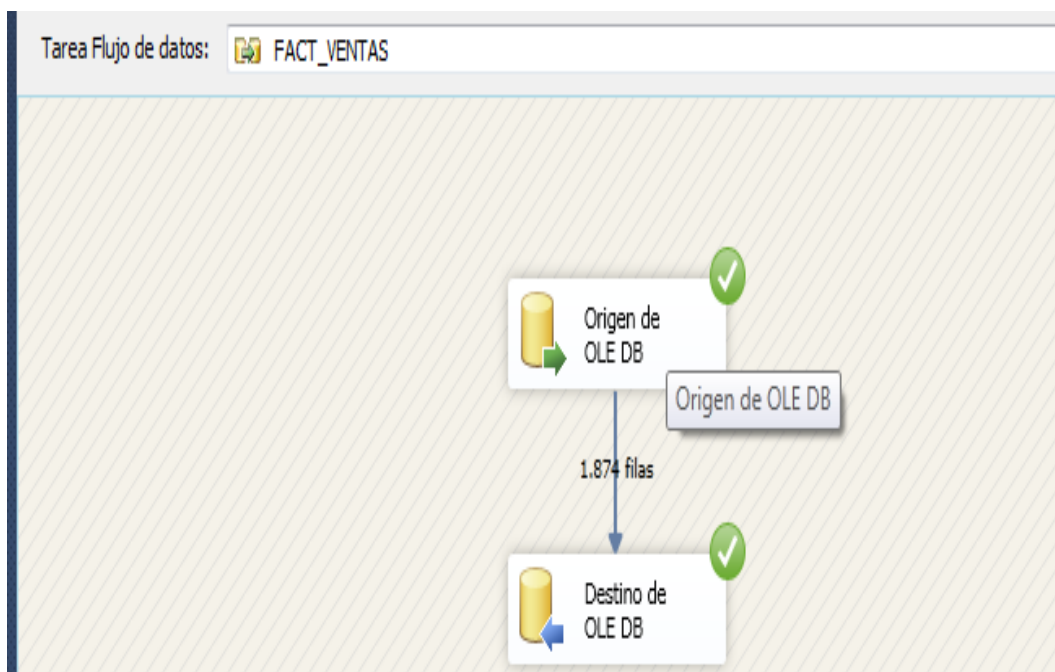
Count (V.Vendedor\_Id) as CantidadDeVendedores,

Count (c.Cliente\_Id) as NroClientes,

Convert (int,sum(P.Cantidad)) AS CantidadProducto,

Count (pa.Pago\_Id) as NroPagos

```
FROM [Copy Ventas].dbo.Cliente CL INNER JOIN [Copy Ventas].dbo.Pedido PS ON  
  
    CL.Pedido_Id = PS.Pedido_Id INNER JOIN [Copy Ventas].dbo.Venta VT ON  
  
    Vt.Venta_Id= CL.Venta_Id  
  
INNER JOIN [North BookC].dbo.Dim_Cliente C on C.Cliente_Id=CL.Cliente_Id  
  
INNER JOIN [North BookC].dbo.Dim_Vendedor V on  
V.Vendedor_Id=PS.Vendedor_Id  
  
INNER JOIN [North BookC].dbo.Dim_Producto P on P.Producto_Id=PS.Producto_Id  
  
INNER JOIN [North BookC].dbo.Dim_Proveedor Pro on  
Pro.Proveedor_Id=PS.Proveedor_Id  
  
INNER JOIN [North BookC].dbo.Dim_Tiempo T ON T.Tiempo_Id=VT.Time_Id  
  
INNER JOIN [North BookC].dbo.Dim_Pago Pa ON Pa.Cod_Pago=Cl.Pago_Id  
  
GROUP BY C.Cliente_Id,P.Producto_Id,V.Vendedor_Id,Pro.Proveedor_Id,T.
```





## VII. SELECCIÓN DE PRODUCTOS




### VII.1. Hardware

La base de datos y los servidores OLAP deben de tener la siguiente configuración de hardware:

- a. RAM: 12GB
- b. Disco Duro (espacio): 500GB.
- c. Procesador: Intel Core i3 -3110M de 2.40 GHz.

### VII.2. Software

Para la implementación se trabajará con MS SQL Server 2012.

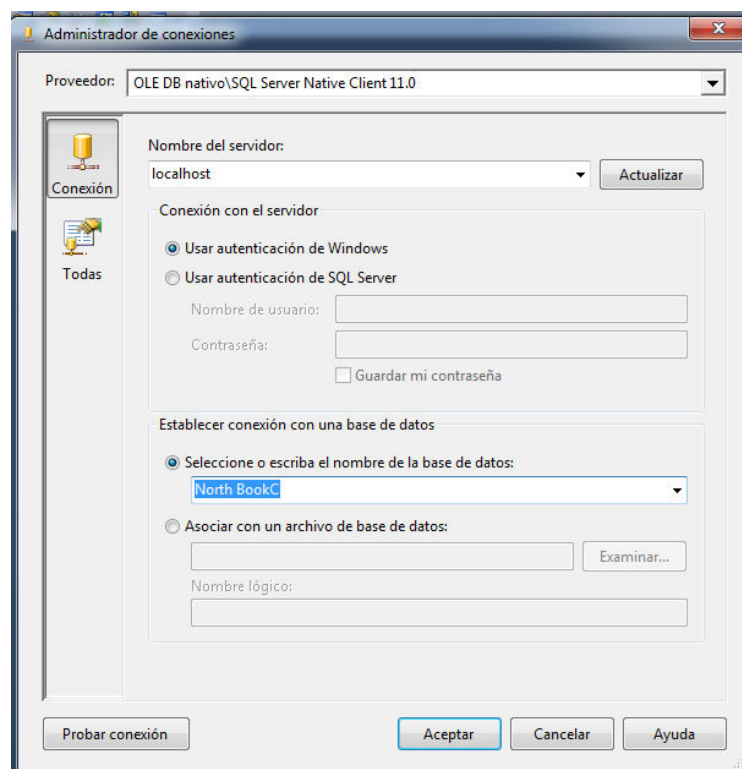
Componente del Proceso	Herramienta
Construcción de Interfaces	 MS Excel 2013
Cubos	 Analysis Services SQL Server Business Intelligence
ETL	
DataMart	 SQL Server 2014

## VIII. ESPECIFICACIÓN DE LA APLICACIÓN DEL USUARIO FINAL

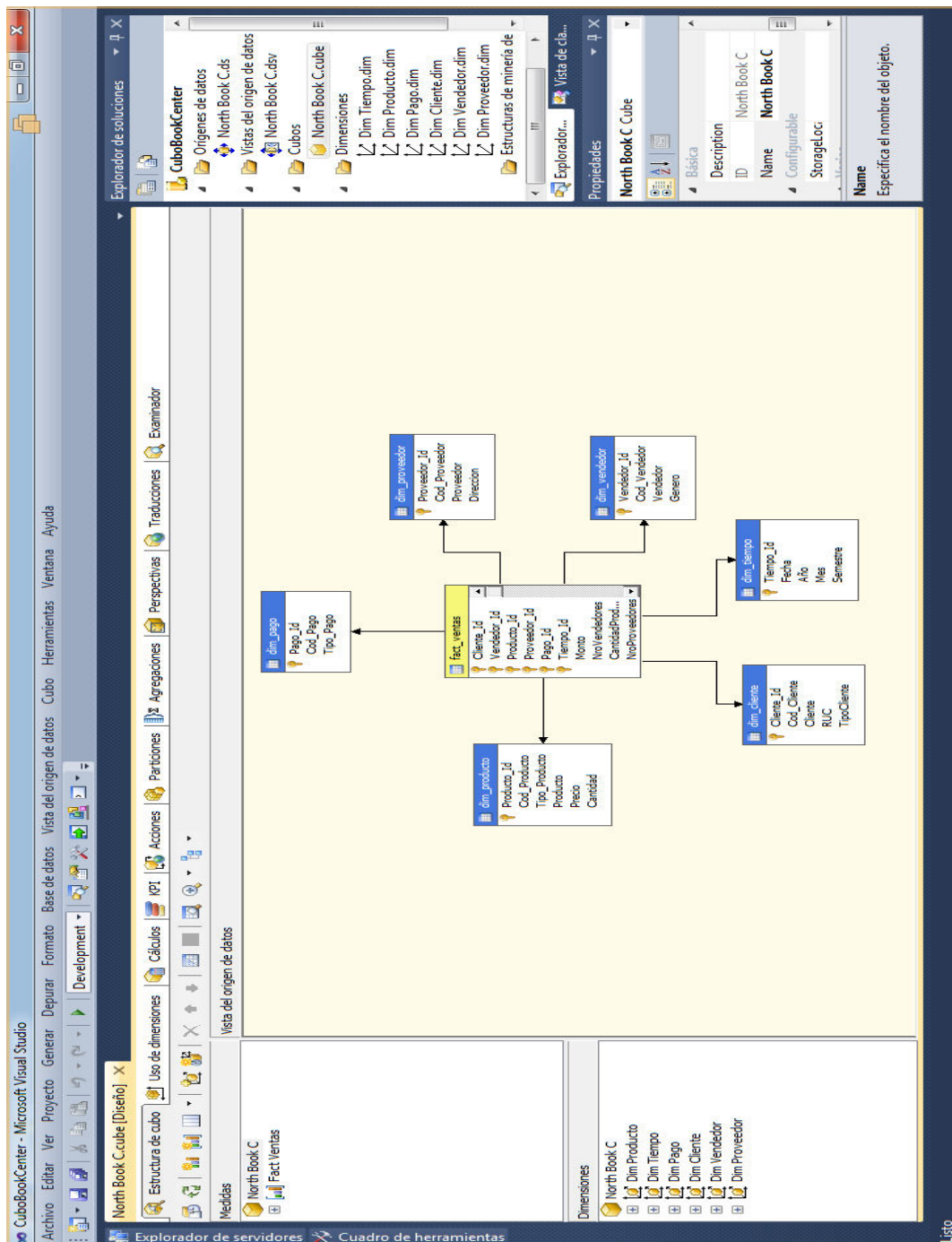
### VIII.1. Estructura de Cubo

La herramienta utilizada para el diseño del cubo es el SSAS (SQL Server Analysis Services) que es el diseñador de cubos de Business Intelligence Development Studio para crear el cubo, incluido las medidas, las dimensiones y sus respectivas relaciones.

Primero definimos el nombre del servidor en este caso “localhost“ y luego seleccionamos la base de dato (datawarehouse)

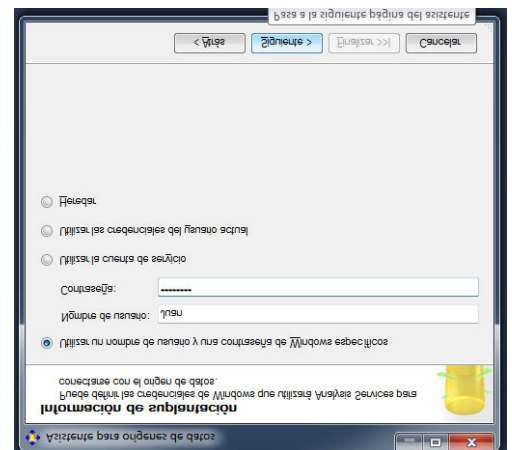
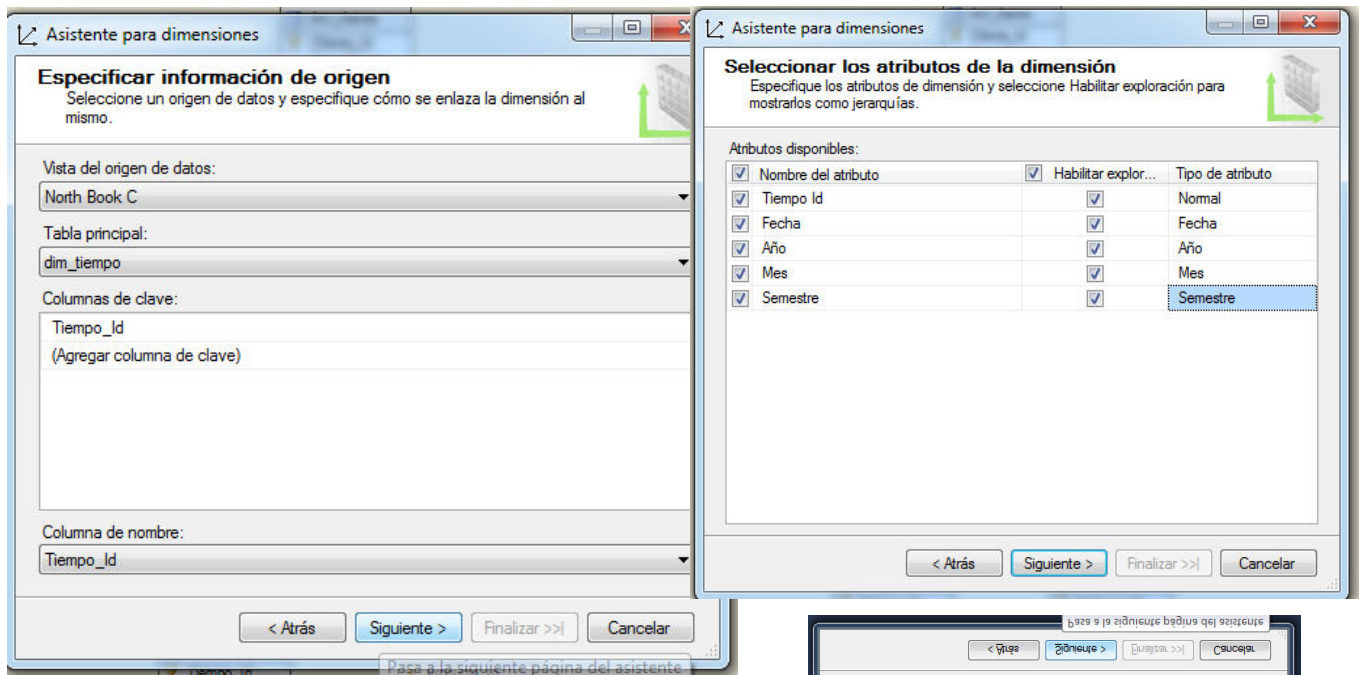


Luego en la información de suplantación colocaremos el nombre de usuario del servidor y su respectiva contraseña

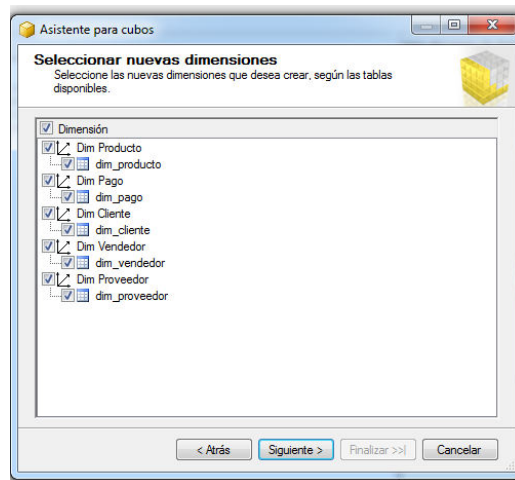
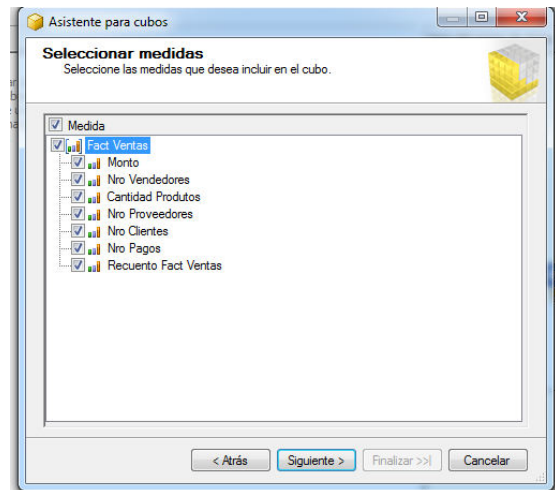
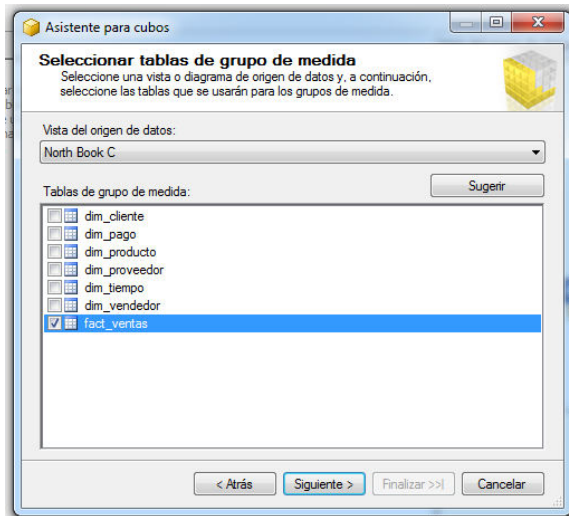


- a) Se realiza la conexión con la base de datos DATA MART BOOK en el Data Source Views, seleccionando las tablas a usar.
- b) Se procede a la creación del cubo definiendo primero la dimensión tiempo y posteriormente las demás dimensiones y las medidas.

### Definiendo Dimensión Tiempo



## Creación Cubo



The screenshot displays the Microsoft Visual Studio interface for the 'CuboBookCenter' project. The central workspace shows a data model with the following components:

- fact\_ventas** (Fact Table):
  - Attributes: Clientes\_Id, Vendedor\_Id, Producto\_Id, Proveedor\_Id, Pago\_Id, Tiempo\_Id, Monto, NoVendedores, CantidadProd..., NoProveedores
- dim\_producto** (Dimension):
  - Attributes: Producto\_Id, Cod\_Producto, Tipo\_Producto, Precio, Cantidad
- dim\_cliente** (Dimension):
  - Attributes: Cliente\_Id, Cod\_Cliente, Cliente, RUC, TipoCliente
- dim\_tiempo** (Dimension):
  - Attributes: Tiempo\_Id, Fecha, Año, Mes, Semestre
- dim\_proveedor** (Dimension):
  - Attributes: Proveedor\_Id, Cod\_Proveedor, Proveedor, Direccion
- dim\_vendedor** (Dimension):
  - Attributes: Vendedor\_Id, Cod\_Vendedor, Vendedor, Genero

The interface also shows a 'Vista del origen de datos' (Data Source View) on the left, a 'Cuadro de herramientas' (Toolbox) at the bottom, and a status bar at the bottom right indicating 'La creación del proyecto 'CuboBookCenter' se ha realizado correctamente.' (The creation of the project 'CuboBookCenter' has been completed successfully.)



**Mostrando Implementación**

The screenshot displays the Microsoft Visual Studio environment for SQL Server Data Tools. The main workspace shows a table with the following data:

Cliente	Producto	Nro Pagos	Monto
Charles Nestor Amador Guevara	Papel carbonico Peilkan HandFilm azul	4	21000
Jorge Gonzalez Gonzalez	Cartulina color de	2	12920
Jorge Miguel Villanueva Diaz	Papel carbonico Carbotype azul	4	15232
Kiara Lopez Vargas	Cartulina color de	4	25840
Ulises Manuel Jara Menes	Papel carbonico Peilkan Ultrafilm negro	4	11040
Wendy Rodriguez Soto	Papel carbonico Carbotype negro	4	19040

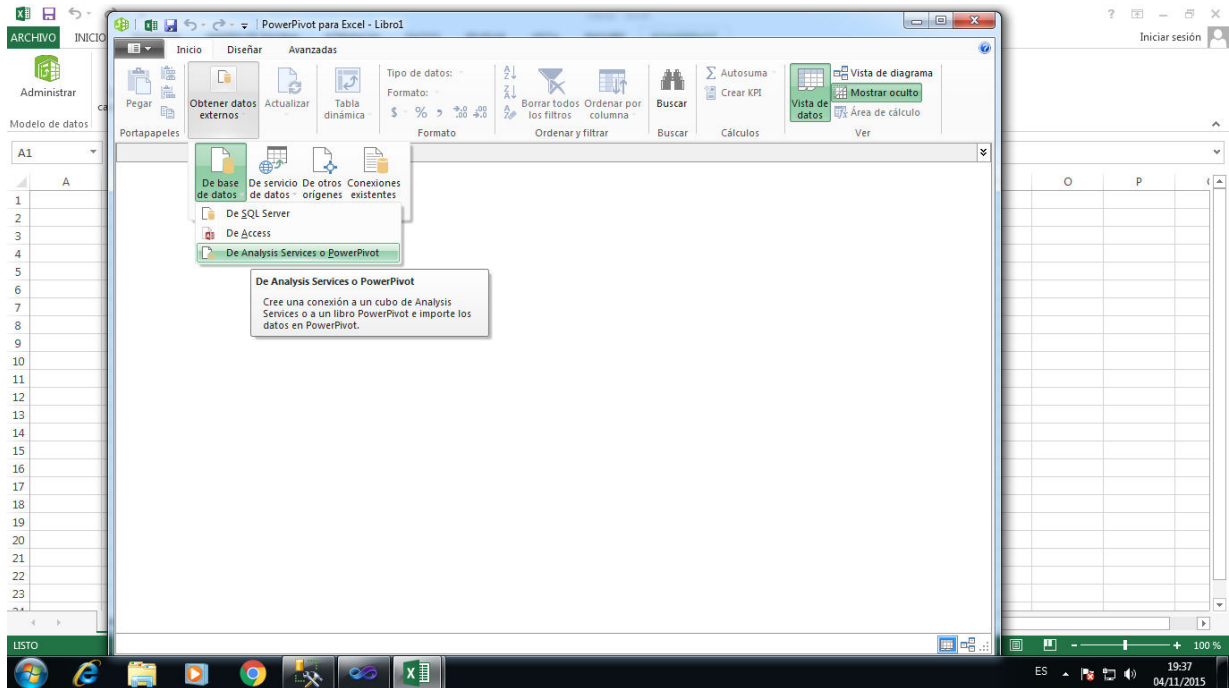
The right-hand pane shows the 'Progreso de la implementación' (Implementation Progress) for the 'CuboBookCenter' cube. It indicates that the implementation was completed successfully, with the message 'La implementación finalizó correctamente' (Implementation finished correctly). The status bar at the bottom right shows 'Progreso de la implementación finalizó correctamente'.

## IX. DESARROLLO DE LA APLICACIÓN DEL USUARIO FINAL

### 1. Elegir la herramienta para procesamiento analítico.

Se han usado Microsoft Visual (Business Intelligence) y Microsoft Excel (PowerPivot)

### 2. Lista los resúmenes de información requeridos por los usuarios.



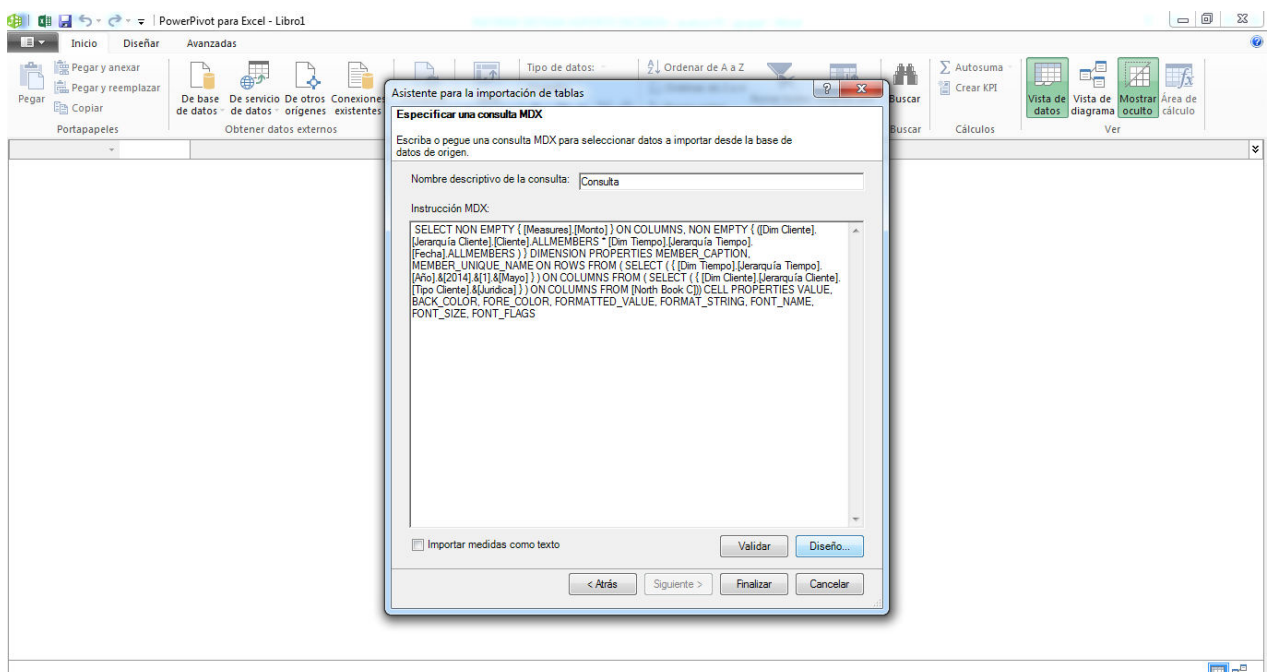
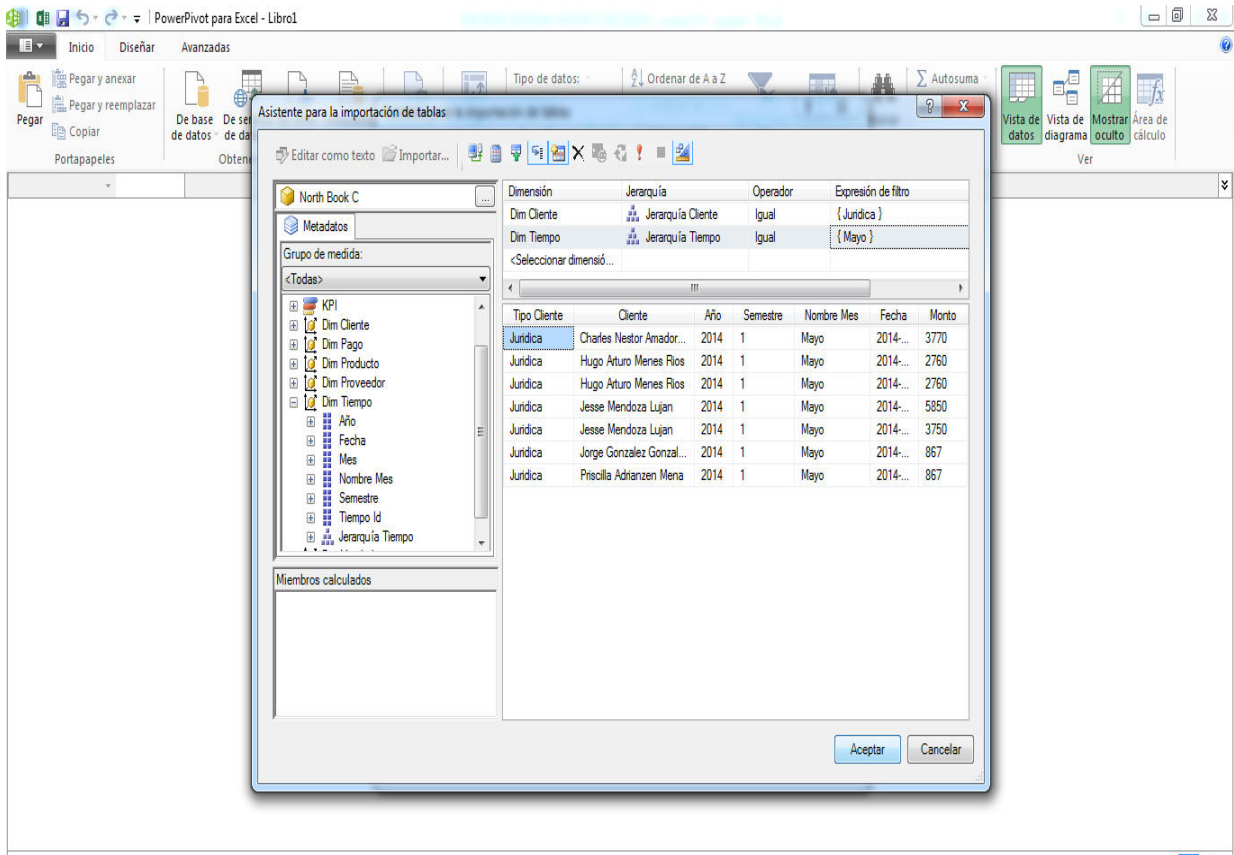
### 3. Determinar los cubos.

“CUBOS VENTAS NORTH BOOK C”

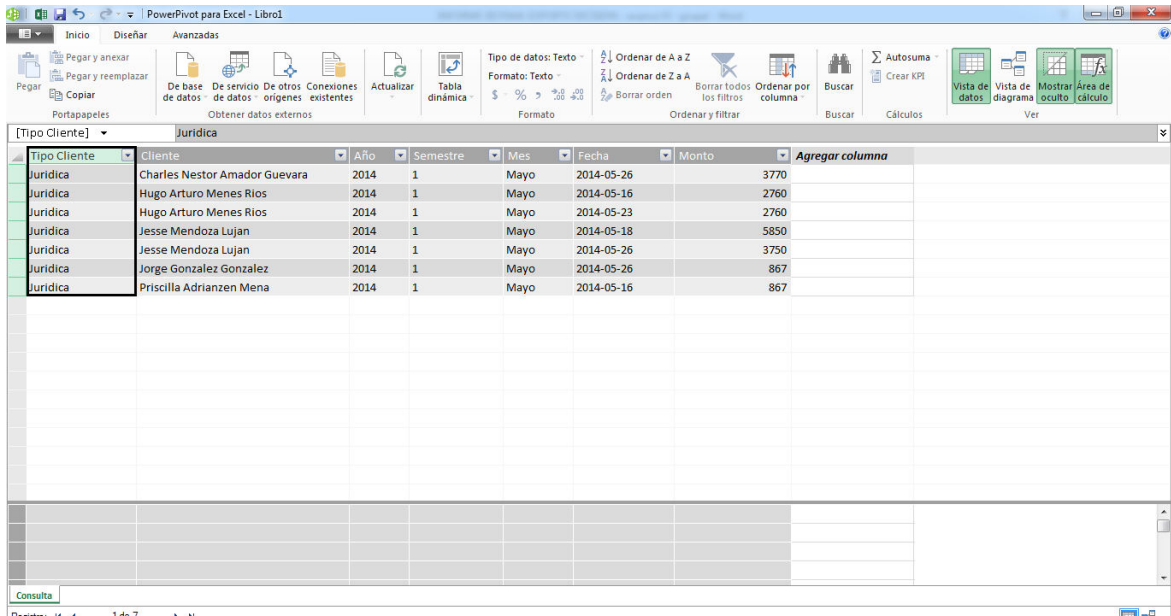


#### 4. Selección de herramienta de aplicación para usuario final: se usara

##### Ms. Excel.



## 5. Implementación de la herramienta Ms Excel.

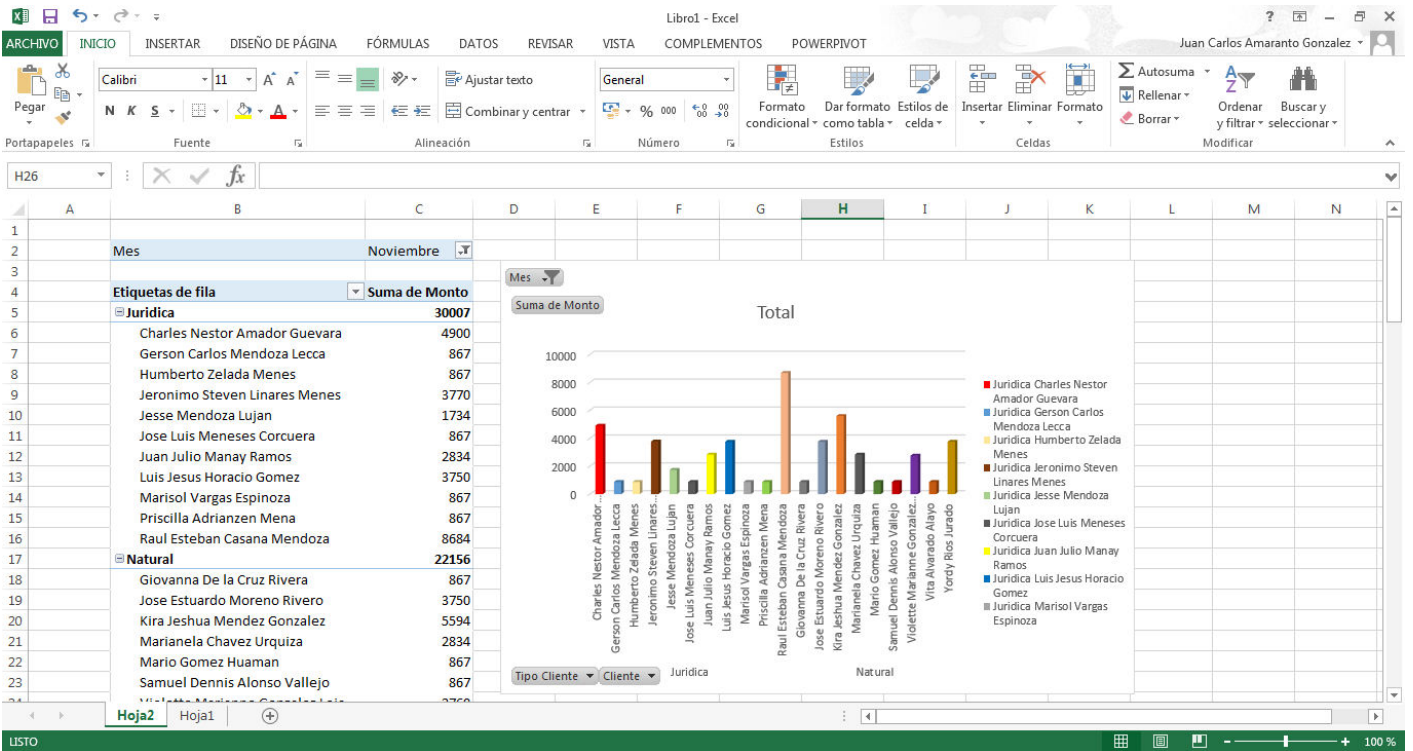


The screenshot displays the Microsoft Excel interface with a PivotTable. The PivotTable is set to 'Jurídica' and has the following columns: Tipo Cliente, Cliente, Año, Semestre, Mes, Fecha, Monto, and Agregar columna. The data rows are as follows:

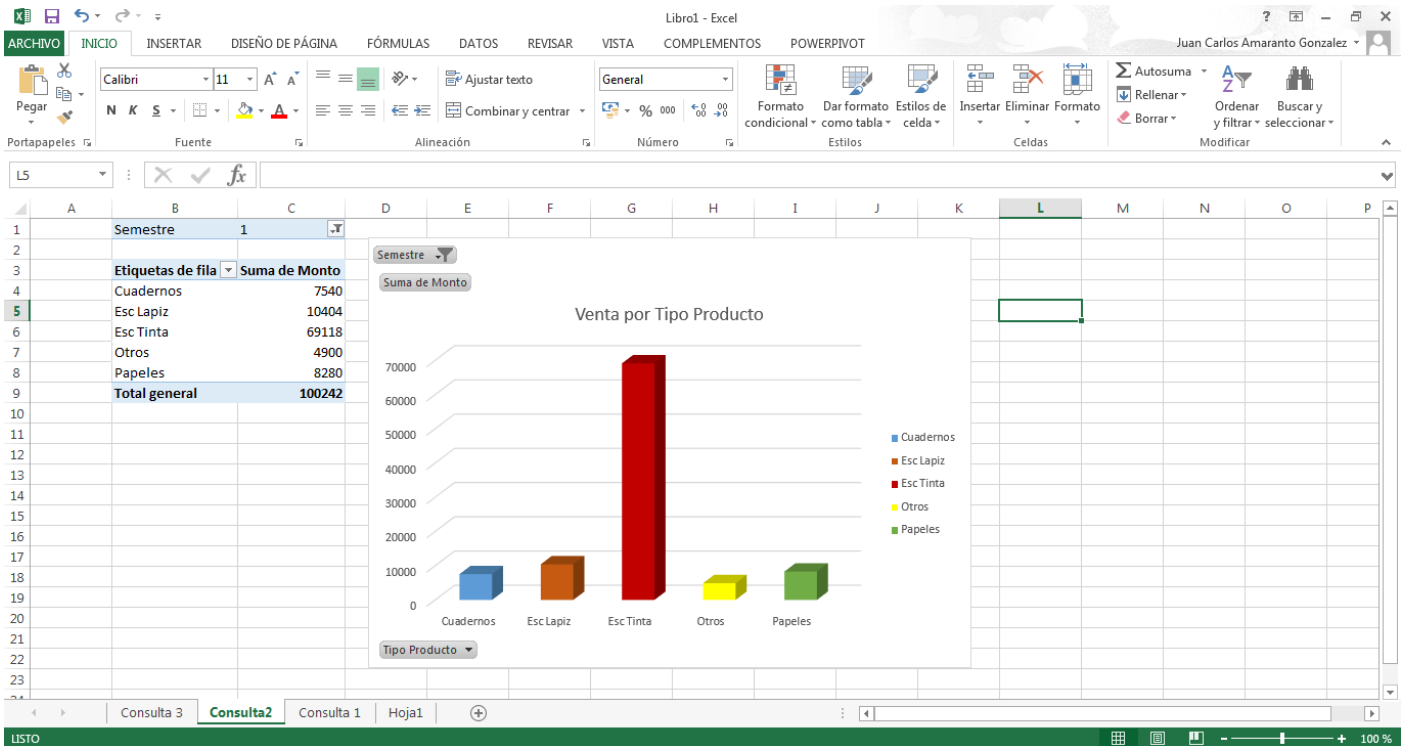
Tipo Cliente	Cliente	Año	Semestre	Mes	Fecha	Monto	Agregar columna
Jurídica	Charles Nestor Amador Guevara	2014	1	Mayo	2014-05-26	3770	
Jurídica	Hugo Arturo Menes Rios	2014	1	Mayo	2014-05-16	2760	
Jurídica	Hugo Arturo Menes Rios	2014	1	Mayo	2014-05-23	2760	
Jurídica	Jesse Mendoza Lujan	2014	1	Mayo	2014-05-18	5850	
Jurídica	Jesse Mendoza Lujan	2014	1	Mayo	2014-05-26	3750	
Jurídica	Jorge Gonzalez Gonzalez	2014	1	Mayo	2014-05-26	867	
Jurídica	Priscilla Adrianzen Mena	2014	1	Mayo	2014-05-16	867	

X. REPORTES EN POWERPIVOT

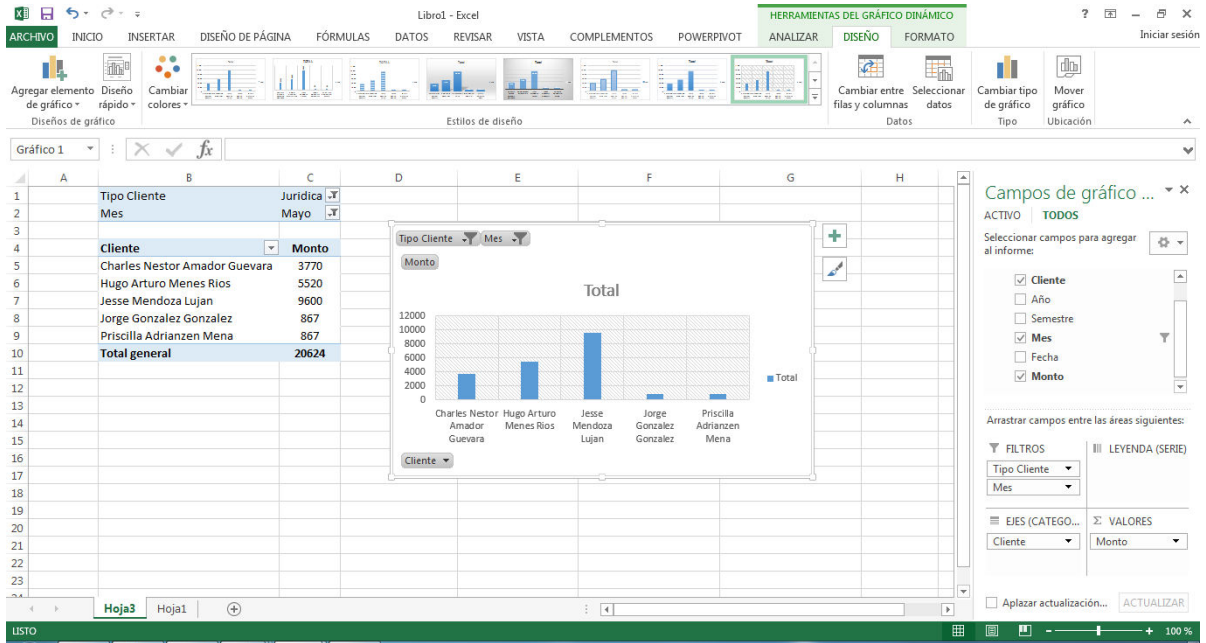
a. Requerimiento 1: ¿Cuál es el monto de un determinado cliente en compras en un mes?



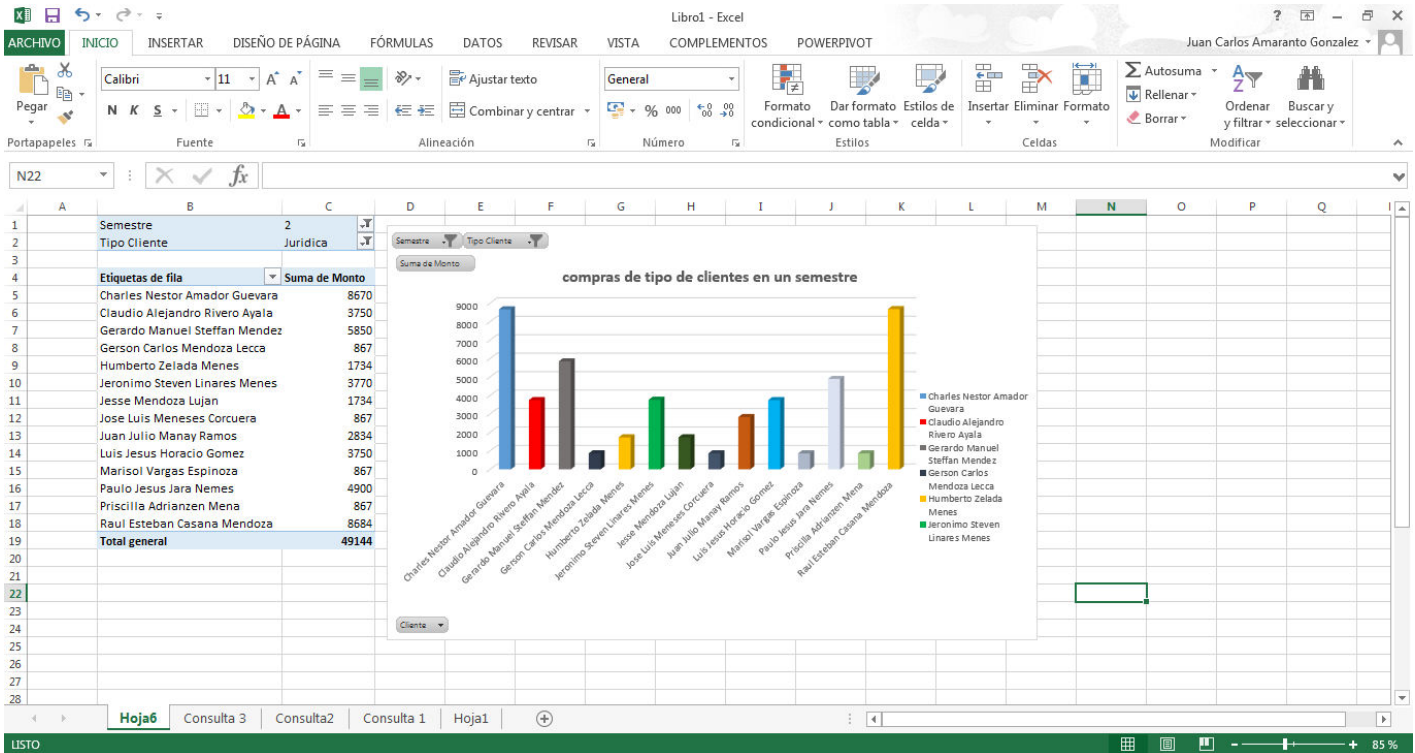
b. Requerimiento 2: ¿Cuál es el Monto de ventas por tipo de producto?



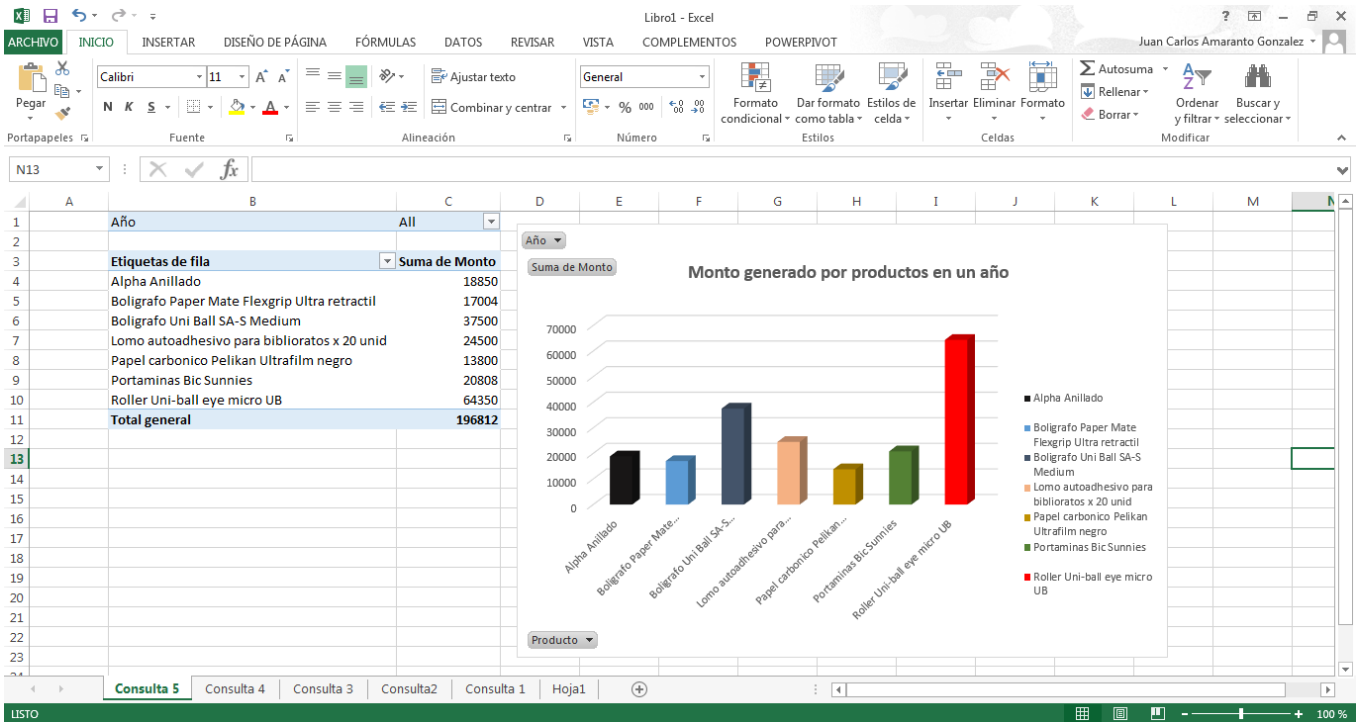
**c. Requerimiento 3: ¿Cuál es el monto de un tipo de producto por proveedor en 3 meses?**



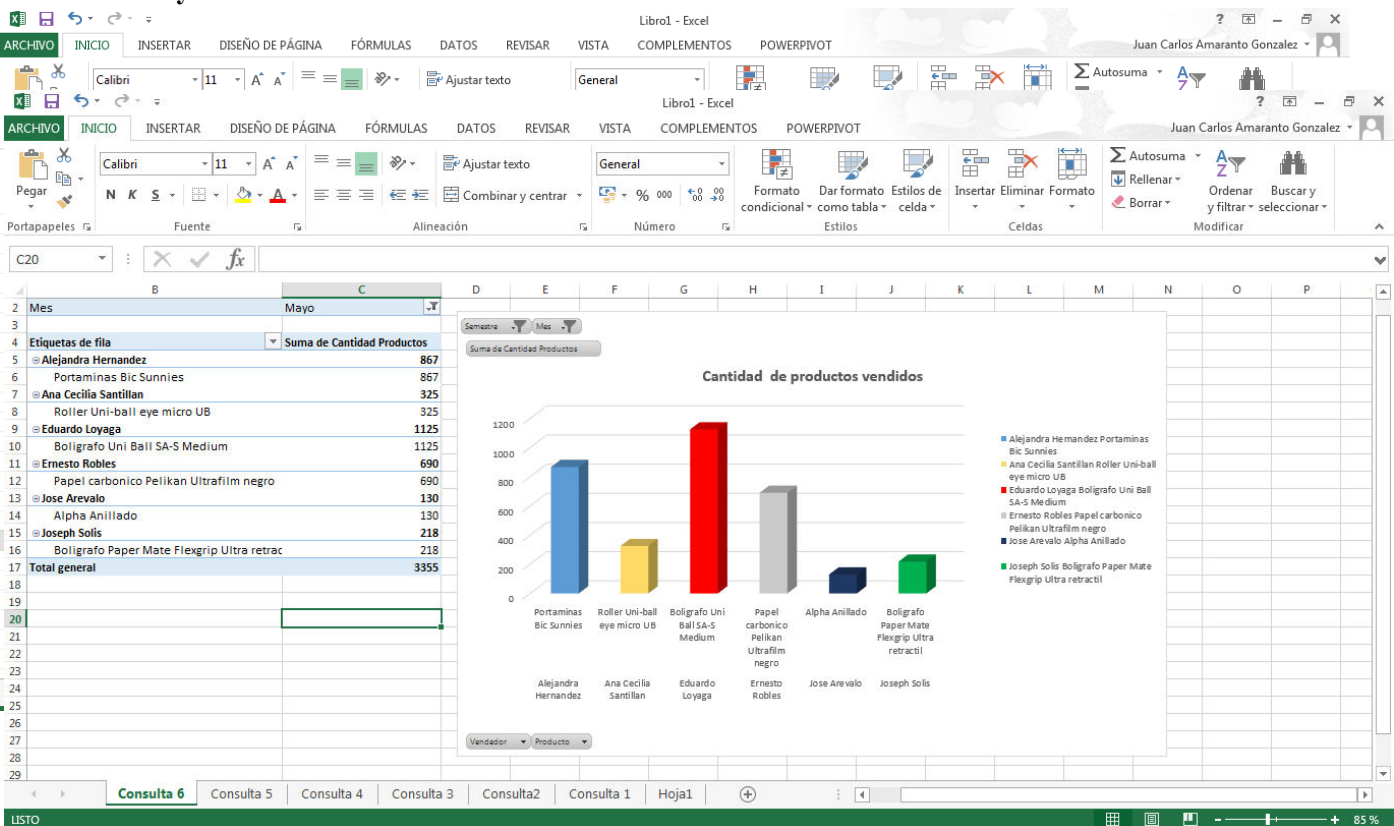
**d. Requerimiento 4: ¿Cuál es el monto de compras de tipos de clientes en un semestre?**



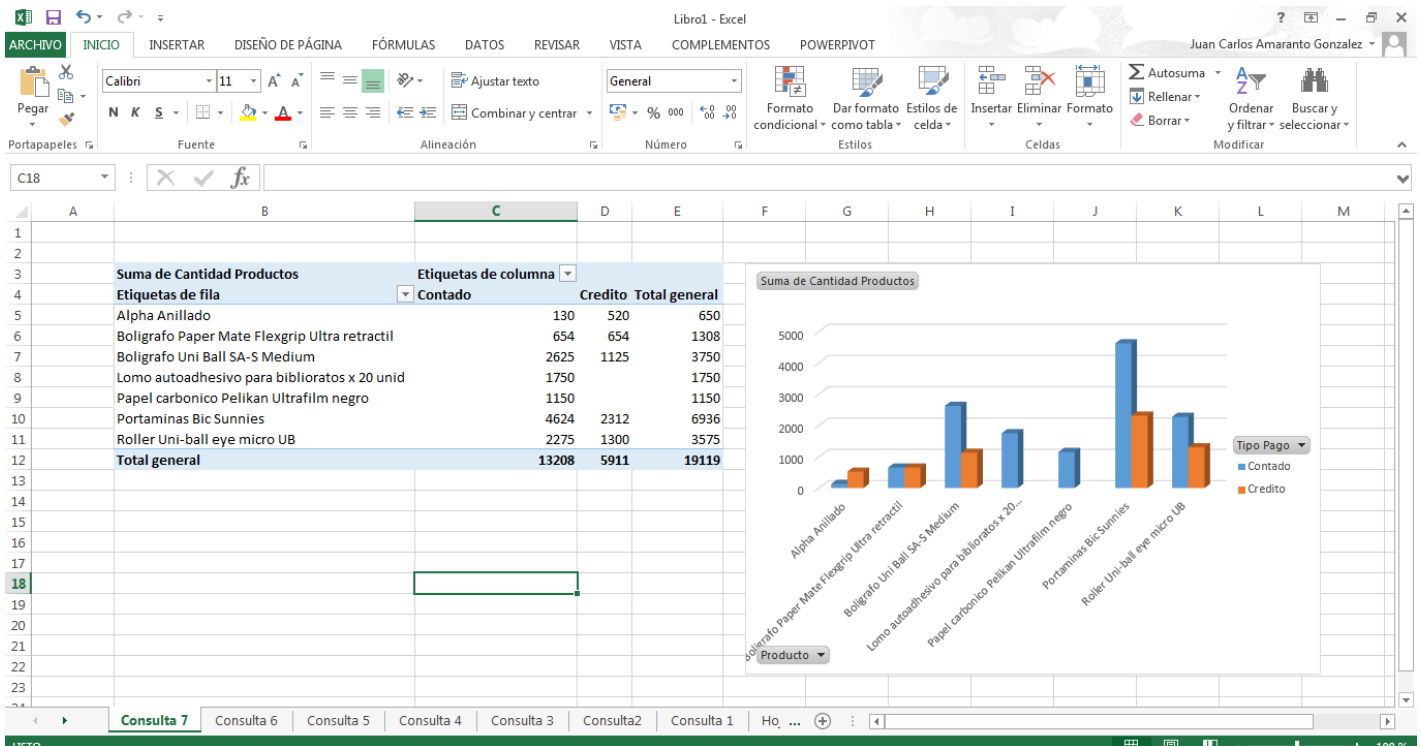
**e. Requerimiento 5: ¿Cuál es el monto que generan los productos en un año?**



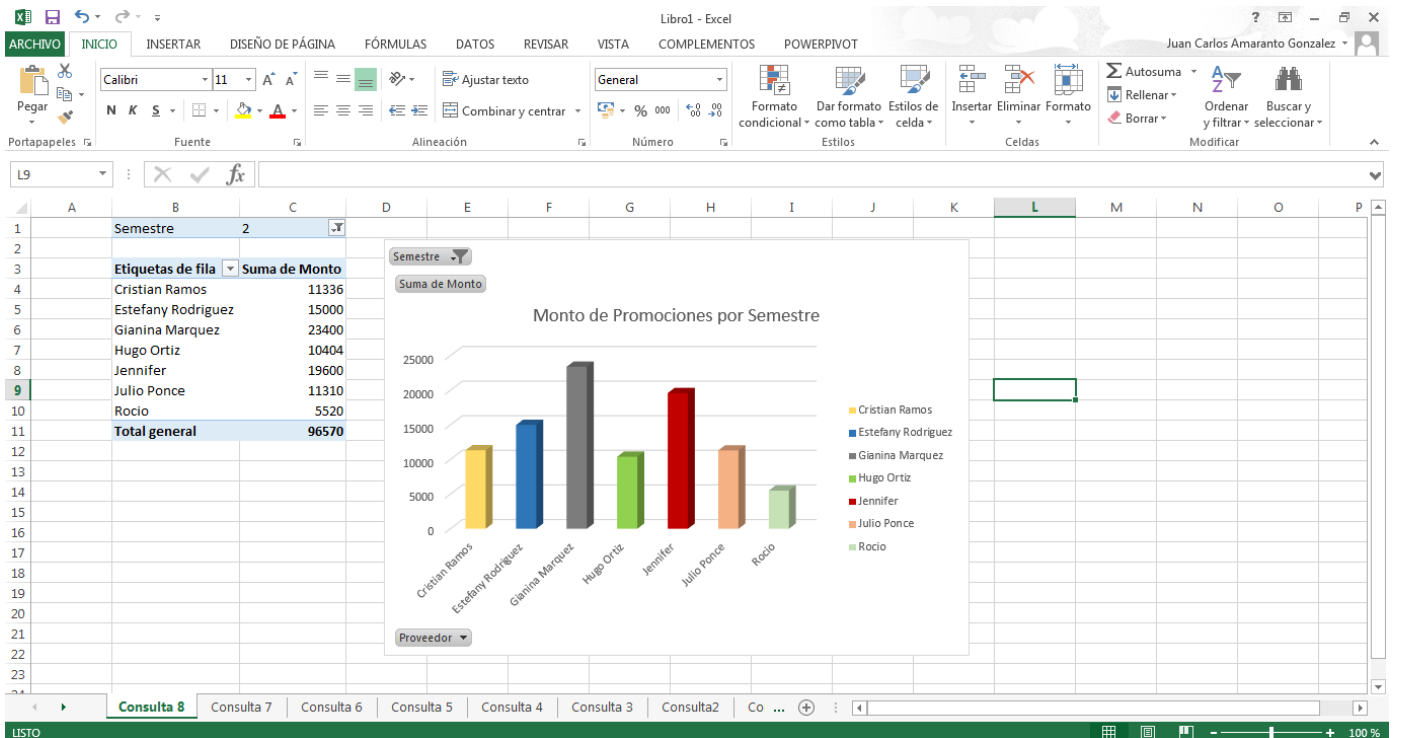
**f. Requerimiento 6: ¿Cuál es la cantidad de productos vendidos por Tipo, vendedor, mes y semestre? Cantidad de Productos**



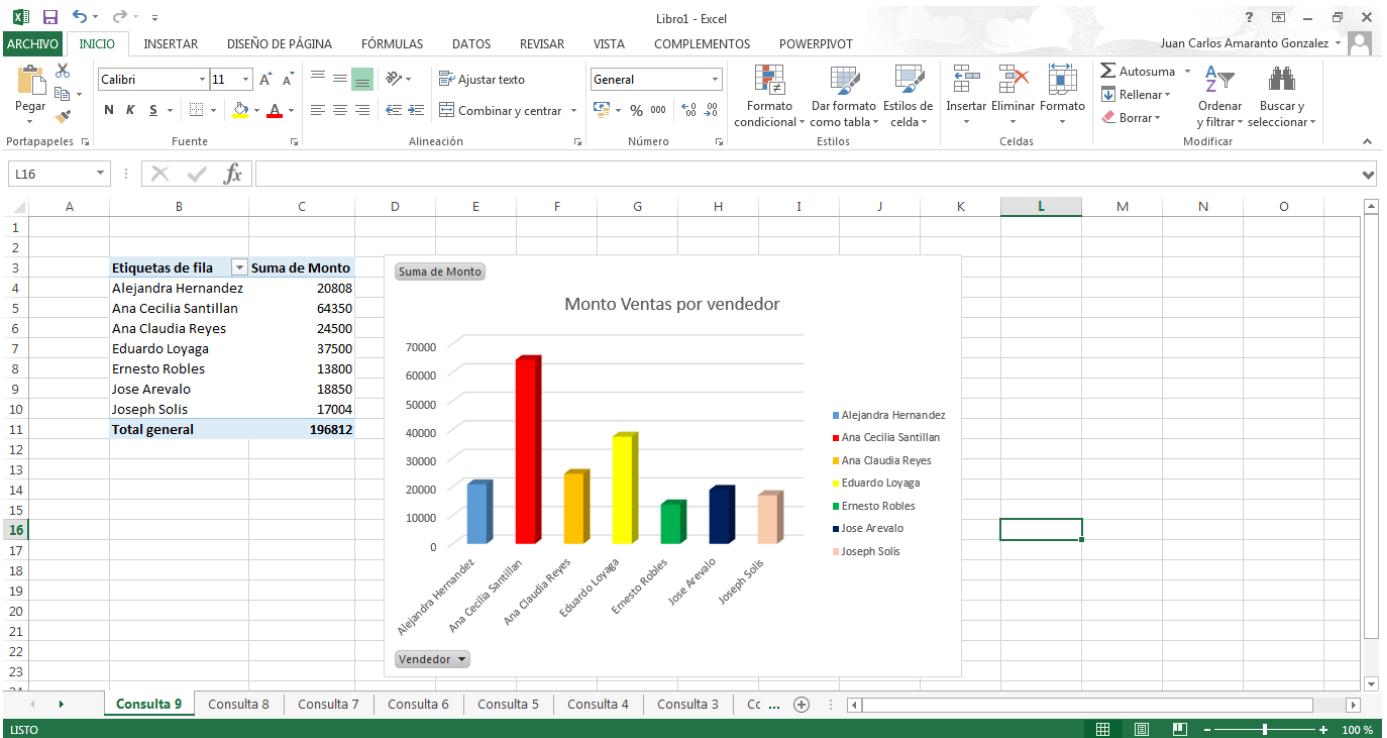
**g. Requerimiento 7: ¿Cuál es la cantidad de productos que se han pagado al crédito y al contado?**



**h. Requerimiento 8: ¿Cuál es el monto que generan las promociones que da un proveedor en un semestre?**

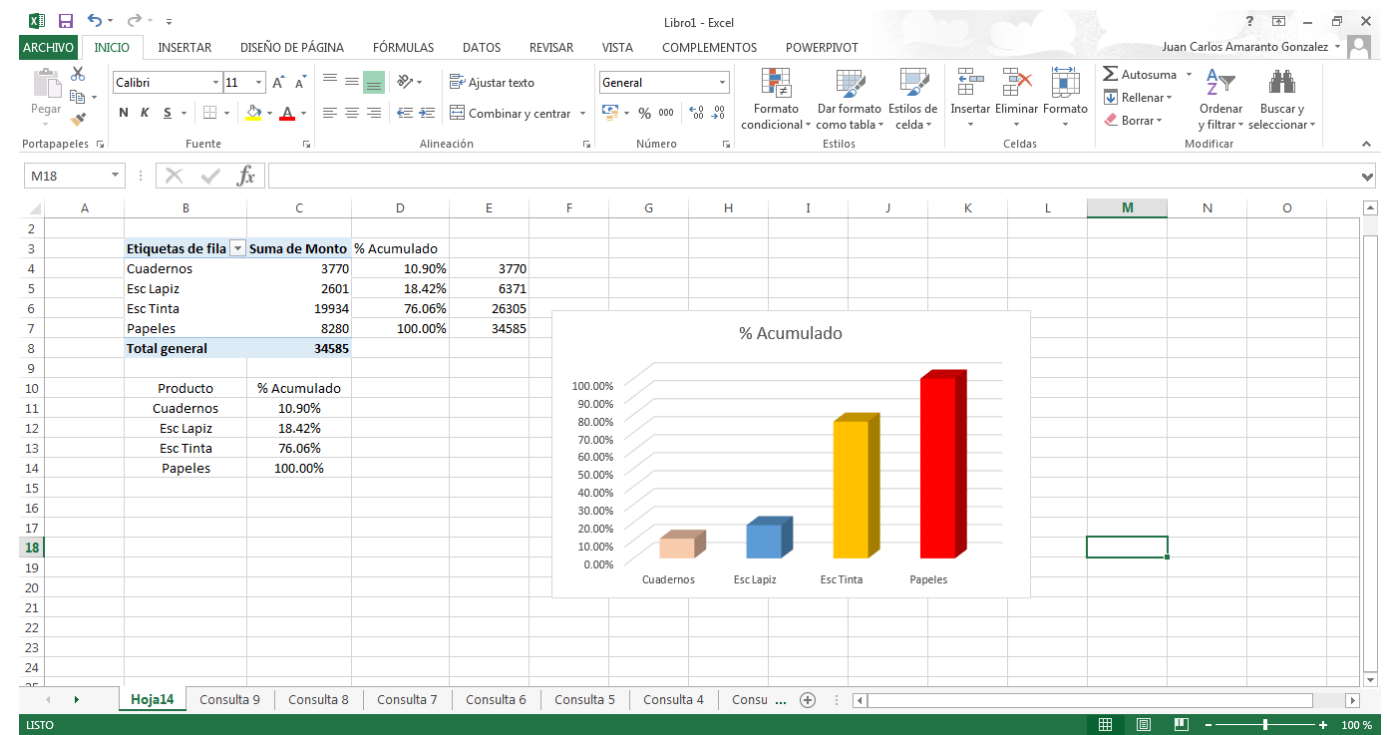






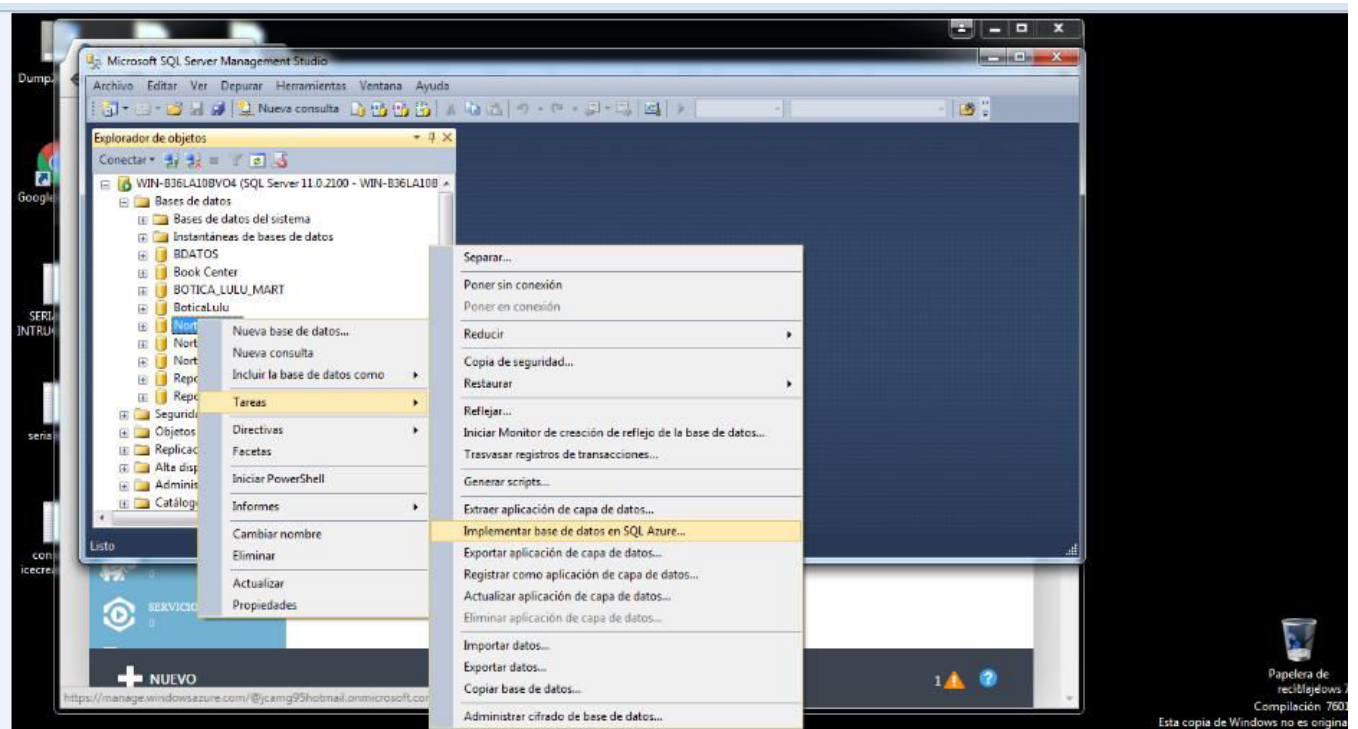
i. **Requerimiento 9: Listar a los Vendedores y el monto total de cada una de las ventas que han generado Monto en soles**

j. **Requerimiento 10: ¿Cuál es el porcentaje de incremento de las ventas de los productos según su tipo por mes?**

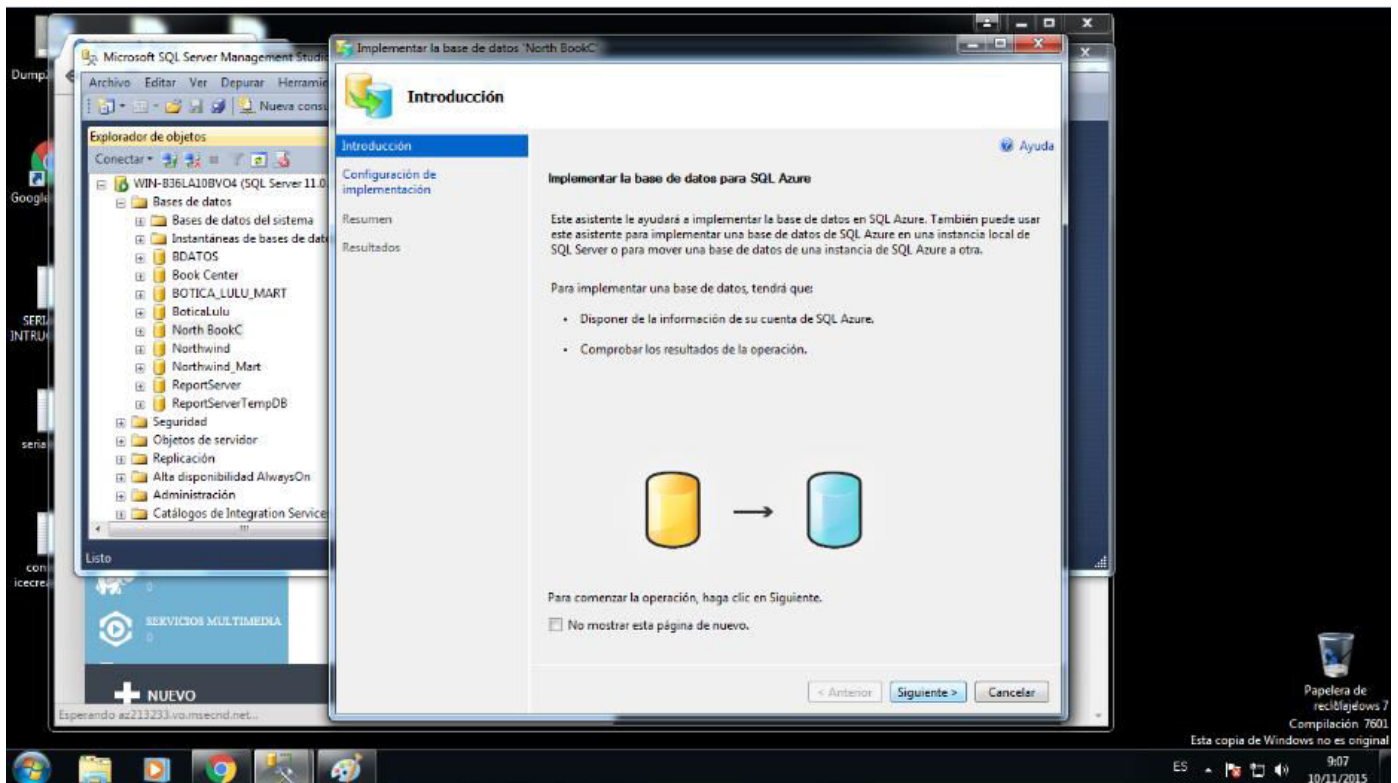


## XI. IMPLEMENTACION EN MICROSOFT AZURE

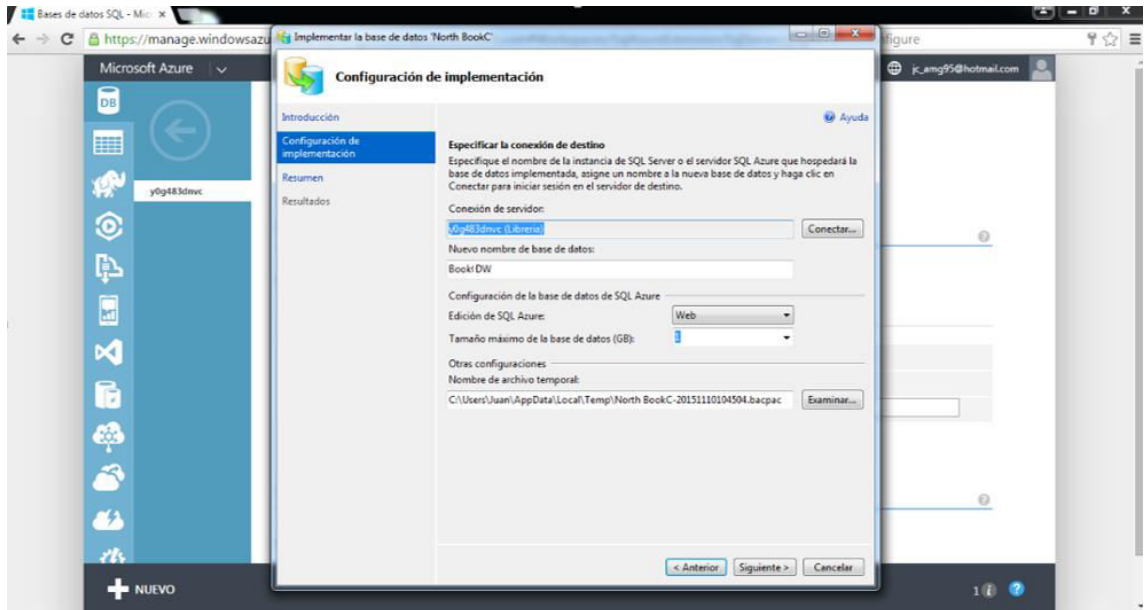
Seleccionamos la base de datos para poder ser implementada para Microsoft Azure



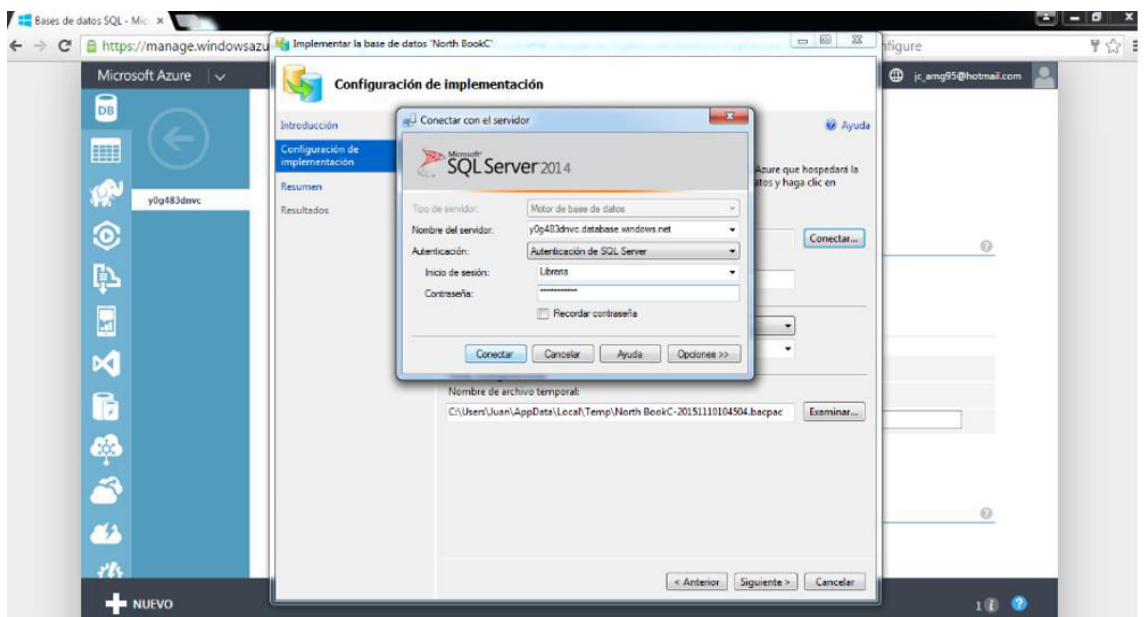
Aparece la Interfaz de Inicio de Implementación de Azure







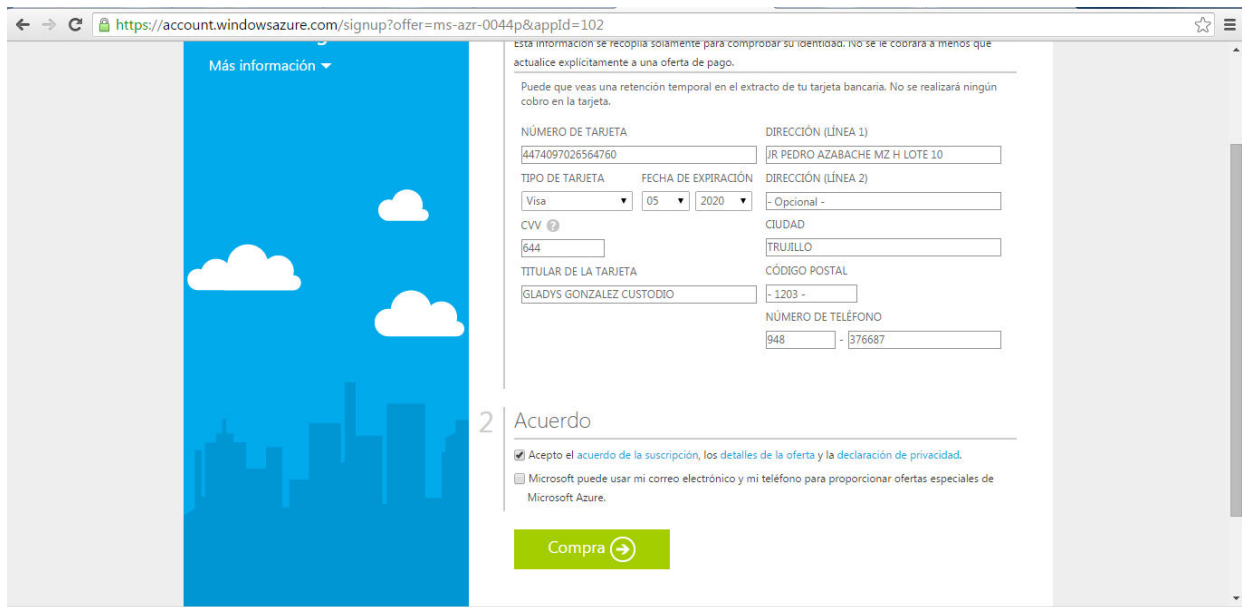
Le asignamos un Nuevo nombre a la base de datos en este caso se llamara BookDW que se subirá en la nube de Windows Azure.



Así mismo se establecerá una conexión con el servidor

- **Inscribirse en la web de Azure**

Rellenar todos los Datos de Acuerdo a lo se pide, se debe tener en cuenta que esta versión de prueba es solo por 30 días , pasado eso se le cargara a su tarjeta el monto



https://account.windowsazure.com/signup?offer=ms-azr-0044p&appId=102

Más información

Esta información se recopila solamente para comprobar su identidad; no se le cobrará a menos que actualice explícitamente a una oferta de pago.

Puede que veas una retención temporal en el extracto de tu tarjeta bancaria. No se realizará ningún cobro en la tarjeta.

NÚMERO DE TARJETA	DIRECCIÓN (LÍNEA 1)	
4474097026564760	JR PEDRO AZABACHE MZ H LOTE 10	
TIPO DE TARJETA	FECHA DE EXPIRACIÓN	DIRECCIÓN (LÍNEA 2)
Visa	05 2020	- Opcional -
CVV	CIUDAD	
644	TRUJILLO	
TITULAR DE LA TARJETA	CÓDIGO POSTAL	
GLADYS GONZALEZ CUSTODIO	- 1203 -	
NÚMERO DE TELÉFONO		
948 - 376687		

## 2 Acuerdo

Acepto el acuerdo de la suscripción, los detalles de la oferta y la declaración de privacidad.

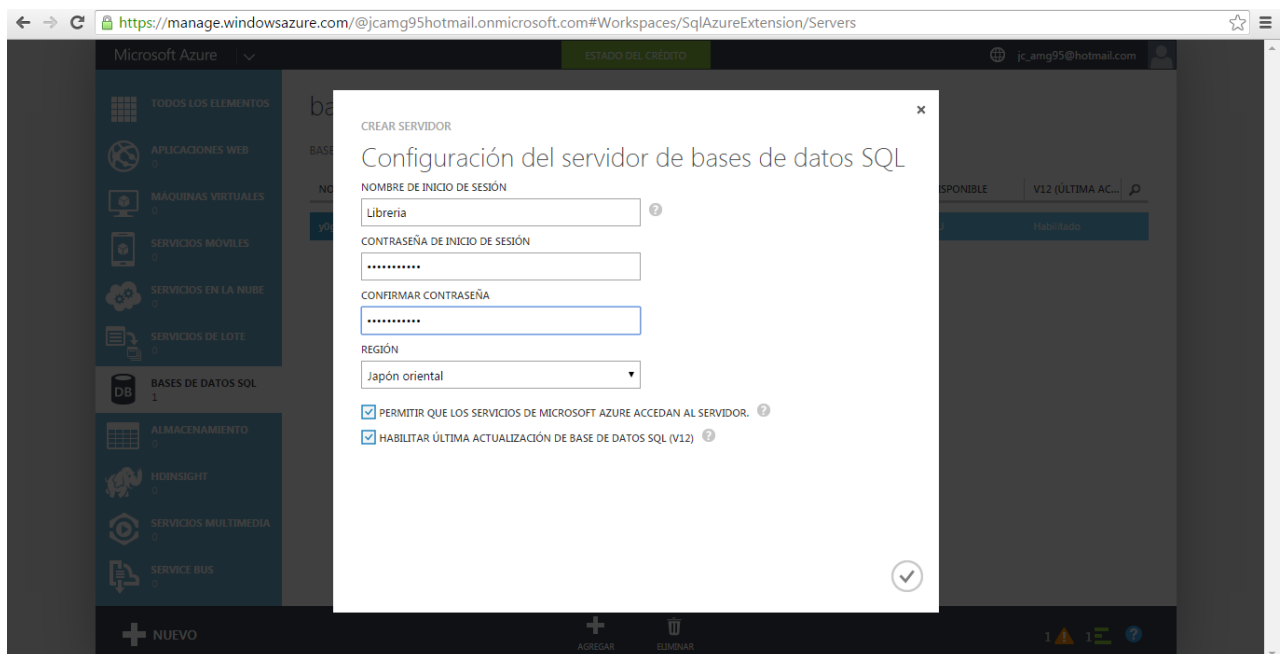
Microsoft puede usar mi correo electrónico y mi teléfono para proporcionar ofertas especiales de Microsoft Azure.

**Compra**

### • Crear Servidor en Azure

Rellenamos los campos correspondientes

En este caso Librería corresponde al nombre de inicio de Sesión asignando al momento de abrir SQL Server Management Studio 2014



https://manage.windowsazure.com/@jcamg95hotmail.onmicrosoft.com#Workspaces/SqlAzureExtension/Servers

Microsoft Azure

ESTADO DEL CREDITO

ic\_amg95@hotmail.com

TODOS LOS ELEMENTOS

- APLICACIONES WEB
- MÁQUINAS VIRTUALES
- SERVICIOS MÓVILES
- SERVICIOS EN LA NUBE
- SERVICIOS DE LOTE
- BASES DE DATOS SQL
- ALMACENAMIENTO
- HEMINSIGHT
- SERVICIOS MULTIMEDIA
- SERVICE BUS

**CREAR SERVIDOR**

### Configuración del servidor de bases de datos SQL

NOMBRE DE INICIO DE SESIÓN

Librería

CONTRASEÑA DE INICIO DE SESIÓN

.....

CONFIRMAR CONTRASEÑA

.....

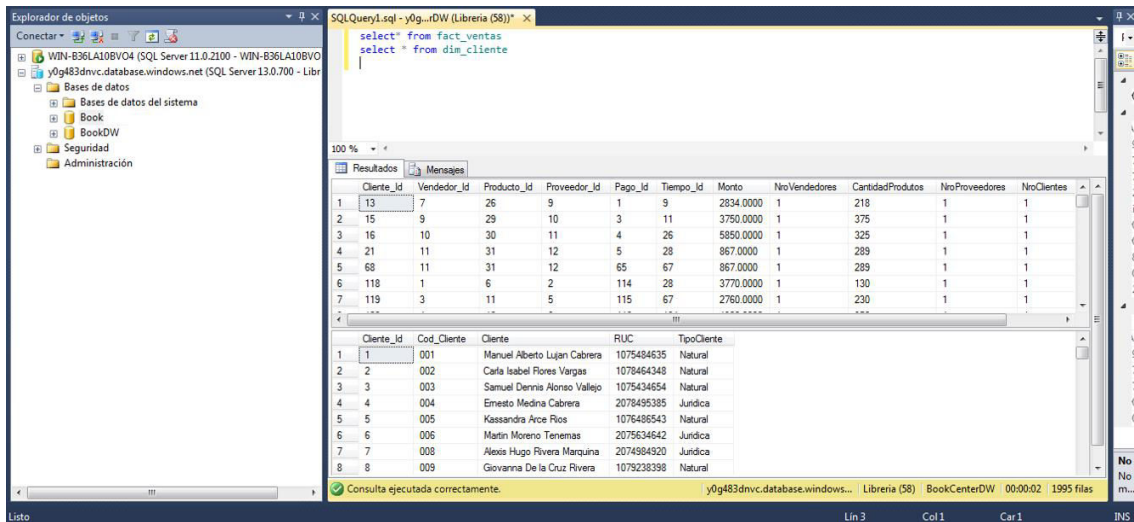
REGIÓN

Japón oriental

PERMITIR QUE LOS SERVICIOS DE MICROSOFT AZURE ACCEDAN AL SERVIDOR.

HABILITAR ÚLTIMA ACTUALIZACIÓN DE BASE DE DATOS SQL (V12)

AGREGAR ELIMINAR



## CAPITULO III

### DISCUSIÓN

Para la contrastación de la hipótesis se ha considerado lo siguiente:

#### Formulación del Problema

¿Cómo se podría mejorar el soporte online en la toma de decisiones para el área de ventas de la empresa Copy Ventas S.R.L., utilizando tecnologías de información?

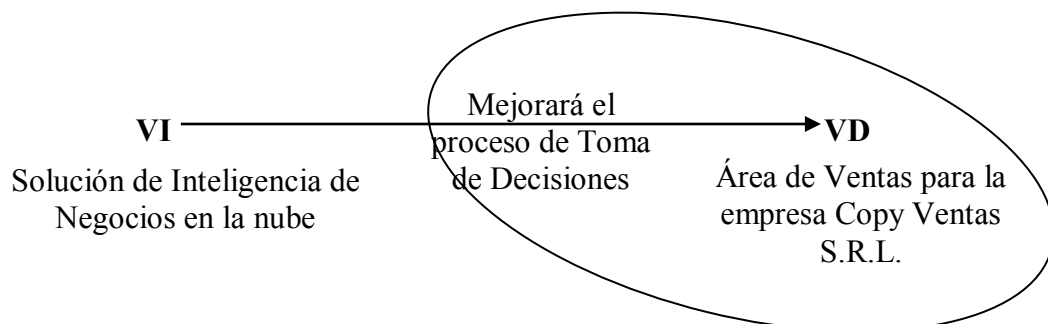
#### Hipótesis

“La implementación de una Solución de Inteligencia de Negocios en la nube mejorará el proceso de Toma de Decisiones en el área de Ventas para la empresa Copy Ventas S.R.L.”

Luego se definen las variables que intervienen en la veracidad o falsedad de la hipótesis:

- Variable Independiente (VI): Solución de Inteligencia de Negocios en la nube
- Variable Dependiente (VD): Proceso de Toma de Decisiones en el área de Ventas para la empresa Copy Ventas S.R.L.”

#### 3.1. MANERA PRESENCIAL



### 3.1.1. DISEÑO PREEXPIMENTAL PRE-PRUEBA Y POST-PRUEBA

**PRE-PRUEBA (O<sub>1</sub>):** Es la medición previa de X a G

**POST-PRUEBA (O<sub>2</sub>):** Corresponde a la nueva medición de X a G

Se determinó usar el Diseño PreExperimental Pre-Prueba y Post-Prueba, porque nuestra hipótesis se adecua a este diseño. Este diseño experimenta con un solo grupo de sujetos el cual es medido a través de un cuestionario antes y después de presentar el estímulo (Data Mart). Este diseño se presenta de la siguiente manera:

**G                      O<sub>1</sub>                      X                      O<sub>2</sub>**

Donde:

X: Tratamiento, estímulo (Data Mart)

O: Medición a sujetos (Cuestionario)

G: Grupo de sujetos (Empleados)

El espacio muestral que se tomó para la medición de los indicadores de la hipótesis, correspondió al total de personas que operarán el Data Mart, siendo estos 2: El jefe de área de ventas y el administrador de la base de datos; a estas dos personas se le aplicó un cuestionario, antes de interactuar con el Data Mart (**O<sub>1</sub>**) y después de interactuar con el mismo (**O<sub>2</sub>**).

Al concluir la investigación se establecen las diferencias entre O<sub>1</sub> y O<sub>2</sub> para determinar si hay o no incremento en los resultados obtenidos.

### 3.1.2. CÁLCULO DE LOS INDICADORES DE LA HIPÓTESIS

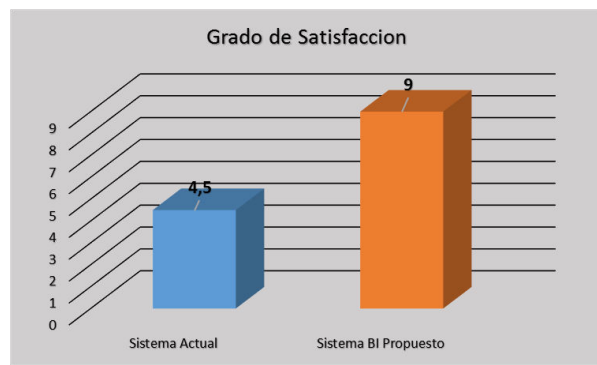
Para el cálculo de los indicadores de la hipótesis en el Sistema de soporte de decisiones (Data Mart) Propuesto (DMP) propuesto y el Sistema Actual (SA), se realizó un cuestionario (Ver Anexo B) donde se evaluó a los usuarios luego de haber interactuado con el Data Mart.

Los valores que los usuarios dieron a las respuestas del cuestionario fueron aplicados según el rango de satisfacción que muestran en la siguiente tabla:

RANGO	GRADO DE SATISFACION
0 – 2.5	Insatisfecho
2.5 – 5.0	Medianamente Satisfecho
5.0 – 7.5	Satisfecho
7.5 – 10.0	Muy Satisfecho

Tabla 3.1. Rango de grado de satisfacción

**Pregunta N° 1: ¿Se puede conocer Cuál es el monto de un determinado cliente en compras en un mes?**



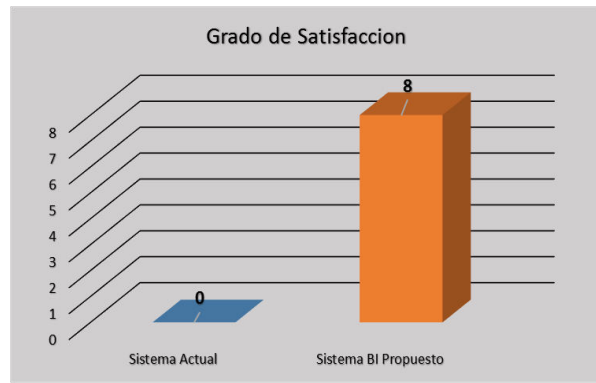
**Resultado:** Los valores obtenidos son : 4.5 (S.A.) y 9 (D.M.P). El Data Mart muestra información sobre el monto de un determinado cliente en compras en un mes mediante tablas y gráficos

**Pregunta N° 2: ¿ Se puede conocer Cuál es el Monto de ventas por tipo de producto?**



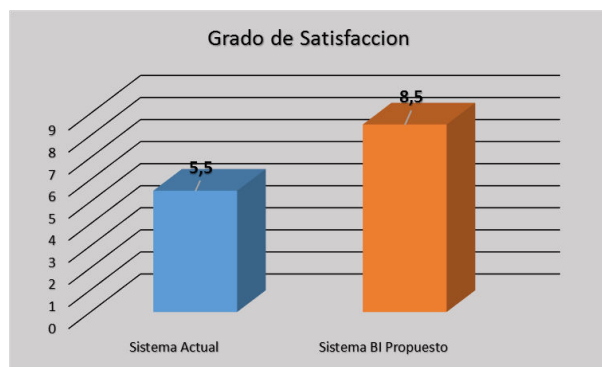
**Resultado:** Los valores obtenidos son : 0 (S.A.) y 8 (D.M.P). El Data Mart muestra información sobre el Monto de ventas por tipo de producto en forma dinámica.

**Pregunta N° 3: ¿Se puede conocerCuál es el monto de un tipo de producto por proveedor en 3 meses?**



**Resultado:** Los valores obtenidos son : 0 (S.A.) y 8 (D.M.P). El Data Mart muestra información sobre el monto de un tipo de producto por proveedor en 3 meses.

**Pregunta N° 4: ¿Se puede conocerCuál es el monto de ventas clientes en un semestre?**



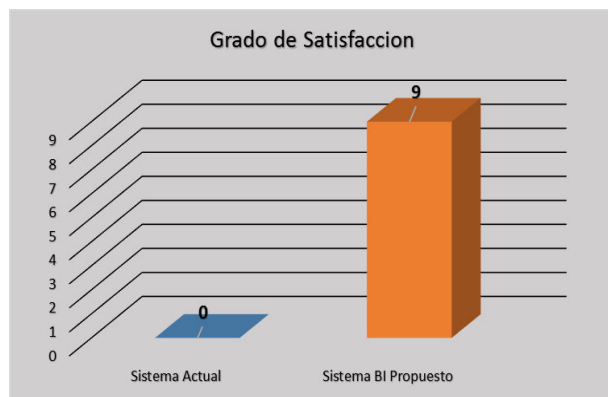
**Resultado:** Los valores obtenidos son: 5.5 (S.A.) y 8.5 (D.M.P). El Sistema Actual permite conocer el monto de ventas, pero no lo presenta en forma dinámica. El Data Mart presenta dinámicamente y con gráficos para un mejor entendimiento.

**Pregunta N°5 ¿Se puede conocerCuál es el monto que generan los productos en un año?**



**Resultado:** Los valores obtenidos son: 0 (S.A.) y 8 (D.M.P). El Data Mart muestra información sobre el monto que generan los productos en un año. El Sistema Actual no presenta esta consulta.

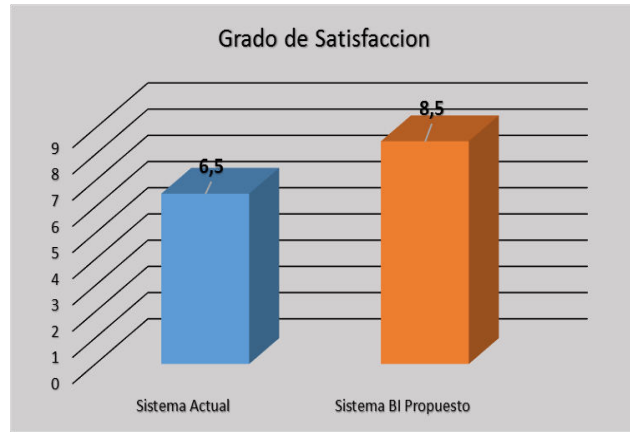
**Pregunta N° 6: ¿Se puede saberCuál es la cantidad de productos vendidos por Tipo, vendedor, mes y semestre?**



**Resultado:** Los valores obtenidos son: 0 (S.A.) y 9 (D.M.P). El Data Mart muestra información sobre la cantidad de productos vendidos por Tipo, vendedor, mes y semestre. El Sistema Actual no presenta esta consulta

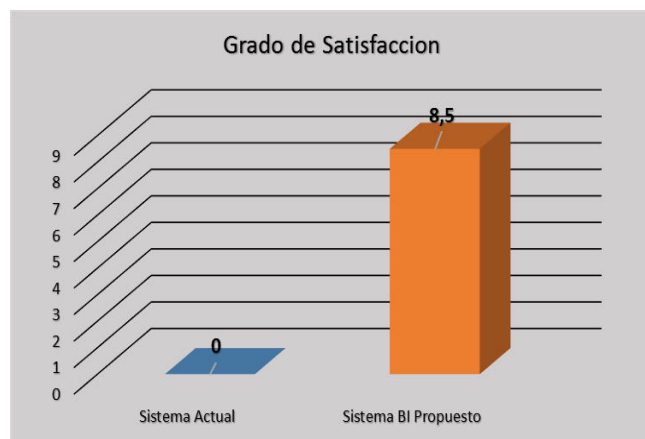


**Pregunta N° 7: ¿Se puede conocer Cuál es la cantidad de Productos que se han pagado al crédito y al contado?**



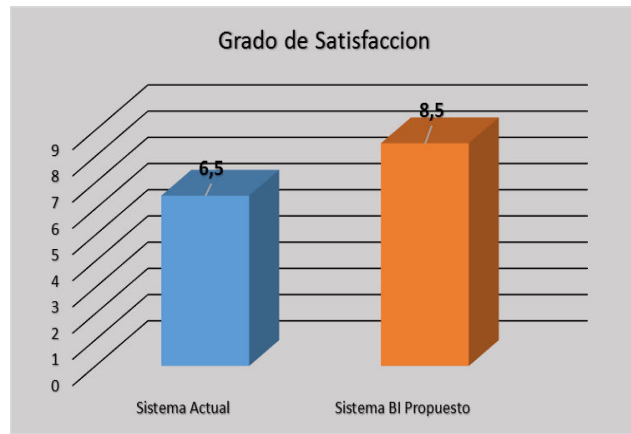
**Resultado:** Los valores obtenidos son: 6.5 (S.A.) y 8.5 (D.M.P). El Sistema Actual permite conocer la cantidad de Productos que se han pagado al crédito y al contado. El Data Mart presenta dinámicamente y con gráficos para un mejor entendimiento

**Pregunta N° 8: ¿Se puede conocer Cuál es el monto que generan las promociones que da un proveedor en un semestre?**



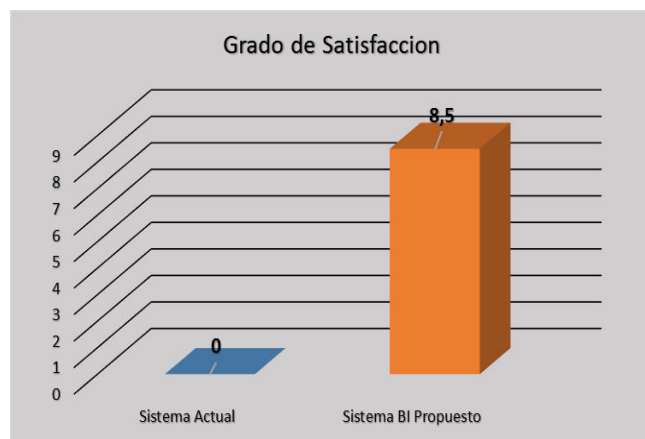
**Resultado:** Los valores obtenidos son: 0 (S.A.) y 8.5 (D.M.P). El Data Mart muestra información sobre el monto que generan las promociones que da un proveedor en un semestre en forma dinámica. El Sistema Actual no presenta esta consulta.

**Pregunta N° 9: ¿Se puede Listar a los Vendedores y el monto total de cada una de las ventas que han generado?**



**Resultado:** Los valores obtenidos son: 6.5 (S.A.) y 8.5 (D.M.P). El Sistema Actual permite Listar a los Vendedores y el monto total de cada una de las ventas que han generado. El Data Mart presenta dinámicamente y con gráficos para un mejor entendimiento

**Pregunta N° 10: ¿Se puede conocerCuál es el porcentaje de las ventas de los productos según su tipo por mes?**



**Resultado:** Los valores obtenidos son: 0 (S.A.) y 8.5 (D.M.P). El Data Mart muestra información sobre el porcentaje de las ventas de los productos según su tipo por mes en forma dinámica. El Sistema Actual no presenta esta consulta.

### 3.1.3. APLICACIÓN DEL RANGO DE SATISFACCIÓN A LOS INDICADORES DE LA HIPÓTESIS

Los valores aplicados a los indicadores de la hipótesis tanto para el sistema Actual como para el Data Mart propuesto se muestran en la siguiente tabla:

#### Evaluación de los indicadores de la hipótesis:

INDICADORES	S.A.	D.M.P
1. ¿Cuál es el monto de un determinado cliente en compras en un mes?	4.5	9
2. ¿Cuál es el Monto de ventas por tipo de producto?	0	8
3. ¿Cuál es el monto de un tipo de producto por proveedor en 3 meses?	0	8
4. ¿Cuál es el monto de ventas clientes en un semestre?	5.5	8.5
5. ¿Cuál es el monto que generan los productos en un año?	0	8
6. ¿Cuál es la cantidad de productos vendidos por Tipo, vendedor, mes y semestre?	0	9
7. ¿Cuál es la cantidad de Productos que se han pagado al crédito y al contado?	6.5	8.5
8. ¿Cuál es el monto que generan las promociones que da un proveedor en un semestre?	0	8.5
9. Listar a los Vendedores y el monto total de cada una de las ventas que han generado	6.5	8.5
10. ¿Cuál es el porcentaje de las ventas de los productos según su tipo por mes?	0	8.5
<b>PROMEDIO</b>	<b>1.65</b>	<b>8.45</b>

*Tabla 3.2. Evaluación de los indicadores de la hipótesis.*

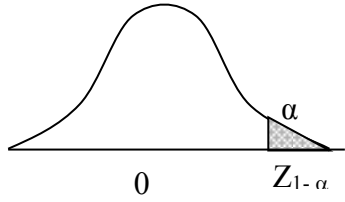
### 3.1.4. ANÁLISIS ESTADÍSTICO PARA LA PRUEBA PRESENCIAL DE LA HIPÓTESIS

**Cálculo de la diferencia de dos medias:**

DESCRIPCIÓN	MEDIAS	VARIANZAS
Fórmula	$\mu_i = \frac{\sum X_i}{N}$	$\sigma^2 = \frac{\sum (X_i - \mu_i)^2}{N}$
Cálculo	$\mu_1 = 1.65$	$\sigma^2_1 = 4.01$
N=10	$\mu_2 = 8.45$	$\sigma^2_2 = 0.12$

*Tabla 3.3. Cálculo de la diferencia de dos medias*

**Cálculo de la Prueba de Hipótesis:**

TIPO DE HIPÓTESIS	ESTADÍSTICA DE PRUEBA	REGIONES DE ACEPTACIÓN Y RECHAZO DE Ho	VALOR CRÍTICO
Hipótesis Nula Ho : $\mu_1 - \mu_2 = 0$ Nivel de signif $\alpha$	$z_0 = \frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{\sqrt{\sigma^2_1/n + \sigma^2_2/n}}$		$\alpha = 0.05$ $Z_{1-\alpha} = 0.97$
Hipótesis Alternativa H <sub>1</sub> : $\mu_1 > \mu_2$	<b>Z<sub>0</sub> = 13.67</b>	<b>Rechazar Ho si, Z<sub>0</sub> &gt; Z<sub>1-\alpha</sub></b>	<b>13.67 &gt; 0.97</b>

*Tabla 3.4. Cálculo de la Prueba de la Hipótesis*

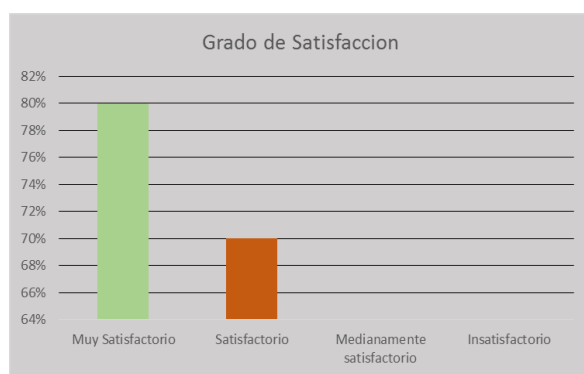
**CONCLUSIÓN:** La implementación de una solución de Inteligencia de Negocios utilizando SQL Server 2014 , Microsoft Azure y la metodología de Ralph Kimball da un mejor apoyo en la Toma de Decisiones en el área de Ventas para la empresa Copy Ventas S.R.L.

### 3.2. MANERA NO PRESENCIAL

La muestra que se tomó fue la misma a los cuales se les aplico un cuestionario (Ver Anexo C) el cual comprobará si el Data Mart obtiene información valiosa para los procesos de Ventas.

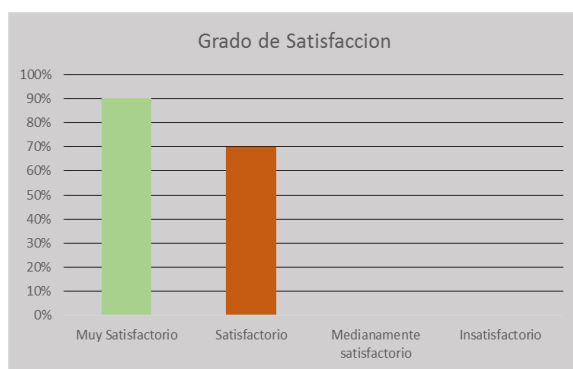
#### CUESTIONARIO DIRIGIDO AL JEFE DE EL ÁREA DE VENTAS y AL JEFE DE AREA INFORMATICA:

**Pregunta N° 1: ¿De que manera se realizó el ingreso a los datos que contiene el Data Mart?**



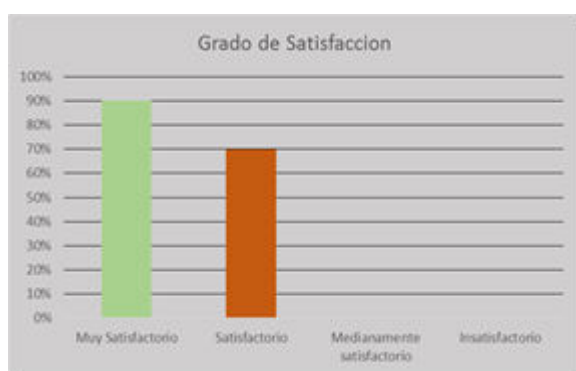
**Resultado:** El Jefe del área de Ventas respondió que el ingreso a los datos que contiene el Data Mart se realizó muy satisfactoriamente (80%). El Jefe del Área de Informática respondió el ingreso a los datos que contiene el Data Mart se realizó satisfactoriamente (70%).

**Pregunta N° 2: ¿El Data Mart permitió realizar consultas que se consideran valiosas para el Área de Ventas?**



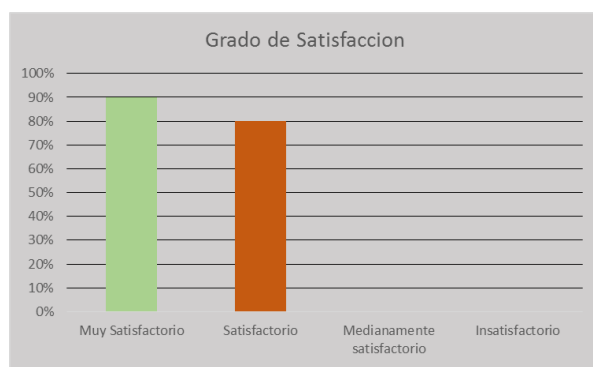
**Resultado:** El Jefe del el área de ventas respondió que el Data Mart permitió realizar consultas que se consideran valiosas para el Área de una manera muy satisfactoriamente (90%). El Jefe del Área de Informática respondió que el Data Mart permitió realizar consultas que se consideran valiosas para el Área de Ventas de una manera satisfactoriamente (70%).

**Pregunta N° 3: ¿El Data Mart puede realizar comparaciones de información de diferentes semestres al mismo tiempo?**



**Resultado:** El Jefe del Área de Ventas respondió que el Data Mart realiza comparaciones de información de diferentes semestres al mismo tiempo de manera muy satisfactoriamente (90%). El Jefe del área de informática respondió que el Data Mart realiza comparaciones de información de diferentes semestres al mismo tiempo de manera satisfactoriamente (70%).

**Pregunta N° 4: ¿El Data Mart permitió un análisis dinámico de los datos?**



**Resultado:** El jefe del Área de Ventas respondió que el Data Mart permitió un análisis dinámico de los datos muy satisfactoriamente (90%). El Jefe del área de Informática respondió que el Data Mart permitió un análisis dinámico de los datos muy satisfactoriamente (80%).

**CONCLUSIÓN:** Luego de haber evaluado a los usuarios del Data Mart concluimos que, el Sistema BI implementado (Data Mart) presenta información valiosa y de forma dinámica para un mejor análisis de los datos que pueda dar un mejor soporte en la toma de decisiones al área de Ventas.

## CONCLUSIONES

- Se determinó que el alcance del proyecto está en el área de Ventas, basándose en las entrevistas hechas a los stakeholders.
- Se determinaron 10 requerimientos que ayudan a la toma de decisiones en el área de ventas que el jefe del área y gerencia necesitan, en base a las entrevistas.
- Con el análisis de los requerimientos se logró identificar la jerarquía de los datos en cada consulta hecha por el usuario determinando el modelo de datos Starnet, los cuales ayudarán al análisis drill down que permite el data mart.
- Al realizar el diseño se identificaron 01 tabla hecho y 7 dimensiones (Ver Pág. 61)
- La construcción del Datamart se realizó utilizando las herramientas de SQL Server 2014, como son SQL Server Data Tools con Integración de Servicios, Analysis Services y Microsoft Azure.
- Se realizaron pruebas basadas en los requerimientos del usuario a fin de garantizar el correcto funcionamiento del Datamart, estas pruebas se realizaron utilizando Power Pivot. (Ver Pág. 91)



## RECOMENDACIONES

- Realizar un Planeamiento Estratégico de Tecnologías de Información con la finalidad de identificar posibles proyectos de TI, para optimizar los procesos en cuanto al ahorro y/o a la reducción de costos.
- Se debe realizar una evaluación periódica de los procesos de Tecnologías de Información de la empresa, para dirigirlos hacia los lineamientos estratégicos de la institución.
- Para proyectos futuros se recomienda el uso de la metodología de Ralph Kimball y de SQL Server 2014 integrado con Windows Azure como herramientas de trabajo para el diseño e implementación de un proyecto de BI.
- Para próximas versiones se recomienda el uso de dashboard que ofrece la Herramienta de Qlikview para un mejor control de los indicadores.

## REFERENCIAS

- CHUNGA, Ana y BOTTON, Jessica. (2013). DATA MART para los procesos de ventas e inventarios de la empresa Consorcio Electronic SAC utilizando la metodología de Ralph Kimball, SQL Server y Qlikview.
- FOWLER, Martin. Refactoring: Improving the Design of Existing Code. Addison Wesley Longman, 1999.
- KENDALL, K y KENDALL, J. 2012. Análisis y Diseño de Sistemas, 8° Edición. Prentice-Hall. Hispanoamericana.
- KIMBALL, R. y ROSS, M., 2013, The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling. 3° Edition, Editorial Wiley. Canadá.
- LANDAURO, Juan. (2008). Metodología para crear y/o agregar valor con la incorporación de estrategias de E-Business en empresas tradicionales.
- LAUDON, K.C. y LAUDON, J.P.(2013), Management Information Systems: Managing the Digital Firm. 13° Edición. Prentice-Hall PTR. Estados Unidos.
- PTD 2012, Web Site con información y gráficos sobre el proceso de toma de decisiones <http://blog.conductuempresa.com/2012/01/proceso-de-toma-de-decisiones.html>
- ROBBINS Stephen & COULTER Mary (2005) Administración. México. Editorial Pearson Educación

- 
- RODRIGUEZ Gómez, Jesús. (2011). Toma de decisiones. Más allá de la intuición, Primera edición. CreateSpace Independent Publishing Platform.
  - ROMERO, Liliam y VINCES, Erika. (2012). Diseño de un Datamart para el Soporte a la Toma de Decisiones en el área de Ahorro de la Caja Municipal de Ahorro y Crédito de Trujillo usando la Metodología Ralph Kimball y SQL Server 2008 BUSINESS INTELLIGENCE.
  - ROOT , R. y MASON C. (2012). Pro SQL Server 2012 BI Solutions. 1º Edición. Apress. Estados Unidos
  - HERNÁNDEZ Celis, Domingo (2009) Costos, presupuestos y toma de decisiones empresariales. Lima. Edición a cargo del autor.
  - VILLALOBOS E. y Constenia J., 2010, Tecnologías de la Información y la Comunicación. Instituto Politécnico Nacional. [En línea]. Disponible en <http://www.dycy.ipn.mx/dycy/quesonlastics.aspx#> [accesado el 25 de Marzo 2016].

# ANEXOS

## ANEXO A

### ENTREVISTAS Y CUESTIONARIOS

#### CUESTIONARIO PARA ENTREVISTA N° 1: Gerente

##### A. LAS RESPONSABILIDADES

- Describe su área y su relación con el resto de la compañía.
- ¿Cuáles son sus responsabilidades primarias?

##### B. LOS OBJETIVOS COMERCIALES Y PROBLEMAS

- ¿Cuáles son los objetivos de su área?
- ¿Qué usted está tratando de lograr con estos objetivos?
- ¿Cuáles de estos objetivos son su prioridad para alcanzar sus metas dentro de su organización?
- ¿Cuáles son sus factores críticos de éxito?
- ¿Cómo usted sabe que usted está haciendo bien?
- ¿Qué tan menudo usted mide los factores de éxito importantes?
- ¿De los departamentos que funcionan? ¿cuáles son cruciales para asegurar que los factores de éxito importantes se logren?
- ¿Qué roles cumplen estos departamentos?
- ¿Cómo ellos trabajan para asegurar el éxito juntos?
- ¿Cuáles son los importantes problemas que usted enfrenta hoy dentro de su función? Y ¿Cuál es el impacto en la organización?
- ¿Cómo usted identifica sus problemas en su Área o sabe que usted se dirige hacia el problema?

### **C. ANALISIS DE LOS REQUISITOS**

- En el análisis de los datos ¿Qué papel juega las decisiones que usted y otros gerentes toman en la ejecución del negocio?
- ¿Qué información importante se exige a hacer o a apoyar las decisiones que usted hace en el proceso de lograr sus metas y superar los obstáculos? ¿Cómo usted consigue esta información hoy?
- ¿Está allí otra información que no está disponible a usted hoy y que usted cree tendría el impacto significativo en ayudar a encontrar sus metas?
- ¿Están allí los cuellos de botella específicos a llegar a la información?
- ¿Que informes usted usa actualmente?
- ¿Qué datos en el informe son importantes?
- ¿Cómo usted usa la información?
- ¿Qué problemas encuentra en estos informes?
- ¿Cuánta información histórica se requiere?
- ¿Si el informe fuera dinámico, en que lo haría diferentemente?
- ¿Qué capacidades analíticas le gustaría tener?
- ¿Qué oportunidades existen para mejorar dramáticamente su negocio basándose en el acceso mejorado de la información?
- ¿Cuál es el impacto financiero usted piensa que tendría?

### **CUESTIONARIO PARA ENTREVISTA N° 2: Jefe área de Ventas**

#### **A. LAS RESPONSABILIDADES**

- Describe su organización y su relación con el resto de la compañía.
- ¿Cuáles son sus responsabilidades primarias?

#### **B. LOS OBJETIVOS COMERCIALES Y PROBLEMAS**

- ¿Cuáles son los objetivos de en el desempeño de su función ?
- ¿Qué usted está tratando lograr con estos objetivos ?
- ¿Cuáles de estos objetivos son sus prioridad para alcanzar sus metas dentro de su función?
- ¿Cuáles son sus factores críticos de éxito?

- ¿Qué tan menudo usted mide los factores de éxito importantes?
- ¿Cuáles son los importantes problemas que usted enfrenta hoy?
- ¿Qué le impide cubrir sus objetivos ?
- ¿Cuál es el impacto de estos problemas en la organización?
- ¿Cómo usted identifica sus problemas en su organización o sabe que usted se dirige hacia el problema?
- ¿Cómo es la relación que tiene con el Gerente?

### **C. ANALISIS LOS REQUISITOS**

- ¿Qué tipo de análisis rutinario usted realiza actualmente? ¿Qué datos se usa?
- ¿Cómo usted consigue los datos actualmente?
- ¿Qué usted hace una vez con la información que usted obtiene?
- ¿Que informes usted usa actualmente?
- ¿Qué datos en el informe es importante?
- ¿Cómo usted usa la información?
- ¿Qué oportunidades existen para mejorar dramáticamente su negocio basándose en el acceso mejorado de la información?

## **CUESTIONARIO PARA ENTREVISTA N° 3: Jefe de Area de Sistemas**

### **A. LAS RESPONSABILIDADES**

- Describe su organización y su relación al resto de la compañía.
- ¿Cuáles son sus responsabilidades primarias?

### **B. EL APOYO AL USUARIO/ LOS ANÁLISIS Y REQUISITOS DE LOS DATOS**

- ¿Cuál es el proceso actual para hacer llegar (obtener) la información?
- ¿Qué herramientas son usadas para acceder y analizar la información hoy? ¿Quién las usa?
- ¿Le piden que realice los análisis rutinarios?
- ¿Usted crea los informes estandarizados?
- Describa las demandas de información ad hoc típicas. ¿Cuánto tiempo toma para cumplir estas demandas?
- ¿Quién son los demandantes más frecuentes de análisis y / o datos ?

- ¿Cómo es el mecanismo de apoyo?
- ¿Cuál es el cuello de botella más grande / los problemas con los datos actuales que encuentran en el proceso?
- ¿Hay un atrasos en enviar información a los demandantes?

### **C. DATOS DISPONIBILIDAD Y CALIDAD**

- ¿Qué sistemas de la fuente se usan para la información frecuentemente-pedida?
- ¿Cuál es la granularidad?
- ¿Qué tan menudo son los datos puesto al día?
- ¿Cuánta información histórica está disponible? Y ¿Cuánta necesita para realizar los informes?
- ¿Cuál es un tamaño estimado de estos datos (preliminar #de filas)?
- ¿Cuáles son los archivos principales que usted tiene? Describa el mantenimiento de éstos archivos.
- ¿Usted tiene los archivos de la fuente comunes actualmente?
- ¿Quién mantiene los archivos de la fuente?
- ¿Cómo las llaves se mantienen? ¿Son las llaves reasignadas?
- ¿Cuál es el cardinalidad (#los valores distintos)?

## ANEXO B

### CUESTIONARIO DIRIGIDO: JEFE DEL AREA DE VENTAS Y AL JEFE DEL AREA DE SISTEMAS

PREGUNTAS	VALORES										
	0	1	2	3	4	5	6	7	8	9	10
1. ¿Cuál es el monto de un determinado cliente en compras en un mes?											
2. ¿Cuál es el Monto de ventas por tipo de producto?											
3. ¿Cuál es el monto de un tipo de producto por proveedor en 3 meses?											
4. ¿Cuál es el monto de ventas clientes en un semestre?											
5. ¿Cuál es el monto que generan los productos en un año?											
6. ¿Cuál es la cantidad de productos vendidos por Tipo, vendedor, mes y semestre?											
7. ¿Cuál es la cantidad de Productos que se han pagado al crédito y al contado?											
8. ¿Cuál es el monto que generan las promociones que da un proveedor en un semestre?											
9. Listar a los Vendedores y el monto total de cada una de las ventas que han generado											
10. ¿Cuál es el porcentaje de las ventas de los productos según su tipo por mes?											

*Tabla B1. Cuestionario Dirigido al Jefe del Área de Ventas y al Jefe del área de Sistemas*



## ANEXO C

### CUESTIONARIO DIRIGIDO AL JEFE DEL AREA DE VENTAS Y EL JEFE DEL AREA DE SISTEMAS

PREGUNTAS	VALORES										
	0	1	2	3	4	5	6	7	8	9	10
1. ¿De qué manera se realizó el ingreso a los datos que contiene el Data Mart?											
2. ¿El Data Mart permitió realizar consultas que se consideran valiosas para el Área de Ventas?											
3. ¿El Data Mart puede realizar comparaciones de información de diferentes meses al mismo tiempo?											
4. ¿El Data Mart permitió un análisis dinámico de los datos?											

*Tabla C1. Cuestionario Dirigido al Jefe del Área de Ventas y el Jefe del Área de Sistemas*