

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



TRABAJO DE TESIS PARA OPTENER EL TITULO DE INGENIERO DE  
COMPUTACION Y SISTEMAS

---

**“SOLUCION DE INTELIGENCIA DE NEGOCIOS EN LA NUBE  
BAJO LA PLATAFORMA DE MICROSOFT AZURE Y POWER BI  
PARA EL PROCESO DE VENTAS DE LA EMPRESA BOOK  
CENTER S.A.C.”**

---

**LINEA DE INVESTIGACION:**

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

**AUTORES:**

Br. Enrique Estuardo Valera Jara  
Br. Jorge Daniel Quispe Chávez

**ASESOR:**

Ing. Agustín Eduardo Ullón Ramírez

**TRUJILLO**  
**2015**

**“SOLUCION DE INTELIGENCIA DE NEGOCIOS EN LA NUBE BAJO LA  
PLATAFORMA DE MICROSOFT AZURE Y POWER BI PARA EL PROCESO DE  
VENTAS DE LA EMPRESA BOOK CENTER S.A.C.”**

**Elaborado por:**

Bach. Enrique Estuardo Valera Jara

Bach. Jorge Daniel Quispe Chávez

**Aprobada por:**

---

**Dr. Walter Aurelio Lazo Aguirre**  
Presidente  
CIP: 36034

---

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

---

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

---

**Ing. Agustin Eduardo Ullón Ramirez**  
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, ponemos a vuestra disposición el presente Trabajo de Tesis: “SOLUCION DE INTELIGENCIA DE NEGOCIOS EN LA NUBE BAJO LA PLATAFORMA DE MICROSOFT AZURE Y POWER BI PARA EL PROCESO DE VENTAS DE LA EMPRESA BOOK CENTER S.A.C.” 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 nuestra formación profesional, además de consulta de fuentes bibliográficas e información obtenida en la empresa.

Los autores.

## **DEDICATORIA**

**A mi esposa Delia y a mis hijas Alyssa Fernanda y Kyara Thais** que son el soporte necesario para mi desarrollo personal y me dieron la fuerza y fe para creer en lo que me parecía imposible: Terminar.

**A mi padre Misael Valera Paisig**, en quien quiero verme reflejado y que desde el cielo sigue guiándome.

**A mi madre Juana Jara Siccha** por apoyarme incondicionalmente en el cuidado de mis hijas, día a día.

**A mis hermanas Jackeline y Katherine** por estar a mi lado en cada momento importante de mi vida.

**Al Dr. José Alfredo Castañeda Nassi**, por haberme impulsado a culminar este proyecto, por su apoyo total y por brindarme su amistad desinteresadamente.

**Enrique Estuardo Valera Jara**

**A mi padre Jorge Quispe Díaz**, por sus consejos y ese afán de no dejar las cosas a medias y seguir adelante.

**A mi madre Irma Chávez Prado**, por ser siempre mi apoyo, mi guía, mi mayor fuente de inspiración, a la fortaleza, lucha, dedicación y amor.

Gracias por su excelente ejemplo de superación y el pilar fundamental en todo lo que soy, e toda mi educación, tanto académica, como de la vida.

**Jorge Daniel Quispe Chávez**

## **AGRADECIMIENTO**

Queremos expresar nuestro agradecimiento al todo el personal de la empresa Book Center y en especial al administrador, quien nos brindó todas las facilidades para conocer el quehacer diario y los principales procesos de la empresa.

Al asesor el Ing. Agustín Ullón, por su apoyo y asesoramiento en el desarrollo de la presente Tesis.

A nuestras familias quienes son el motor y la fuerza que nos impulsa a seguir desarrollándonos personal y académicamente.

Y a todas aquellas personas que prefirieron permanecer en el anonimato pero que de alguna u otra manera estuvieron siempre con nosotros durante todo el camino de este trabajo de tesis.

Muchas Gracias.

**Los autores.**

## **RESUMEN**

### **“SOLUCION DE INTELIGENCIA DE NEGOCIOS EN LA NUBE BAJO LA PLATAFORMA DE MICROSOFT AZURE Y POWER BI PARA EL PROCESO DE VENTAS DE LA EMPRESA BOOK CENTER S.A.C.”**

Por:

Bach. Enrique Estuardo Valera Jara  
Bach. Jorge Daniel Quispe Chávez

Books Center librerías es el día de hoy una empresa referente en el ámbito de venta de libros. Con más de 15 años de experiencia, Books Center se ha convertido en un espejo para otras muchas librerías 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 Book Center S.A.C necesita mantenerse a la vanguardia de la tecnología para ser líder en el norte del País brindando una mejor atención a sus clientes con diagnóstico temprano y certero a la disposición oportuna de información, 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.

Teniendo como problemática que los sistemas con los que cuenta la empresa, no fueron desarrollados con el fin de brindar síntesis, análisis, consolidación, búsquedas de datos y algunos datos se encuentran en diferentes fuentes. Los reportes que se emiten son entregados fuera de tiempo y, a la vez también son analizados fuera de fecha, lo que ocasiona un retraso en la toma de decisiones.

Ante este problema la solución que se propone solución de inteligencia de negocios en la nube bajo la plataforma de Microsoft Azure y Power BI para el proceso de ventas de la empresa que apoye a la Gerencia brindando información valiosa que dé soporte a la hora de tomar decisiones. Para cumplir con esta necesidad, se utilizó la metodología de Ralph Kimball, SQL Server 2012, Microsoft Azure en la implementación del Data Mart dando como resultado el trabajo que se presenta en este proyecto. Permitted brindar un mejor análisis de los datos que maneja la empresa. Este análisis es dinámico, permite que los datos sean accesibles de la forma como el usuario piensa analizarlo. El Data Mart permite a la Gerencia desarrollar estrategias más efectivas.

## **ABSTRACT**

### **BUSINESS INTELLIGENCE SOLUTION IN THE CLOUD IN THE MICROSOFT AZURE PLATFORM AND POWER BI SALES PROCESS FOR ENTERPRISE BOOK CENTER SAC**

**By:**

Bach. Enrique Estuardo Valera Jara

Bach. Jorge Daniel Quispe Chávez

Books Center libraries is today a leading company in the field of book sales. With over 15 years of experience, Books Center has become a mirror for many other libraries for their care system, modern and easily adaptable to market conditions management, prioritizing, before, now and then, customer needs, giving to it, the opportunity to get the best price culture.

Book Center SAC The company needs to stay ahead of the technology to be leader in the north of the country providing a better service to its clients with accurate early diagnosis and the timely provision of information, enabling better control and management of them. That is why we present this technological tool; that would provide support to enable decision making for the company with reliable, accurate and timely.

Taking as problematic that the systems available to the company were not developed in order to provide synthesis, analysis, consolidation, data searches and some data are from different sources. The reports that are issued are delivered out of time and, at the same time are also analyzed out of date, causing a delay in the decision making.

Faced with this problem the solution business intelligence solution in the cloud proposed under the Microsoft Azure platform and BI Power to the sales process of the company to support the Management providing valuable information that supports the decision-making .

To meet this need, the methodology of Ralph Kimball, SQL Server 2012, Microsoft Azure was used in the implementation of the Data Mart resulting in the work presented in this project. Allowing provide a better analysis of the data handled by the company. This analysis is dynamic,

allows data to be accessible to the way the user thinks analysis. The Data Mart allows management to develop more effective strategies.

## 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 .....	xiii

<b>INTRODUCCIÓN.....</b>	<b>xv</b>
--------------------------	-----------

### CAPITULO I: FUNDAMENTO TEÓRICO , MATERIALES Y METODOS

<b>1.1. FUNDAMENTO TEÓRICO.....</b>	<b>01</b>
1.1.1. Proceso de Toma de Decisiones .....	01
1.1.2. Tecnologías de Información .....	02
1.1.3. Sistema de Información .....	03
1.1.4. Procesamiento Transaccional en Línea (OLTP) .....	08
1.1.5. Datawarehouse.....	09
1.1.6. Introducción a Datamarts .....	11
1.1.7. Arquitectura del Datawarehouse .....	12
1.1.8. Introducción al Procesamiento Analítico en Línea .....	13
1.1.9. Arquitectura OLAP .....	15
1.1.10. El modelo de datos OLAP .....	16
1.1.11. Implementación del OLAP .....	19
1.1.12. MICROSOFT AZURE .....	22

1.1.13. Business Intelligence (BI) .....	26
1.1.14. Microsoft SQL Server 2012 .....	37
1.1.15. Analysis Services 2012.....	29
<b>1.2. METODOLOGÍA .....</b>	<b>31</b>
1.2.1. Metodología de Ralph Kimball.....	31

## **CAPITULO II: DESARROLLO DEL PROYECTO**

<b>2.1. PLANIFICACIÓN DEL PROYECTO .....</b>	<b>36</b>
2.1.1. Selección de la Estrategia de Implementación .....	36
2.1.2. Selección de la Metodología de Desarrollo .....	36
2.1.3. Selección del Ámbito de Implementación.....	36
2.1.4. Selección del enfoque arquitectónico .....	37
2.1.5. Desarrollo de un Programa y del Presupuesto del Proyecto .....	38
2.1.5.1. Presupuesto .....	38
2.1.5.2. Cronograma de actividades .....	39
2.1.6. Desarrollo del escenario del uso empresarial .....	39
<b>2.2. DETERMINACIÓN DE LOS REQUERIMIENTOS.....</b>	<b>41</b>
2.2.1. Requerimientos del Propietario .....	41
2.2.2. Requerimientos del Usuario Final.....	42
2.2.3. Requerimientos no Funcionales .....	43
2.2.4. Análisis de los requerimiento.....	44
<b>2.3. DISEÑO TÉCNICO DE LA ARQUITECTURA.....</b>	<b>50</b>
2.3.1. Nivel de Datos .....	50
2.3.2. Nivel Técnico .....	51
<b>2.4. MODELADO DIMENSIONAL .....</b>	<b>51</b>
2.4.1. Identificación de los Componentes del Modelo .....	52
2.4.2. Diagrama de la Tabla de Hechos .....	55
2.4.2.1. Esquema Estrella.....	60
<b>2.5. DISEÑO FÍSICO .....</b>	<b>61</b>
2.5.1. Determinación de las agregaciones .....	66

2.5.2. Construcción de las Tablas y la Base de Datos en SQL.....	67
<b>2.6. PROCESO DE EXTRACCIÓN, TRANSFORMACIÓN Y CARGA DE DATOS .....</b>	<b>70</b>
<b>2.7. SELECCIÓN DE PRODUCTOS.....</b>	<b>91</b>
2.7.1. Hardware .....	91
2.7.2. Software .....	92
<b>2.8. ESPECIFICACIÓN DE LA APLICACIÓN DEL USUARIO FINAL .....</b>	<b>92</b>
2.8.1. Estructura de Cubo .....	92
<b>2.9. DESARROLLO DE LA APLICACIÓN DEL USUARIO FINAL .....</b>	<b>96</b>
<b>CAPITULO III: DISCUSIÓN DE LA HIPOTESIS.....</b>	<b>111</b>
<b>CONCLUSIONES.....</b>	<b>120</b>
<b>RECOMENDACIONES.....</b>	<b>121</b>
<b>REFERENCIAS.....</b>	<b>122</b>
<b>ANEXOS .....</b>	<b>123</b>

## ÍNDICE DE FIGURAS

	Pág.
<b>Figura Nº 1:</b> Tipos de Sistemas de Información.....	04
<b>Figura Nº 2:</b> TPS Nomina de Pagos .....	05
<b>Figura Nº 3:</b> Modelo de un ESS .....	08
<b>Figura Nº 4:</b> Estructura básica Datawarehouse .....	13
<b>Figura Nº 5:</b> Arquitectura básica para OLAP .....	15
<b>Figura Nº 6:</b> Un esquema de estrella .....	17
<b>Figura Nº 7:</b> Un esquema de copo de nieve ... ..	18
<b>Figura Nº 8:</b> Cubos OLAP .....	22
<b>Figura Nº 9:</b> Esquemas Relacionales y cubo ... ..	23
<b>Figura Nº 10:</b> Jerarquías de los cubos OLAP .....	24
<b>Figura Nº 11:</b> Medidas de un cubo.....	25
<b>Figura Nº 12:</b> Business Intelligence .....	26
<b>Figura Nº 13:</b> Metodología de implementación .....	27
<b>Figura Nº 14:</b> Cronograma de actividades .....	39
<b>Figura Nº 15:</b> Análisis del requerimiento R1 .....	44
<b>Figura Nº 16:</b> Análisis del requerimiento R2 .....	44
<b>Figura Nº 17:</b> Análisis del requerimiento R3 .....	45
<b>Figura Nº 18:</b> Análisis del requerimiento R4 .....	45
<b>Figura Nº 19:</b> Análisis del requerimiento R5 .....	45
<b>Figura Nº 20:</b> Análisis del requerimiento R6 .....	46
<b>Figura Nº 21:</b> Análisis del requerimiento R7.....	46
<b>Figura Nº 22:</b> Análisis del requerimiento R8 .....	46
<b>Figura Nº 23:</b> Análisis del requerimiento R9 .....	47
<b>Figura Nº 24:</b> Análisis del requerimiento R10 .....	47
<b>Figura Nº 25:</b> Modelo Startnet Ventas .....	48
<b>Figura Nº 27:</b> Diagrama E-R de la base de datos.....	50
<b>Figura Nº 28:</b> Enfoque Arquitectónico del Datamart .....	51
<b>Figura Nº 29:</b> Diagrama de Tabla de Hechos Ventas .....	55
<b>Figura Nº 31:</b> Tabla de Hechos del Data Mart .....	57
<b>Figura Nº 32:</b> Jerarquía de Atributos de la Dimensión Tiempo.....	58
<b>Figura Nº 33:</b> Jerarquía de los Atributos de la Dimensión Cliente .....	59
<b>Figura Nº 34:</b> Jerarquía de los Atributos de la Dimensión Producto.....	59
<b>Figura Nº 35:</b> Tabla de Hechos Ventas .....	57
<b>Figura Nº 37:</b> Esquema Estrella del Data Mart .....	61

<b>Figura Nº 38:</b> Claves Foráneas de las Tablas de Hechos .....	65
<b>Figura Nº 39:</b> Diseño de la Físico de la Base de Datos del Data Mart .....	66
<b>Figura Nº 40:</b> Tabla de la Dimensión Tiempo Ventas .....	67
<b>Figura Nº 43:</b> Tabla de la Dimensión Cliente.....	68
<b>Figura Nº 44:</b> Tabla de la Dimensión Vendedor.....	68
<b>Figura Nº 45:</b> Tabla de la Dimensión Pago .....	68
<b>Figura Nº 46:</b> Tabla de la Dimensión Producto .....	69
<b>Figura Nº 47:</b> Data Mart.....	69
<b>Figura Nº 48:</b> Workflow de los Pasos de Transformación.....	71
<b>Figura Nº 49:</b> Diagrama Workflow con Restricciones de Precedencia .....	73
<b>Figura Nº 50:</b> ETL del DataMart... ..	74
<b>Figura Nº 51:</b> Ventana de Código para el Paso Limpieza Total .....	76
<b>Figura Nº 52:</b> Poblamiento de la Dimensión Tiempo Ventas .....	78
<b>Figura Nº 55:</b> Poblamiento de la Dimensión Vendedor.....	82
<b>Figura Nº 56:</b> Poblamiento de la Dimensión Cliente.....	83
<b>Figura Nº 57:</b> Poblamiento de la Dimensión Pago.....	84
<b>Figura Nº 58:</b> Poblamiento de la Dimensión Producto.....	86
<b>Figura Nº 59:</b> Poblamiento de la Tabla de Hechos Ventas.....	89
<b>Figura Nº 61:</b> Diseño de Cubo de Ventas .....	93
<b>Figura Nº 62:</b> Medidas y Dimensiones del Cubo Ventas .....	94
<b>Figura Nº 65:</b> Diseño de Interfaces .....	96

## ÍNDICE DE TABLAS

	Pág.
<b>Tabla Nº 1:</b> Diferencia entre Sistemas Transaccionales y Datawarehouse .....	10
<b>Tabla Nº 2:</b> Personal involucrado en el Proyecto .....	40
<b>Tabla Nº 3:</b> Descripción de usuarios del Data Mart .....	41
<b>Tabla Nº 4:</b> Consultas y medidas .....	42
<b>Tabla Nº 5:</b> Matriz de Intersecciones.....	53
<b>Tabla Nº 6:</b> Objetivos de las tablas de Hecho .....	54
<b>Tabla Nº 7:</b> Dimensiones de la tabla de Hechos Ventas.....	54
<b>Tabla Nº 9:</b> Elección de Hechos .....	55
<b>Tabla Nº 10:</b> Detalle de las claves de las dimensiones .....	56
<b>Tabla Nº 11:</b> Detalle de las Medidas de las Tablas de Hecho .....	57
<b>Tabla Nº 13:</b> Detalle de Dimensión Vendedor.....	58
<b>Tabla Nº 14:</b> Detalle de Dimensión Cliente .....	59
<b>Tabla Nº 15:</b> Detalle de Dimensión Producto .....	60
<b>Tabla Nº 17:</b> Detalle de Dimensión Pago .....	60
<b>Tabla Nº 18:</b> Nombres estándares para las Tablas Hechos y Dimensiones.....	62
<b>Tabla Nº 19:</b> Nombres estándares para los atributos de las Dimensiones .....	62
<b>Tabla Nº 20:</b> Nombres estándares para los atributos de Tablas de Hechos .....	63
<b>Tabla Nº 21:</b> Tipo de Dato para Dim_Tiempo_ Ventas .....	63
<b>Tabla Nº 22:</b> Tipo de Dato para Dim_Vendedor .....	63
<b>Tabla Nº 23:</b> Tipo de Dato para Dim_Cliente .....	64
<b>Tabla Nº 25:</b> Tipo de Dato para Dim_Producto .....	64
<b>Tabla Nº 26:</b> Tipo de Dato para Dim_Pago .....	64
<b>Tabla Nº 27:</b> Determinación de las Agregaciones.....	66
<b>Tabla Nº 28:</b> Software para el proceso de Diseño del DataMart.....	92

## INTRODUCCIÓN

En todas las organizaciones son tomadas decisiones a diario, trascendentes o intrascendentes, pero todas ellas sin estar exentas de riesgo. Quienes deben tomar las decisiones, requieren de minimizar este riesgo, teniendo a mano la mayor cantidad de información, la cual debe ser oportuna, eficiente y además, que agregue valor.

Es fundamental también que en toda empresa los procesos estén definidos y más aún cuando están correctamente relacionados entre ellos. Las Herramientas que permiten que los procesos de la empresa puedan ser integrados satisfactoriamente son las tecnologías de información.

Books Center librerías es el día de hoy una empresa referente en el ámbito de venta de libros. Con más de 15 años de experiencia, Books Center se ha convertido en un espejo para otras muchas librerías 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 Book Center S.A.C necesita mantenerse a la vanguardia de la tecnología para ser líder en el norte del País brindando una mejor atención a sus clientes con diagnóstico temprano y certero a la disposición oportuna de información, 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.

Book Center S.A.C ha decidido evaluar la implementación de un sistema 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. La empresa Book Center S.A.C. no quiere permanecer ajena al uso de estas herramientas, ya que tiene la necesidad de información confiable y oportuna que permita a la gerencia tomar mejores decisiones, en cualquier momento y desde cualquier lugar para ello se busca crear una solución de inteligencia de negocios en la nube.

Cada vez es más frecuente que los requerimientos cambien a medida que un proyecto progresa. Al mismo tiempo el cliente demanda resultados rápidos, que puedan implementarse y medirse en períodos cortos de tiempo.

Es por ello que en los últimos tiempos, han ganado mucho terreno los esquemas de trabajo, también llamados metodologías, de naturaleza evolutivos y ágiles, cuyas principales premisas son el trabajo altamente colaborativo, iterativo e incremental.

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 Book Center SAC 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 procedimiento que se realiza para estos reportes solicitados es:

- El Jefe de Ventas solicita datos sobre el área de ventas, que deben ser reportados con datos exactos y con gráficos en Excel, esto se realiza con la ayuda del personal de sistemas, ya que son los que manejan la Base de Datos de la empresa.
- Obtener los datos solicitados demanda un tiempo importante, pudiendo ser hasta de 3 o 4 días.
- Luego de esta espera, ya con los datos recibidos del personal de sistemas se procede a realizar otro análisis en Excel con sus respectivos gráficos.
- Esto pasa al jefe de ventas, quien procede al análisis de la información, pero es muy posible que se necesite contar con más información de meses anteriores y poder realizar comparaciones, lo cual significa repetir el ciclo mencionado.
- Este proceso se tiene que realizar cada vez que se requiera de un reporte, lo que trae consigo los siguientes problemas:
  - ✓ Pérdida de tiempo y esfuerzo por parte del personal del área de sistemas en la emisión de los datos solicitados, los cuales son entregados para realizar el reporte.
  - ✓ Demora en la entrega de la información, lo cual genera retraso en la toma de decisiones.
  - ✓ La generación de reportes sobre la producción de la agencia es muy dificultosa e ineficiente.
  - ✓ Difícil acceso a reportes históricos.
  - ✓ La información se encuentra dispersa dentro de la bases de datos y, en consecuencia, no se puede explotar en conjunto.
  - ✓ Solo es posible el análisis de los datos actuales, no mantiene valores históricos para tratar evolutivos.
  - ✓ Se obstaculiza el trabajo transaccional diario de los sistemas de información originales, no permite el análisis on-line.
  - ✓ La base de datos está diseñada para el trabajo transaccional, no para el análisis de los datos.
  - ✓ La información no es validada y en muchas ocasiones hay pérdida de ella en los informes.

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.

Estos problemas pueden ser expresados en una sola pregunta:

**¿Cómo mejorar el soporte online en la toma de decisiones para el área de ventas de la empresa Book Center S.A.C, 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 Book Center S.A.C.**

El objetivo general es **Implementar una solución de Inteligencia de Negocios en la nube bajo la Plataforma de Microsoft Azure y Power BI para el proceso de Ventas de la empresa Book Center S.A.C.**

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
- ✓ Implementar una solución ETL para los problemas de integración de datos usando SQL Server 2012.
- ✓ 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:** 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. 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".

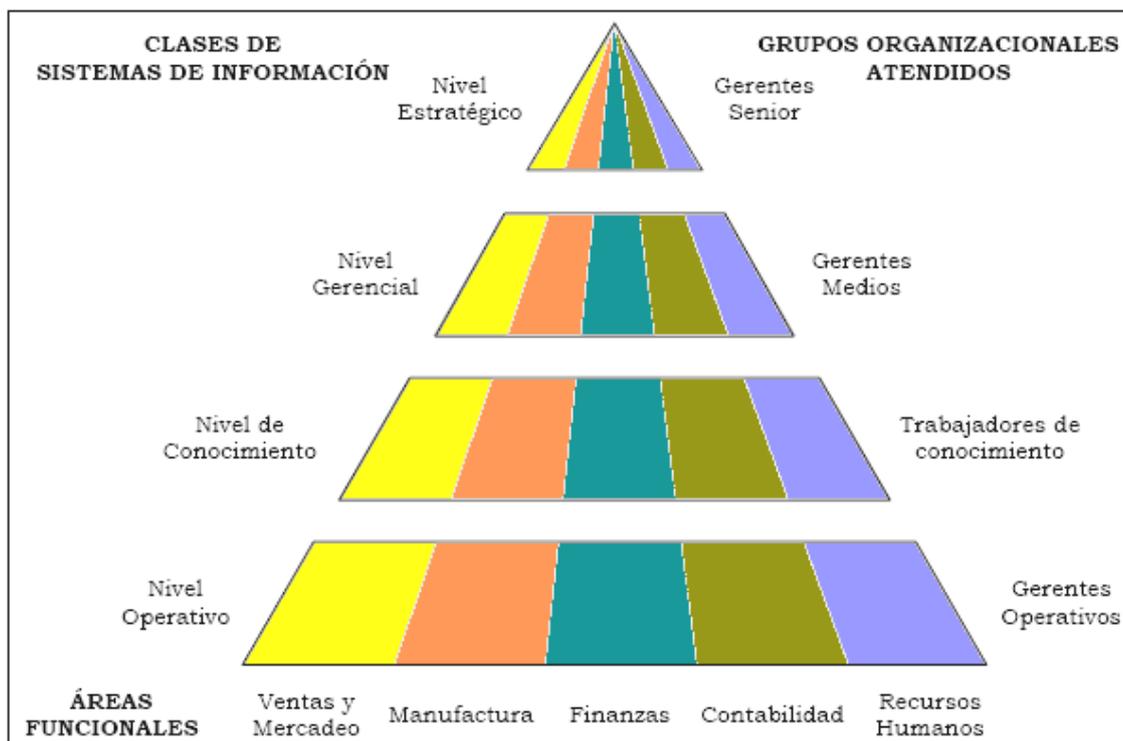
Por lo tanto podemos definir un sistema de información como un conjunto de subsistemas que incluyen hardware, software, medios de almacenamiento de datos ya sea primarios, secundarios y bases de datos relacionadas entre sí con el fin de procesar entradas para realizar transformaciones a esas entradas 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, 2012)**

#### **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 nomina, facturación, entrega de mercancía y deposito de cheques. Estas transacciones varían de acuerdo al tipo de empresa.

Los sistemas de procesamiento de transacción o TPS (Transaccion Procesation System) por sus siglas en ingles, 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

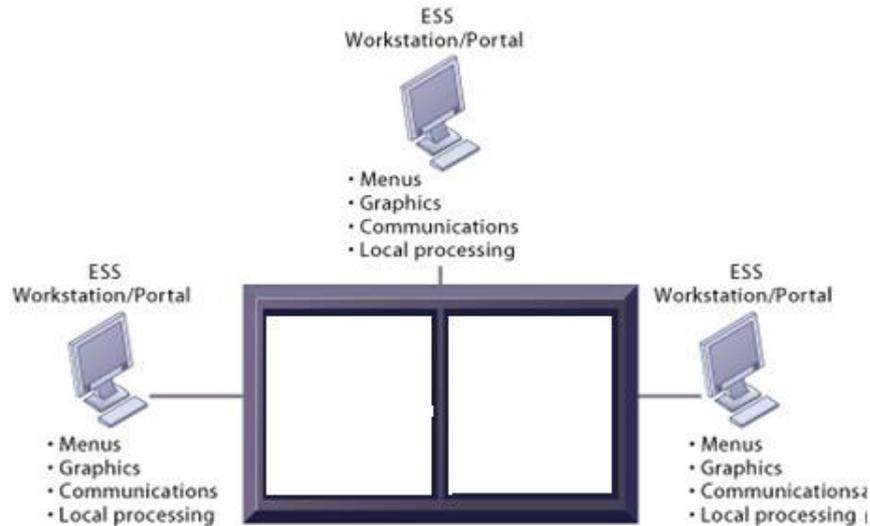
directamente, estos sistemas incluyen explícitamente el software de uso fácil. DSS son interactivos, el usuario puede cambiar las hipótesis, pedir nuevas preguntas, e incluir nuevos datos. (Laudon y Laudon, 2013)

#### **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)

### 1.1.2. 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 cambió 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 cuando 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".

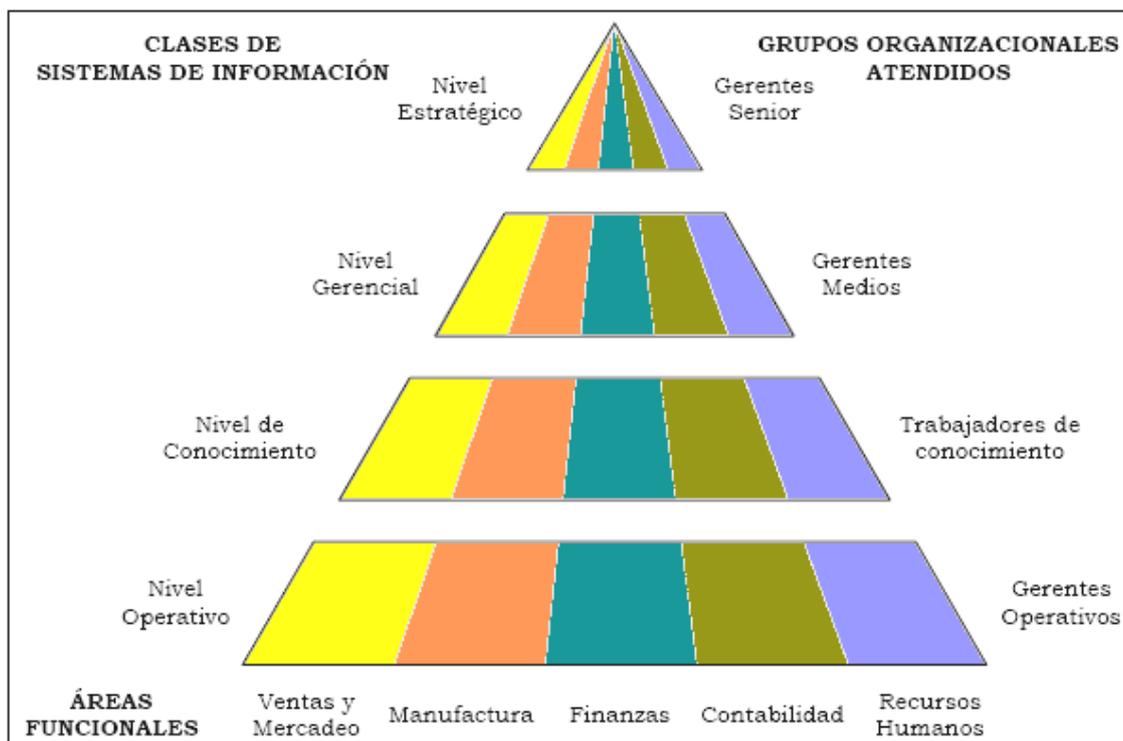
Por lo tanto podemos definir un sistema de información como un conjunto de subsistemas que incluyen hardware, software, medios de almacenamiento de datos ya sea primarios, secundarios y bases de datos relacionadas entre sí con el fin de procesar entradas para realizar transformaciones a esas entradas 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, 2012)

#### 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]

**e. 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 nomina, facturación, entrega de mercancía y deposito de cheques. Estas transacciones varían de acuerdo al tipo de empresa.

Los sistemas de procesamiento de transacción o TPS (Transaccion Procesation System) por sus siglas en ingles, 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)

**f. 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)

**g. 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

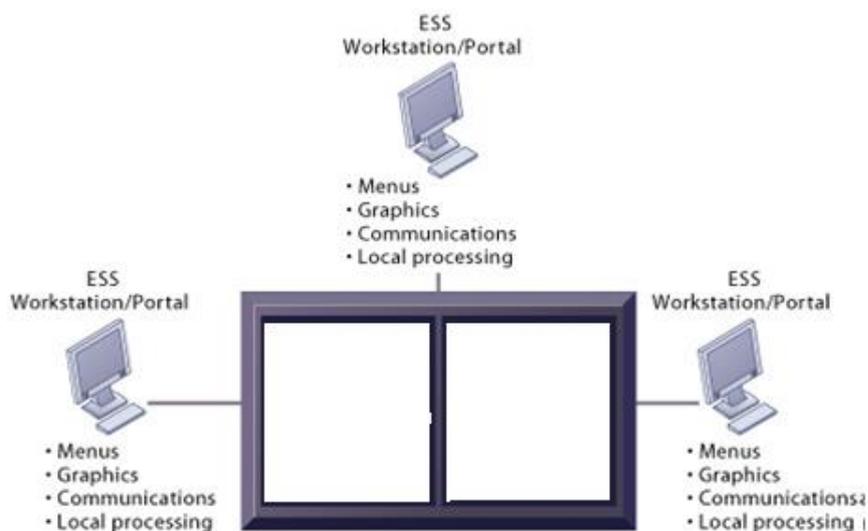
directamente, estos sistemas incluyen explícitamente el software de uso fácil. DSS son interactivos, el usuario puede cambiar las hipótesis, pedir nuevas preguntas, e incluir nuevos datos. (Laudon y Laudon, 2013)

#### **h. 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 inglés), 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

<b>Transaccionales</b>	<b>Basados en Datawarehouse</b>
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, sumarizados, 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:

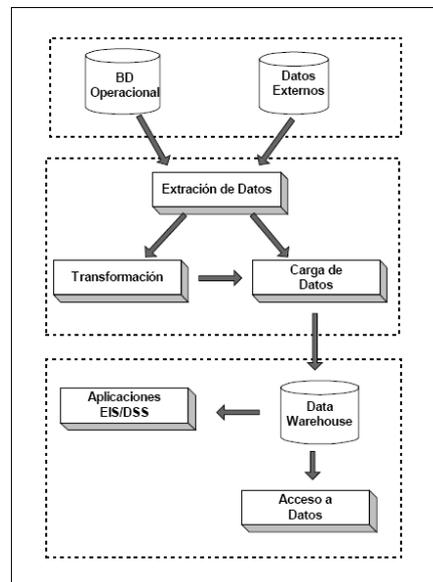
- Acelerar las consultas reduciendo el volumen de datos a recorrer
- Estructurar los datos para su adecuado acceso por una herramienta
- Dividir los datos para imponer estrategias de control de acceso
- Segmentar los datos en diferentes plataformas hardware
- Permite el acceso a los datos por medio de 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.

En segundo lugar, el OLAP acelera la entrega de información a los usuarios finales que ven estas estructuras de datos como cubos denominadas multidimensionales debido a que la información es vista en varias dimensiones.

Esta entrega es optimizada ya que se preparan algunos valores calculados en los datos por adelantado, en vez de realizar el cálculo al momento de la solicitud. La combinación de navegación fácil y rápida le permite a los usuarios ver y analizar información más rápida y eficientemente que lo que es posible con tecnologías de bases

de datos relacionales solamente. El resultado final: se pasa más tiempo analizando los datos y menos tiempo analizando las bases de 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.

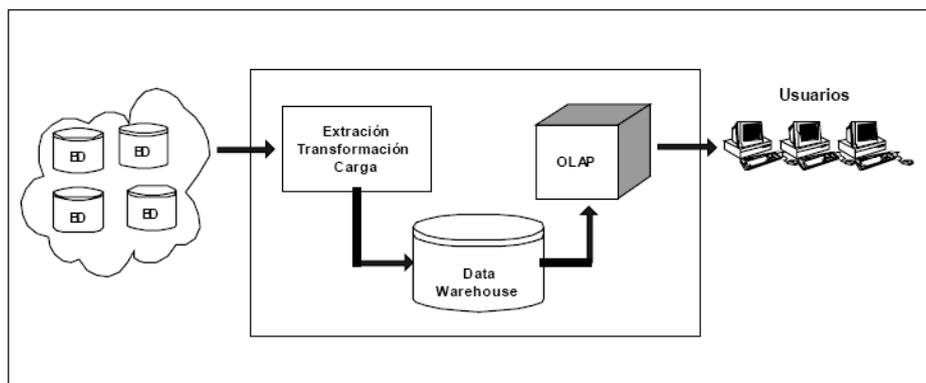
El OLAP es un componente clave en el proceso de almacenamiento de datos (data warehousing) y los servicios OLAP proporcionan la funcionalidad esencial para una gran variedad de aplicaciones que van desde reportes corporativos hasta soporte avanzado de decisiones.

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.

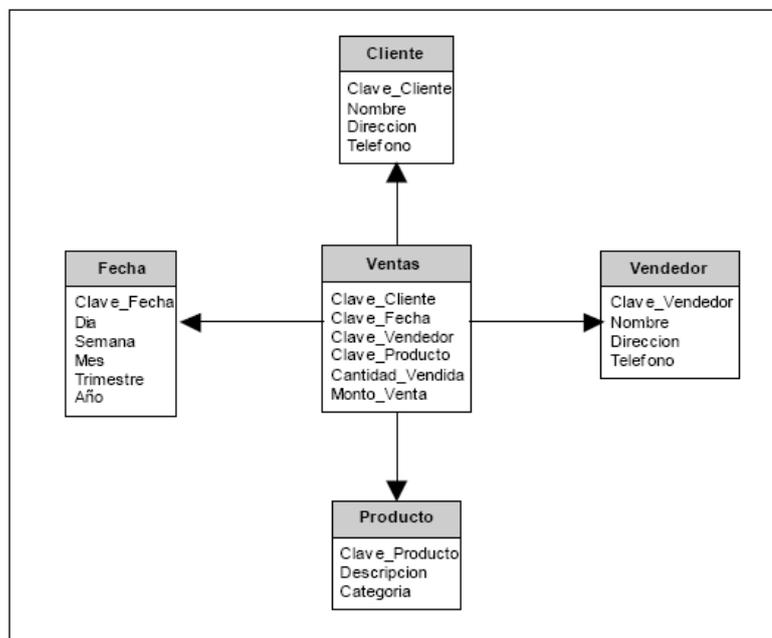
Este es un paso vital en el proceso, que asegura que los datos que son vistos por el usuario OLAP son correctos, consistentes y que llenan las definiciones organizacionales para los datos.

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°6:** Un esquema de estrella.

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

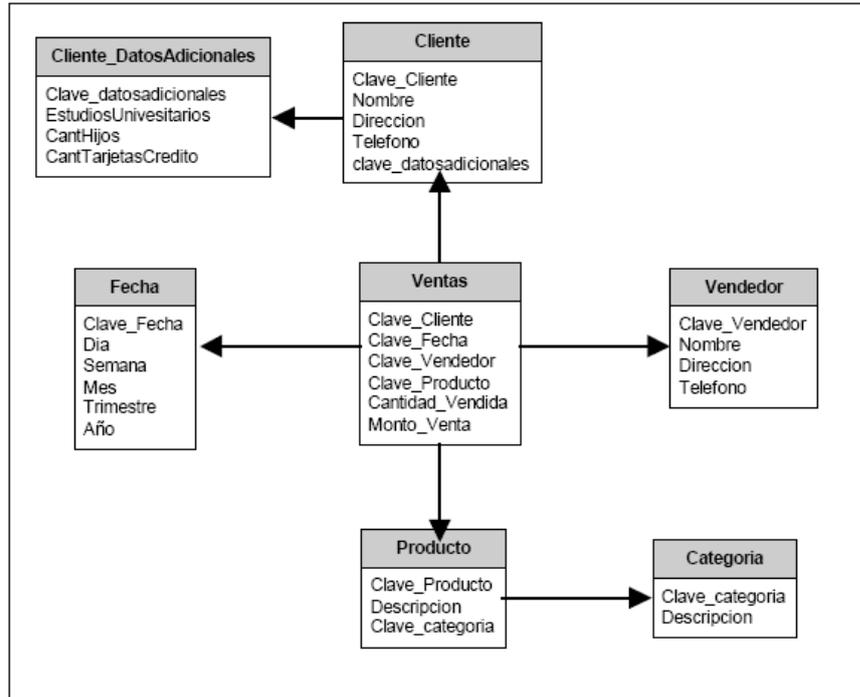
A continuación se enumeran algunas de las principales ventajas del esquema estrella.

- 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 el esquema de copo de nieve son mejorar el desempeño de consultas debido a un 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° 7:** 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 , 2012)**

### 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.

La modelación multidimensional de datos es una forma de facilitar el análisis empresarial en línea y de mejorar el rendimiento de las consultas.

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 , 2012)**.

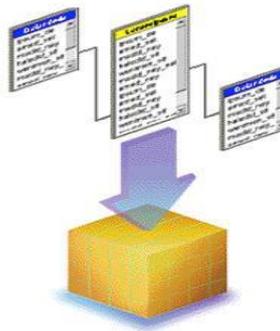


**Figura N° 8:** Cubos OLAP.

**Fuente:** (Root y Mason , 2012)

## 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 , 2012).



**Figura N° 9: Esquemas Relacionales y cubo**

**Fuente:** (Root y Mason , 2012)

## DIMENSIONES DE UN CUBO

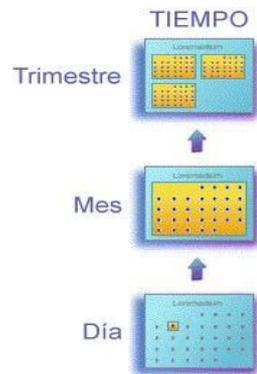
Las dimensiones de un cubo representan distintas categorías para analizar datos empresariales. Categorías tales como fecha, geografía o línea de productos son dimensiones típicas de cubos.

Nota: los cubos no están limitados a tres dimensiones. Pueden contener hasta 64 dimensiones. (Root y Mason , 2012)

## 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° 10:**

Jerarquías de los cubos

OLAP

**Fuente:** (Root y Mason , 2012)

### 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 , 2012)

#### **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.

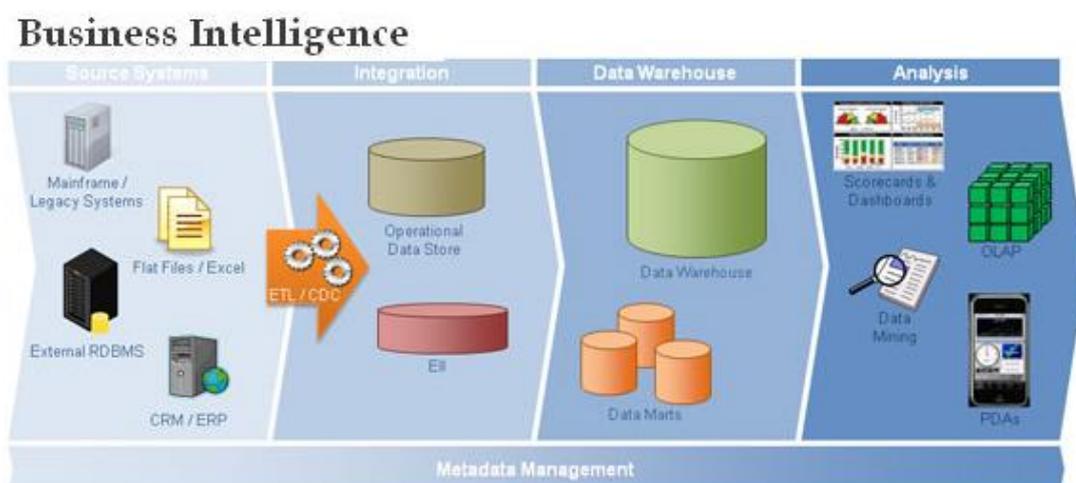
Los cubos son la forma en que la tecnología OLAP organiza los datos de resumen en estructuras multidimensionales. Las dimensiones y sus niveles jerárquicos reflejan las consultas que se pueden hacer al cubo. Los agregados se almacenan en la estructura multidimensional en celdas que tienen las coordenadas especificadas por las dimensiones. (Root y Mason , 2012)

#### **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

Microsoft SQL Server 2014 se basa en las funciones críticas ofrecidas en la versión anterior, proporcionando un rendimiento, una disponibilidad y una facilidad de uso innovadores para las aplicaciones más importantes. Microsoft SQL Server 2014 ofrece nuevas capacidades en memoria en la base de datos principal para el procesamiento de transacciones en línea (OLTP) y el almacenamiento de datos, que complementan nuestras capacidades de almacenamiento de datos en memoria y BI existentes para lograr la solución de base de datos en memoria más completa del mercado.

SQL Server 2014 también proporciona nuevas soluciones de copia de seguridad y de recuperación ante desastres, así como de arquitectura híbrida con Windows Azure, lo que permite a los clientes utilizar sus actuales conocimientos con características locales

que aprovechan los centros de datos globales de Microsoft. Además, SQL Server 2014 aprovecha las nuevas capacidades de Windows Server 2012 y Windows Server 2012 R2 para ofrecer una escalabilidad sin parangón a las aplicaciones de base de datos en un entorno físico o virtual.

### 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 SQL AZURE

**Microsoft SQL Azure** es un servicio de base de datos relacional (RDBMS) en la nube basado en tecnologías de SQL Server. Se trata de un servicio de base de datos de alta disponibilidad y escalable, en modalidad de multi-propiedad en la nube, gestionado por Microsoft. La Base de Datos SQL Azure facilita las tareas de aprovisionamiento y despliegue de múltiples bases de datos. Los desarrolladores no tienen que instalar, configurar, aplicar parches o administrar software alguno, de todo eso se encarga Microsoft con su solución de “Plataforma como Servicio” (PaaS). La alta disponibilidad y tolerancia a fallos son características propias de la solución y no necesitan administración física.

Los clientes pueden utilizar su conocimiento actual en desarrollo sobre T-SQL y el modelo conocido de datos relacionales que han adquirido y aplican a sus bases de datos in-situ. Además pueden poner en marcha cualquier solución sobre SQL Azure de manera rápida utilizando las mismas herramientas de desarrollo y gestión que emplean en sus servicios de bases de datos en la empresa.

#### **Utilice la base de datos SQL Azure para:**

- Ampliar sus oportunidades de negocio y alcanzar nuevos mercados creando aplicaciones en modo “Software como Servicio” (SaaS) de manera rápida y eficiente.
- Crear aplicaciones a medida, para la empresa y la Web, que aprovechen las capacidades y la escalabilidad de la base de datos.
- Ampliar el alcance de sus activos de datos y consolidar orígenes de datos en la nube, y con ello mejorar la colaboración entre partners, oficinas remotas, empleados en movilidad y dispositivos móviles.
- Optimizar recursos consolidando en la nube las bases de datos departamentales actualmente en sus oficinas, y aprovisionando nuevas bases de datos en un tiempo mínimo.
- Desarrollar soluciones que ofrezcan nuevas panorámicas sobre el rendimiento y los procesos internos de su empresa, integrando múltiples fuentes de datos con sus actuales herramientas de análisis.

**Ventajas principales de SQL Azure:**

- No requiere administración física: la instalación de software y actualizaciones ya están incluidas en el servicio, puesto que se trata de una solución PaaS (*Platform as a Service*)
- Alta disponibilidad y tolerancia a fallos integrados en la solución.
- Proceso sencillo de aprovisionamiento y despliegue de múltiples bases de datos.
- Escalabilidad de las bases de datos, que se adaptan a sus necesidades de negocio.
- Multipropiedad.
- Integración con SQL Server y las herramientas de Visual Studio®.
- Soporte para el conocido modelo de base de datos relacional basado en T-SQL.
- Opción de contrato de pago por uso (“*pay-as-you-go*”).

**CARACTERÍSTICAS PRINCIPALES:****Servicios de Gestión de Bases de Datos Relacionales (RDBMS)**

- Creación, acceso y manipulación de tablas, vistas, índices, roles, procedimientos almacenados, triggers y funciones.
- Ejecución de queries complejas y sentencias join entre múltiples tablas.
- Funciones Insert, Update y Delete.
- Restricciones.
- Transacciones.
- Tablas temporales.
- Funciones básicas (agregados, funciones matemáticas, de cadena, fecha/hora).
- Un subconjunto de los actuales procedimientos almacenados y vistas del sistema integrados en SQL Server.
- Soporte para métricas de facturación controlables en tiempo real y para análisis históricos.

**Capacidad de programación**

- Acceso a datos ADO.NET gestionado.
- ODBC nativo.
- Soporte para PHP.

## Herramientas

- Aprovisionamiento de servidores y bases de datos lógicas mediante el portal de registro de SQL Azure Database .
- SQL Server Management Studio: un entorno integrado con herramientas gráficas para el acceso y configuración de bases de datos de SQL Server y SQL Azure: 32-bit | 64 bit
- Herramienta Web de gestión de bases de datos Project Houston.
- Soporte para el despliegue mediante la ejecución de scripts de configuración de SQL con SQLCMD.Si desea más información sobre las herramientas disponibles, lea este artículo.

### 1.1.19. MICROSOFT AZURE

MICROSOFT Azure es un sistema operativo de servicios en la nube que actúa como entorno de desarrollo, hospedaje y administración de servicios de la plataforma Windows Azure. Windows Azure proporciona a los desarrolladores servicios informáticos y de almacenamiento a petición para hospedar, escalar y administrar aplicaciones web de Internet a través de los centros de datos de Microsoft. (Brunetti, 2011)

Windows Azure es una plataforma flexible que admite varios idiomas y se integra con el entorno local existente. Para compilar aplicaciones y servicios en Windows Azure, los desarrolladores pueden usar su experiencia previa con Microsoft Visual Studio. Asimismo, Windows Azure es compatible con los estándares, protocolos y lenguajes más frecuentes, como SOAP, REST, XML, Java, PHP y Ruby.

#### Implementación

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 customizada 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.

### **Copias de seguridad**

Dado que la tecnología puede fallar, Windows Azure ofrece una manera de proteger la información importante con una copia de seguridad automática dentro de un servicio de almacenamiento. Las copias de seguridad quedan cifradas antes de la transmisión y se almacenan cifradas en Windows Azure. Estas copias de seguridad están fuera de sitio, lejos de su centro de datos, lo que reduce la necesidad de asegurar y proteger los medios de copia de seguridad en el lugar.

La administración de copias de seguridad en la nube usa herramientas de copia de seguridad conocidas en Windows Server, Windows Server Essentials, o el Administrador de System Center Data Protection. Estas herramientas proporcionan experiencias similares al configurar, supervisar y recuperar copias de seguridad ya sea en el disco local o el almacenamiento de Windows Azure, o puede utilizar el software propio del agente.

Después de que los datos se copian a la nube, los usuarios autorizados pueden recuperar fácilmente copias de seguridad de cualquier servidor. También se pueden utilizar Copias de seguridad incremental para asegurar el uso eficiente de almacenamiento y un menor consumo de ancho de banda, al mismo tiempo que permite la recuperación de punto en el tiempo de varias versiones de los datos.

### **Servicio de Microsoft Azure**

Dentro de la plataforma, el servicio de *Windows Azure* es el encargado de proporcionar el alojamiento de las aplicaciones y el almacenamiento no relacional. Dichas aplicaciones deben funcionar sobre Windows Server 2008 R2. Pueden estar desarrolladas en .NET, PHP, C++, Ruby, Java. Además del servicio de ejecución, dispone de diferentes mecanismos de almacenamiento de datos: tablas NoSQL, blobs, blobs para streaming, colas de mensajes o 'drives' NTFS para operaciones de lectura / escritura a disco.

### **Características de Microsoft Azure**

- **Proceso:** el servicio de proceso de Windows Azure ejecuta aplicaciones basadas en Windows Server. Estas aplicaciones se pueden crear mediante .NET Framework en lenguajes como C# y Visual Basic, o implementar sin .NET en C++, Java y otros lenguajes.
- **Almacenamiento:** objetos binarios grandes (blobs) proporcionan colas para la comunicación entre los componentes de las aplicaciones de Windows Azure y ofrece un tipo de tablas con un lenguaje de consulta simple.
- **Controlador de tejido:** Windows Azure se ejecuta en un gran número de máquinas. El trabajo del controlador de tejido es combinar las máquinas en un solo centro de datos de Windows Azure formando un conjunto armónico. Los servicios de proceso y almacenamiento de Windows Azure se implementan encima de toda esta eficacia de procesamiento.
- **Red de entrega de contenido (CDN):** el almacenamiento en caché de los datos a los que se accede frecuentemente cerca de sus usuarios agiliza el acceso a esos datos.
- **Connect:** organizaciones interactúan con aplicaciones en la nube como si estuvieran dentro del propio firewall de la organización.

### **Azure Platform Componentes:**

- **Windows Azure Compute** es una plataforma para hospedar y administrar aplicaciones en los centros de datos de Microsoft. Una aplicación de Windows Azure consta de uno o varios componentes denominados 'roles.' Los roles pueden ser de tres tipos: rol web, rol de trabajo y rol de máquina virtual (VM).

- Windows Azure Storage tiene servicios de básicos como parte de la cuenta de almacenamiento de Windows Azure. Los blobs, tablas y colas están accesibles a aplicaciones o instancias de aplicaciones simultáneamente.
- Microsoft SQL Azure es un servicio de base de datos en la nube basado en las tecnologías de SQL Server. Los servicios de SQL Azure incluyen: Base de datos SQL Azure, SQL Azure Reporting y SQL Azure Data Sync Aspectos destacados de la base de datos de SQL Azure.
- Content Delivery Network (CDN) de Windows Azure coloca copias de los datos cerca de donde estos se encuentran. La CDN de Windows Azure entrega actualmente muchos productos de Microsoft, como Windows Update, vídeos de Zune y Bing Maps, que los clientes conocen y usan todos los días. Gracias a la incorporación de la CDN a los servicios de Windows Azure, ahora esta red a gran escala está disponible a todos los usuarios de Windows Azure.
- Azure Market Place es un mercado en línea global compartir, comprar y vender aplicaciones SaaS completas y conjuntos de datos. La sección de datos de Windows Azure Marketplace incluye datos, imágenes y servicios Web en tiempo real de proveedores de datos comerciales, líderes en el sector y orígenes de datos públicos acreditados.

Azure Virtual Network es una serie de funciones de red. Windows Azure Connect es la primera característica de Azure Virtual Network que configura la conectividad de red basada en IP entre recursos locales y de Windows Azure. Windows Azure Traffic Manager equilibra la carga del tráfico en servicios hospedados.

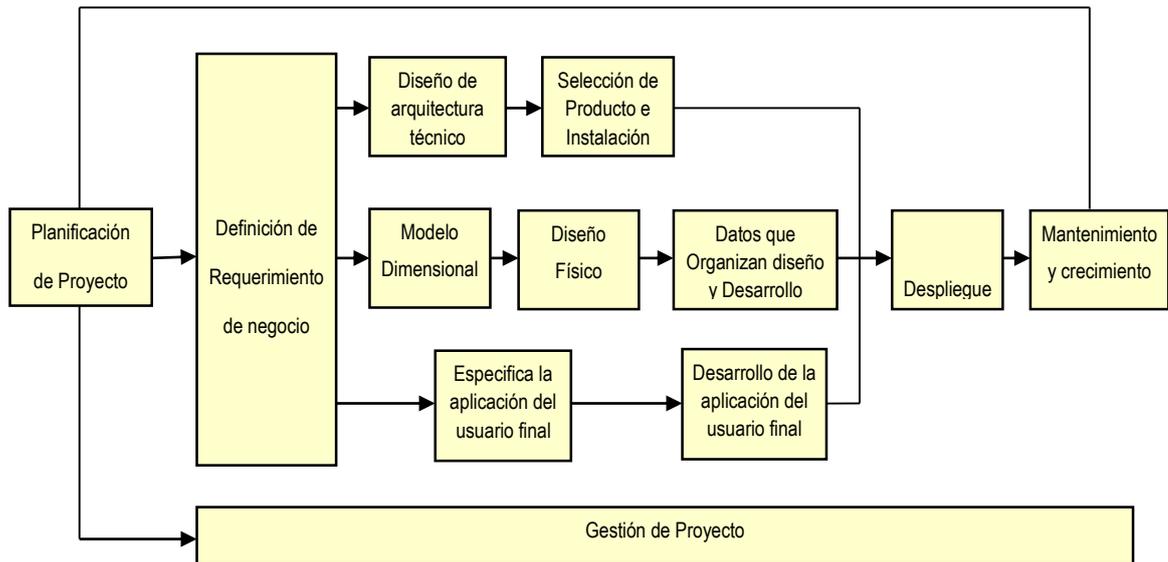
## **1.2. METODOLOGÍA**

### **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.

**1. 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>
<b>Determinación de Requerimientos</b>	<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>
<b>Análisis</b>	<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.

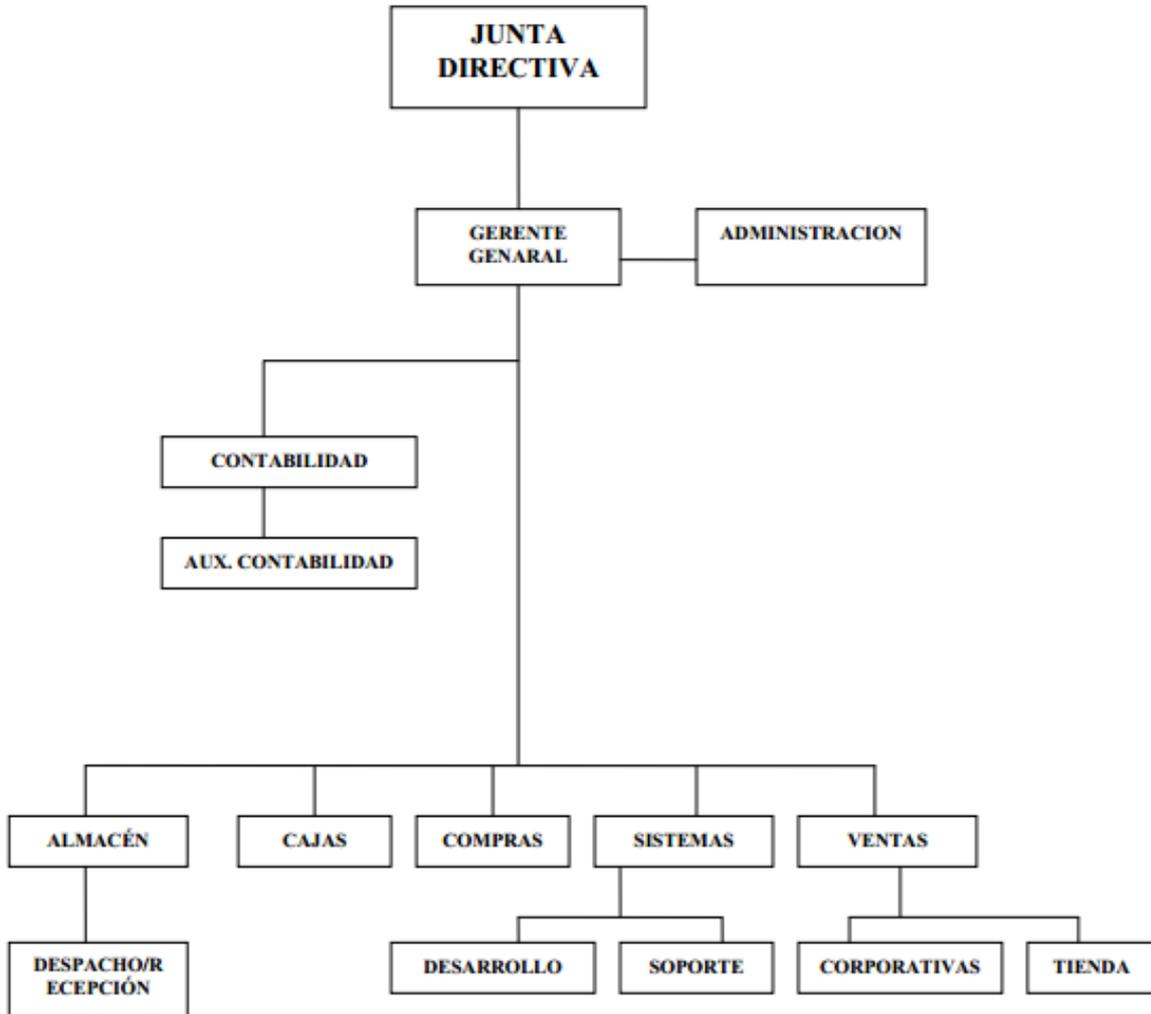
### DESARROLLO DEL PROYECTO DE TESIS

#### 2.1. Planificación del Proyecto

##### 2.1.1. El negocio

- **Descripción :**
  - ✓ Razón Social: BOOK CENTER S.A.C
  - ✓ 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
  - ✓ Competidores : CopyVentas
  
- **VISIÓN:** Ser la organización N° 1 en el norte del 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.

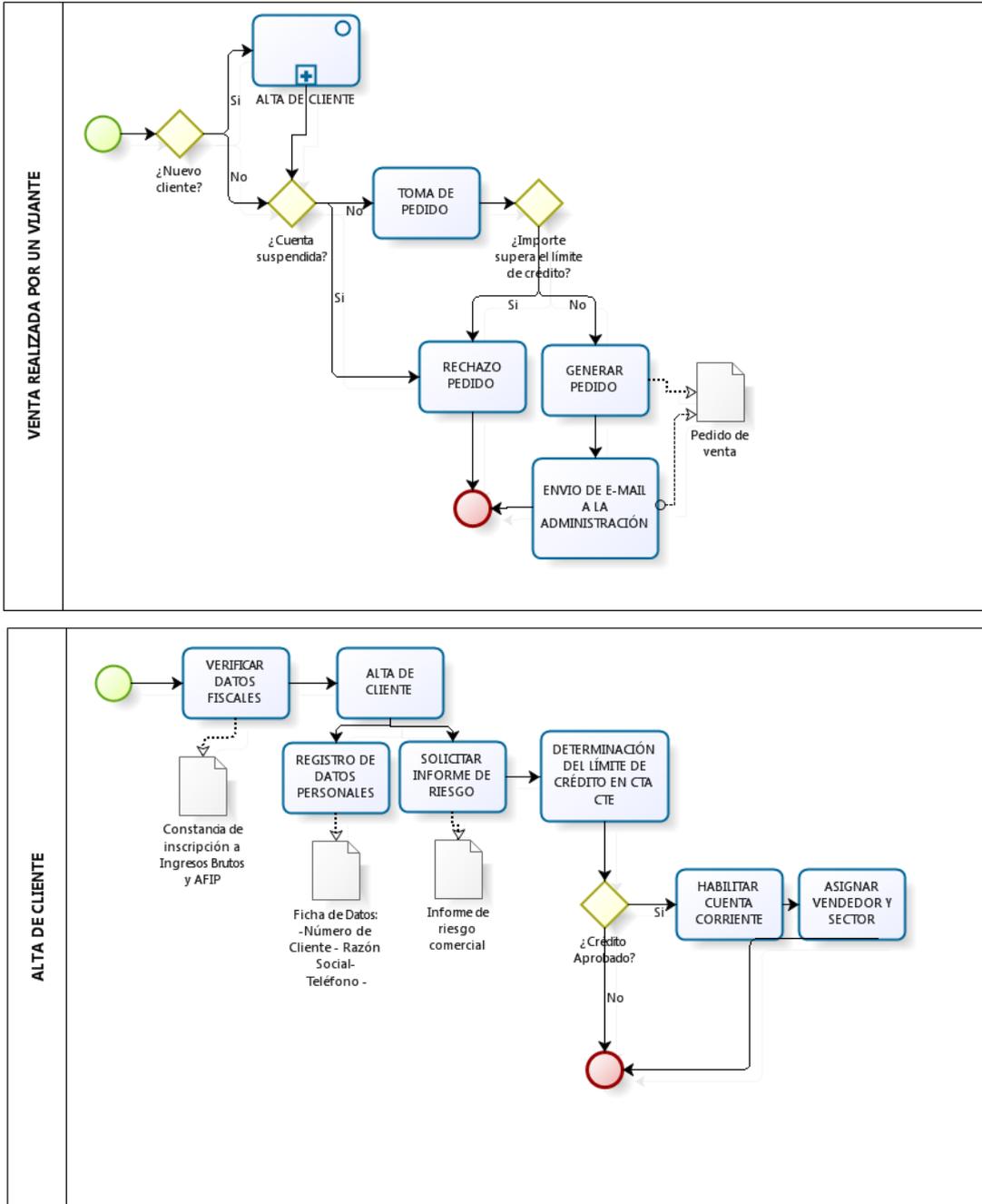
- **ORGANIGRAMA:**



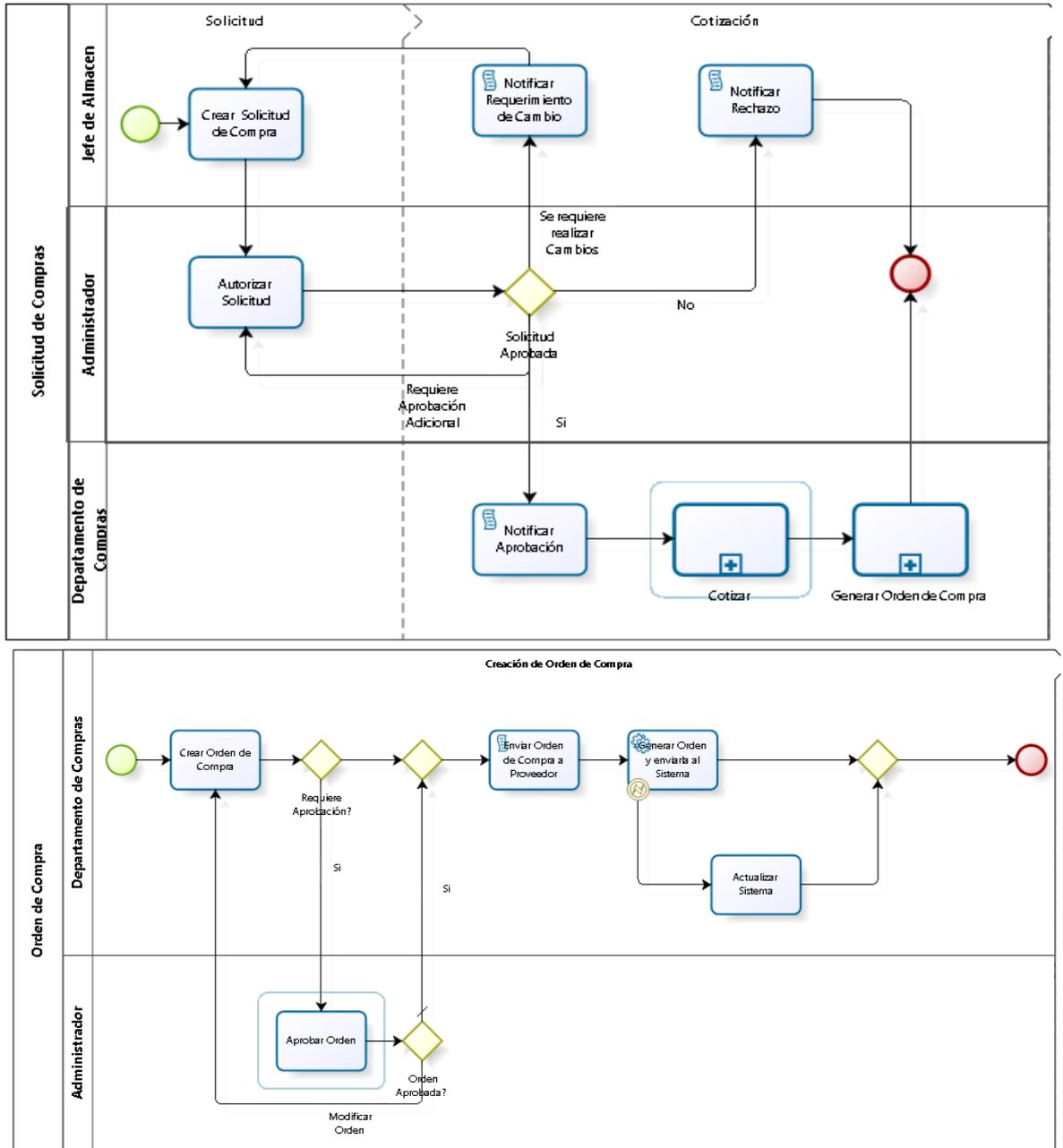
• **Proceso a escoger y Modelar el Proceso ( Bizagi)**

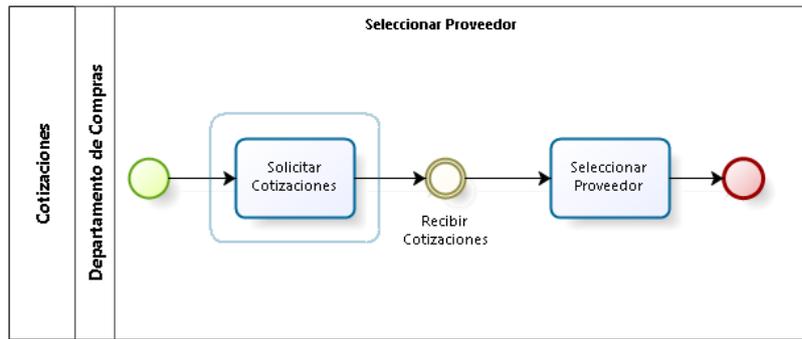
Hemos considerado:

a) **Ventas**



**b) Compras**





### 2.1.2. Selección de la Estrategia de Implementación

La estrategia de implementación del proyecto contiene las siguientes tareas:

1. Conocer la realidad de la empresa y familiarizarnos con la cultura organizacional.
2. Analizar la documentación, manuales y procedimientos de la empresa
3. Analizar los sistemas transaccionales.
4. Recolectar los requerimientos empresariales.
5. Analizar los requerimientos.
6. Realizar el diseño y la arquitectura para los Data Marts en la nube.
7. Implementar los Data Marts.

### 2.1.2. Selección de la Metodología de Desarrollo

La implementación de un Data Mart es un proceso complejo, es por esto que deben usarse las mejores prácticas existentes. 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.

### 2.1.3. 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 medio o administrativo de la empresa Book Center S.A.C.; y el análisis con el apoyo del equipo desarrollador.

Para ello se formuló las siguientes preguntas:

- a) ¿Cuál es el área más recomendable estratégicamente para aplicar los data marts?
- b) ¿Cuál es la priorización de desarrollo de los Data Marts en los procesos de las principales áreas de la empresa?
- c) ¿Cuál es el rango de consultas empresariales a los que se debe responder inicialmente los Data Marts?

Después de haber realizado el análisis del funcionamiento de la organización, las respuestas obtenidas fueron las siguientes:

- a) Las áreas de Ventas, es la más recomendable, en el cual debe ser aplicada los Data Marts, por ser la área que brindan mayor información acerca de la situación de la empresa.
- b) De acuerdo a las entrevistas, manuales y documentación de los sistemas, la priorización de desarrollo será la siguiente
  - 1° Proceso de Ventas
  - 2° Proceso de Compras
- c) El rango de consultas empresariales que debe responder inicialmente los data marts, es de 10 consultas.

#### 2.1.4. Selección del enfoque arquitectónico

El enfoque arquitectónico para los data marts estarán compuestos de los siguientes niveles

##### ➤ **Back Room**

Está compuesto por:

- 1 Servidor de Base de Datos HP ProLiant XL220a, cuya estructura se encuentra en SQL Server.
- 1 Servidor de Base de Datos SQL Server 2014, que servirá como repositorio de la base de datos de la empresa.
- Conexión a Microsoft Azure.

##### ➤ **Front Room**

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

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

#### 2.1.5.1. Presupuesto

##### a. Recursos Humanos

RECURSOS HUMANOS		COSTO S/.
01	Investigador	2,000.00
01	Asesor	1,500.00
<b>TOTAL</b>		<b>3,500.00</b>

##### b. Bienes: materiales, equipos, software

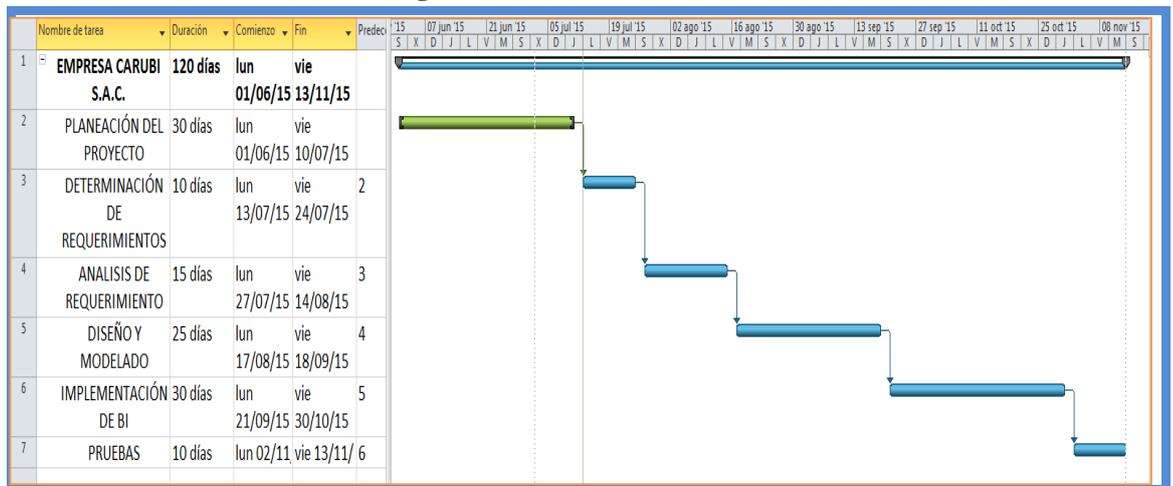
BIENES		COSTO S/.	
MATERIALES	01	Millar de papel bond A-4	44.00
	05	Lapiceros	10.00
	02	Correctores	4.00
	01	Lápiz	2.00
	02	Cartuchos de tinta Negra	110.00
	02	Cartuchos de tinta Color	120.00
	01	Tóner 35 <sup>a</sup>	220.00
	10	Folder manila A-4	10.00
	<b>TOTAL</b>		<b>521.00</b>
EQUIPOS	01	Computadora Intel i3 – 2GB RAM	1,700.00
	01	Impresora HP deskjet D2400	160.00
	01	Impresora HP Laser P1006	480.00
	01	Grabador de DVD	100.00
	01	Memoria USB 8GB	20.00
	<b>TOTAL</b>		<b>2,460.00</b>

SOFTWARE	01	SQL Server 2014	300.00
	01	Licencia Excel (Office 2007)	400.00
	01	Windows Azure	400.00
	<b>TOTAL</b>		<b>1100.00</b>
SERVICIOS		Servicio de transporte	100.00
		Servicio de tipos	140.00
		Servicio de Fotocopiados	10.00
		Servicios de Internet	50.00
	<b>TOTAL</b>		<b>300.00</b>

**c. Presupuesto Total**

PRESUPUESTO TOTAL GENERAL	
CATEGORÍA	COSTO TOTAL S/.
Recursos Humanos	3,500.00
Materiales	521.00
Equipo	2,460.00
Software	1100.00
Servicios	300.00
<b>COSTO TOTAL :</b>	<b>7,881.00</b>

**2.1.5.2. Cronograma de actividades**



**Figura N° 14: Cronograma de actividades**

### 2.1.6. 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**

**Tabla N° 2:** 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	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.
<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>
--	---	--

Tabla N° 3: Descripción de usuarios del Data Mart

Nombre	Representa	Stakeholder
<b>GERENTE GENERAL</b>	Es aquel que se encarga de plantear las estrategias para el buen funcionamiento de la empresa	Acceso total de las dimensiones del Data Mart
<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>

## 2.2. Determinación de los Requerimientos

### 2.2.2. Requerimientos del Propietario.

Responder a las preguntas que proponen los propietarios del Data Mart.

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

Para desarrollar estrategias como lograr una mayor participación en el mercado y administrar riesgos, haciendo uso del análisis de sus datos históricos y actuales, beneficiando de esta manera un mejor desempeño de la empresa.

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

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

- **¿Cómo afecta nuestras inversiones actuales en cómputo?**

Al contar con la máquina para los usuarios finales sólo se necesitará hacer una inversión en la instalación del Data Mart que es de \$3000.00.

- **¿Cuáles son los riesgos?**

Los riesgos al implementarse el Data Mart serian que la base de datos operacional tenga complicaciones, es decir, genere problemas de consistencia.

### 2.2.3. Requerimientos del Usuario Final

Para determinar los requerimientos empresariales se realizaron entrevistas a los usuarios finales que utilizarán los data marts basados en la metodología de Ralph Kimball.

#### Área de Ventas

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 las 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 N° 4:** Consultas y medidas

NRO	CONSULTAS	UNIDAD
01	¿Cuál es el monto de un determinado cliente en compras en un mes?	Ventas soles
02	¿Cuál es el Monto de ventas por tipo de producto?	Ventas soles
03	¿Cuál es el monto de un tipo de producto por proveedor en 3 meses?	Ventas soles
04	¿Cuál es el monto de ventas clientes en un semestre?	Ventas soles
05	¿Cuál es el monto que generan las productos en un año?	Ventas soles
06	¿Cuál es la cantidad de productos vendidos por Tipo, vendedor, mes y semestre?	Cantidad productos
07	¿Cuál es la cantidad de Productos que se han pagado al crédito y al contado?	Cantidad productos

08	¿Cuál es el monto que generan las promociones que da un proveedor en un semestre?	Ventas soles
09	Listar a los Vendedores y el monto total de cada una de las ventas que han generado	Ventas soles
10	¿Cuál es el porcentaje de las ventas de los productos según su tipo por mes?	% Ventas

#### 2.2.4. Requerimientos no Funcionales

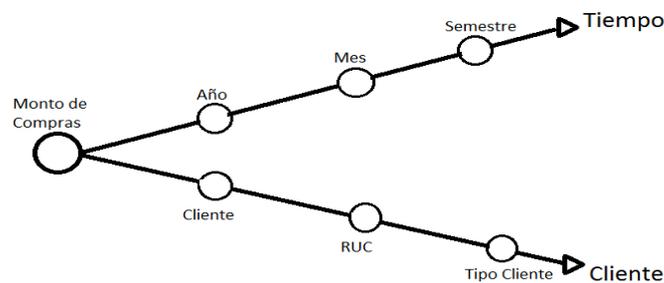
- Brindar mejor información a los usuarios, en el menor tiempo.
- Establecer seguridad adecuada para la administración de los data marts.
- Los Data Marts deben funcionar sobre la plataforma de Microsoft Windows 2012 Server, con el manejador de base de datos Microsoft SQL Server 2012 y utilizando el Microsoft SQL Server 2012 Analysis Services como herramienta de diseño y construcción de los cubos.
- Utilizar MS Excel 2013 para mostrar los reportes de los cubos.

#### 2.2.5. Análisis de los requerimiento

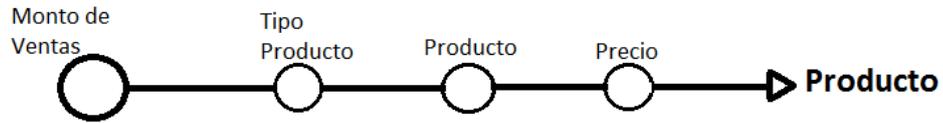
En esta parte se hará un análisis de cada requerimiento del usuario final y definiendo algunas tablas de la Base de Datos Operacional de la empresa, que actuarían como dimensión dentro del Data Mart. También se definirá las posibles medidas dentro del Data Mart.

##### Área de Ventas

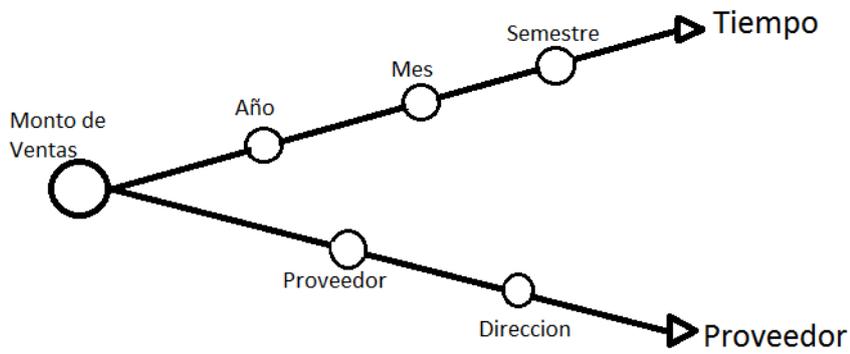
**R1: ¿Cuál es el monto de un determinado cliente en compras en un mes?**



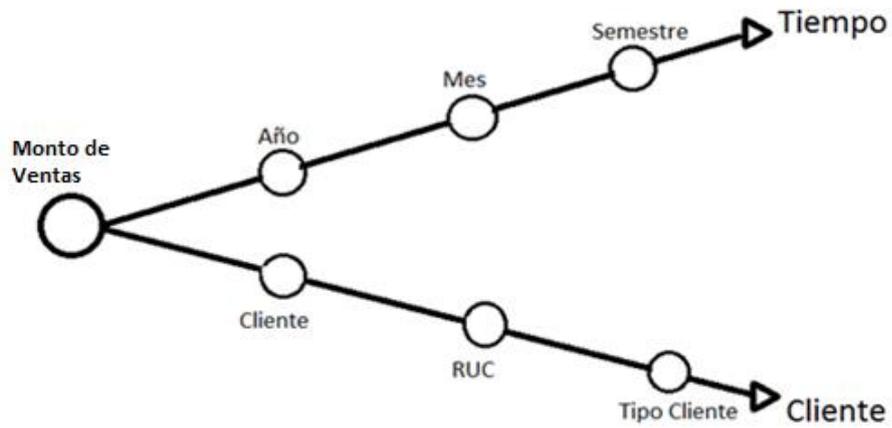
**R2: ¿Cuál es el Monto de ventas por tipo de producto?**



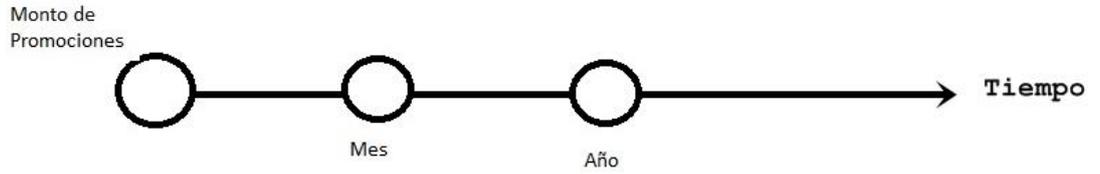
**R3: ¿Cuál es el monto de un producto para un determinado proveedor en 3 meses?**



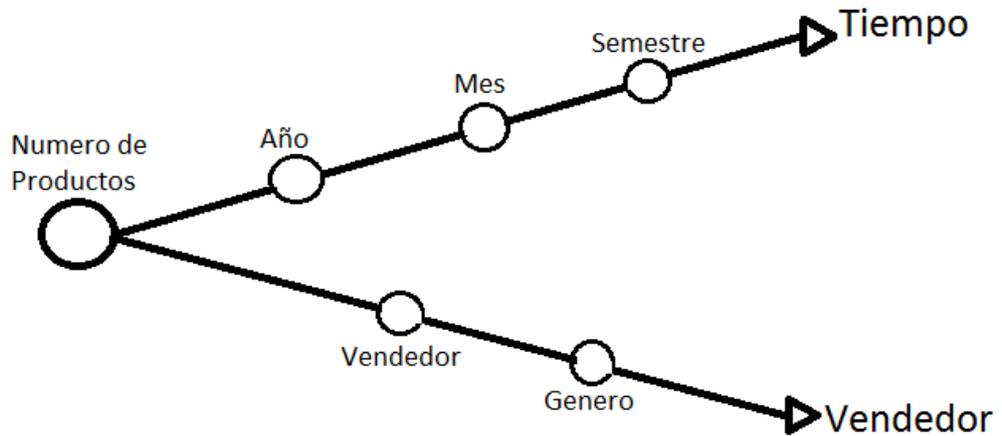
**R4: ¿Cuál es el monto de ventas de los clientes en un semestre?**



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



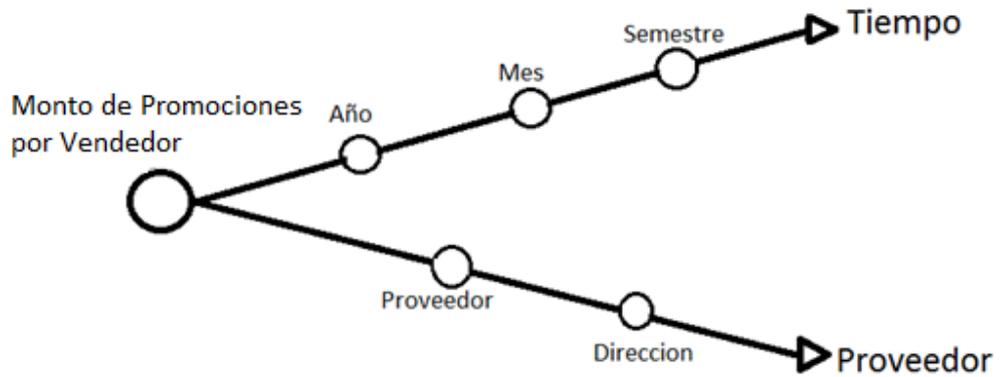
**R6: ¿Cuál es la cantidad de productos vendidos por Tipo, vendedor, mes y año?**



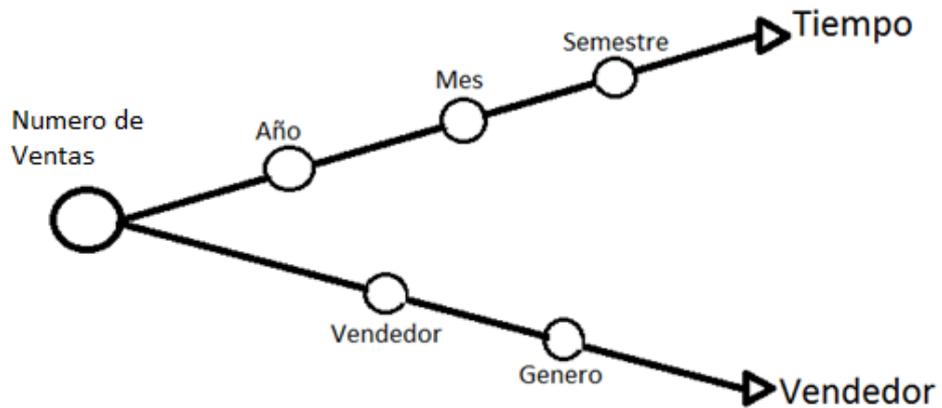
**R7: ¿Cuál es la cantidad de Ventas que se han pagado al crédito y al contado?**



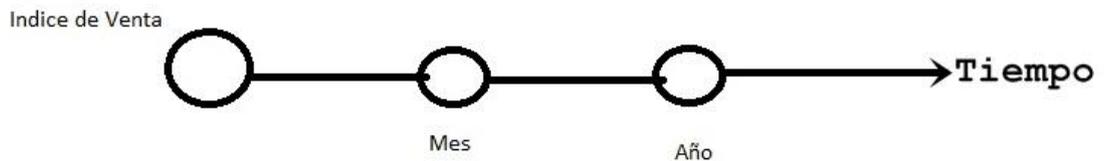
**R8: ¿Cuál es el monto que generan las promociones que da un proveedor en un año?**

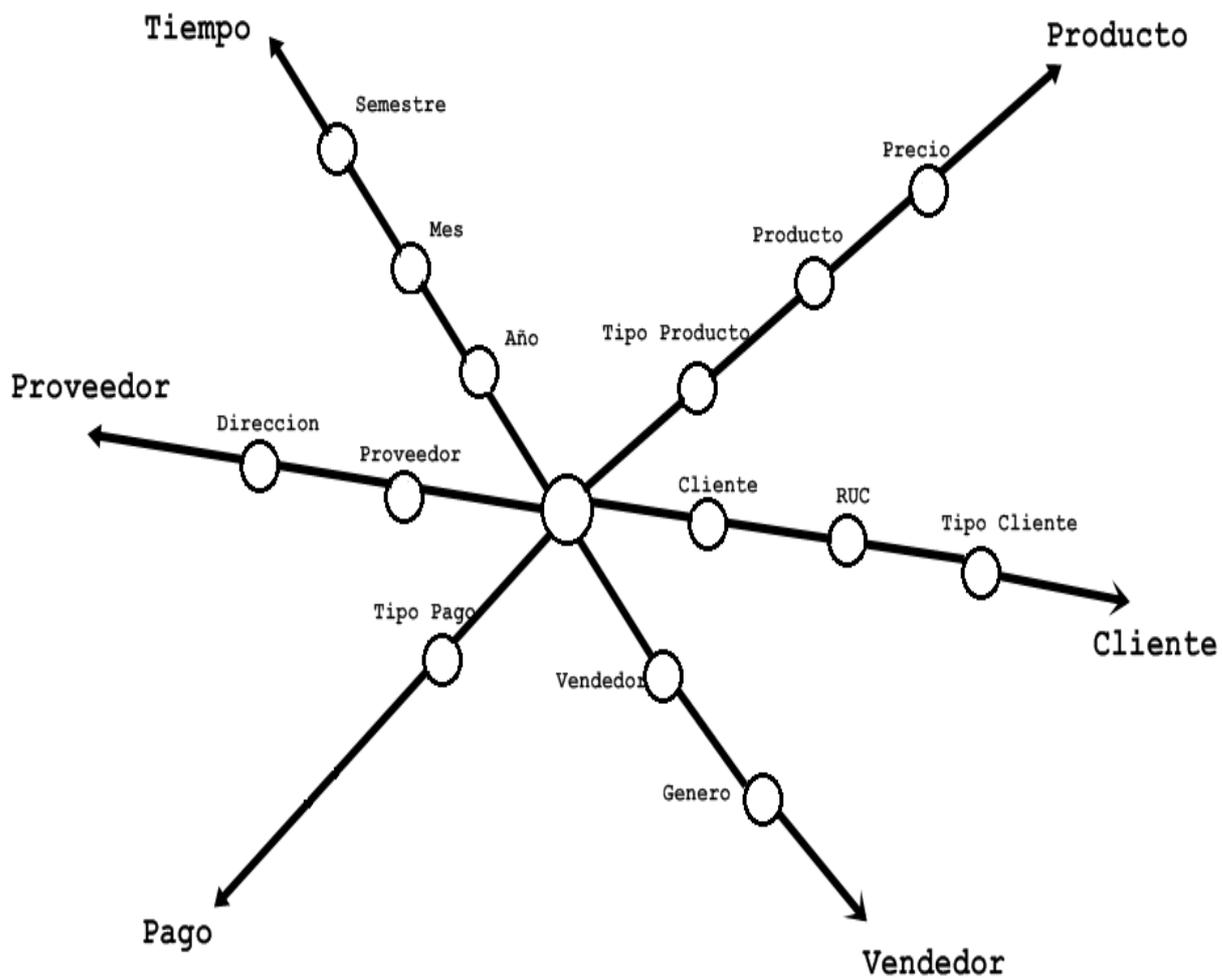


**R9: Listar a los Vendedores y el monto total de cada una de las ventas que han generado**



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



**MODELO STARTNET VENTAS****Figura N° 26:** Modelo Startnet Ventas

## 2.3 Diseño Técnico de la Arquitectura

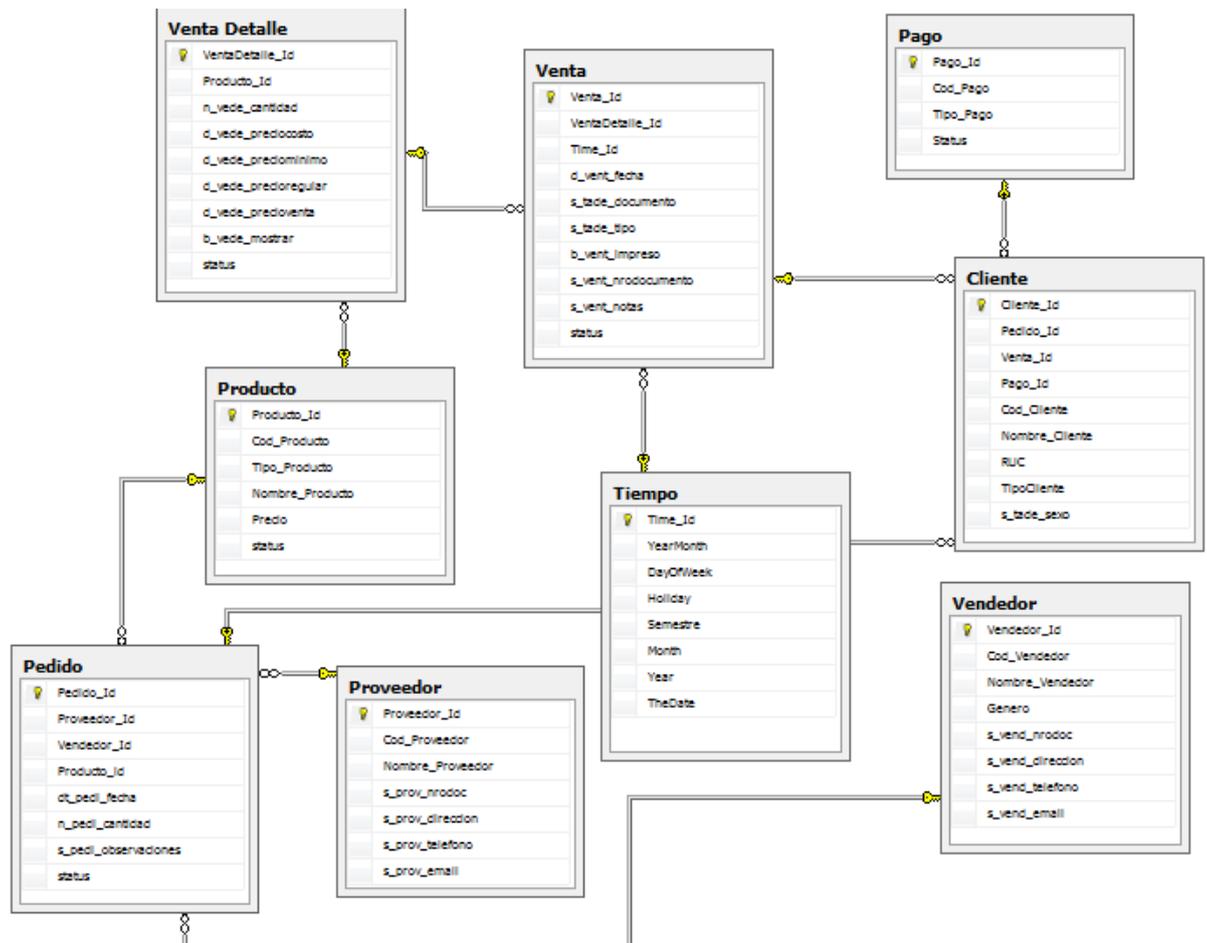
### 2.3.1 Nivel de Datos

Para el análisis de los datos, se comienza por analizar los datos fuentes que manejan los procesos de la empresa, el tipo de la base de datos y la estructura de las tablas.

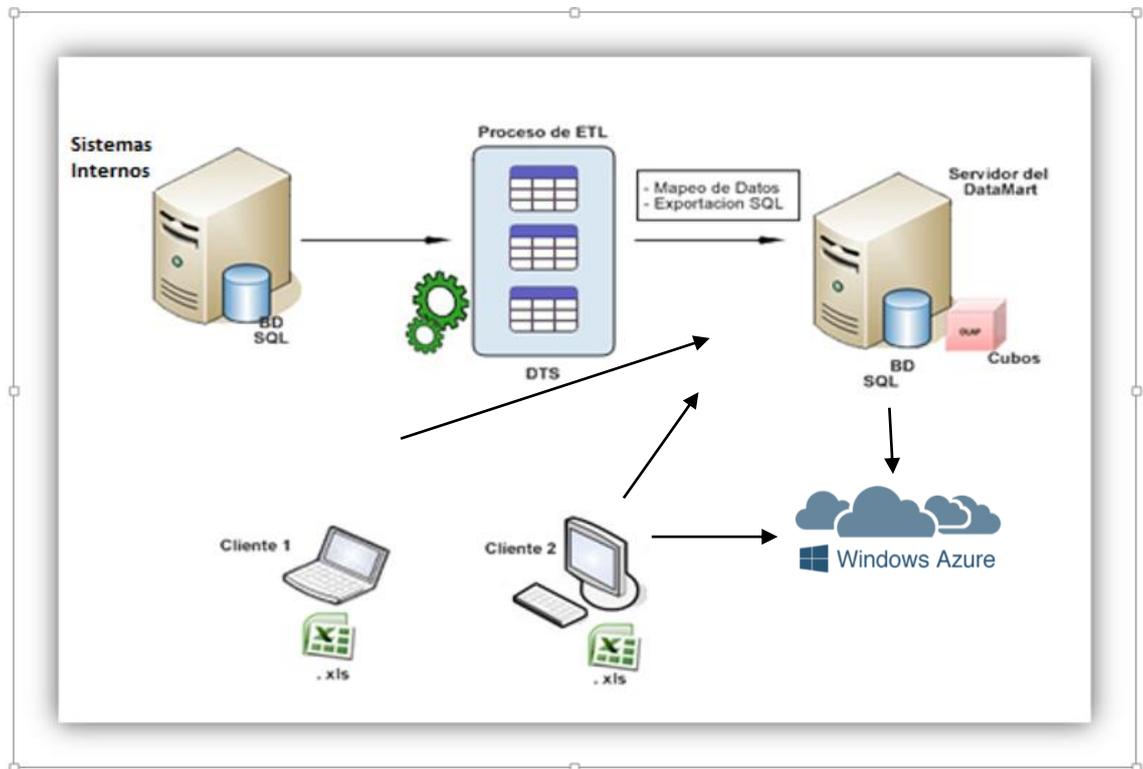
**Base de Datos Fuente:** Se utilizó la base de datos BookCenter cuyo modelo esta descrito en la Figura N° 27.

En este diagrama se muestran las tablas que se relacionan con el área de Ventas, y que son de interés para el Datamart. La mayoría de estas tablas servirán para crear las consultas para el poblamiento de nuestro Datamart.

**Figura N° 27:** Diagrama E-R de la base de datos



### 2.3.2 Nivel Técnico



**Figura N° 28:** Enfoque Arquitectónico del DataMart

El enfoque arquitectónico para los data marts estarán compuestos de los siguientes niveles

➤ **Back Room**

Está compuesto por

- 1 Servidor de Base de Datos HP ProLiant XL220a, cuya estructura se encuentra en SQL Server.
- 1 Servidor de Base de Datos SQL Server 2012, que servirá como repositorio de los datamarts de la empresa.

➤ **Front Room**

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

## 2.4. Modelado Dimensional

Esta parte llamada Modelamiento Dimensional, 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.

### 2.4.1. Identificación de los Componentes del Modelo

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:

#### **1. Elección del Data Mart**

En este paso identificaremos los posibles Data Mart que se pueden construir en la empresa Book Center SAC. En nuestro caso identificaremos el área que está en estudio por este proyecto, el cual ya tenemos conocimiento sobre su modelo general de datos.

Entonces, para identificar nuestro Data Mart en estudio recurrimos al siguiente método:

#### **Método de la Matriz**

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

**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

**Listado de las Dimensiones**

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

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

**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.

**Tabla N° 5:** Matriz de Intersecciones

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

Para el fácil manejo de las consultas hechas por el usuario, en el Data Mart se consideró la Tabla de Hecho: Ventas.

## 2. Elección del Objetivo de las Tablas 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 figurará en el diseño del modelado del Data Mart.

Tomamos la siguiente definición para la tabla de hechos:

**Tabla N° 6:** Objetivos de las tablas de Hecho

Tabla de Hechos	Objetivo
Ventas	“Tener un mejor control y gestión de las ventas”.

Al enunciar el objetivo de la Data Mart, van a intervenir las posibles dimensiones que se explicarán a continuación.

## 3. Elección de las 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 sí mismo establece cuales serán nuestras dimensiones.

A continuación por cada objetivo escogeremos las dimensiones:

**Tabla N° 7:** Dimensiones de la tabla de Hechos Ventas

Tabla de Hechos	Objetivos	Dimensiones
Ventas	“Tener un mejor control y gestión de las ventas”	Pago Tiempo Vendedor Proveedor Cliente Producto

Como se puede ver la mayoría de las dimensiones elegidas pertenecen a las Entidades del Modelo de Datos General descritos anteriormente. Otras dimensiones son la unión de entidades.

#### 4. Elección de los Hechos

El Objetivo de cada tabla de hechos, permite definir los hechos y hace claro el alcance que estos hechos deben tener.

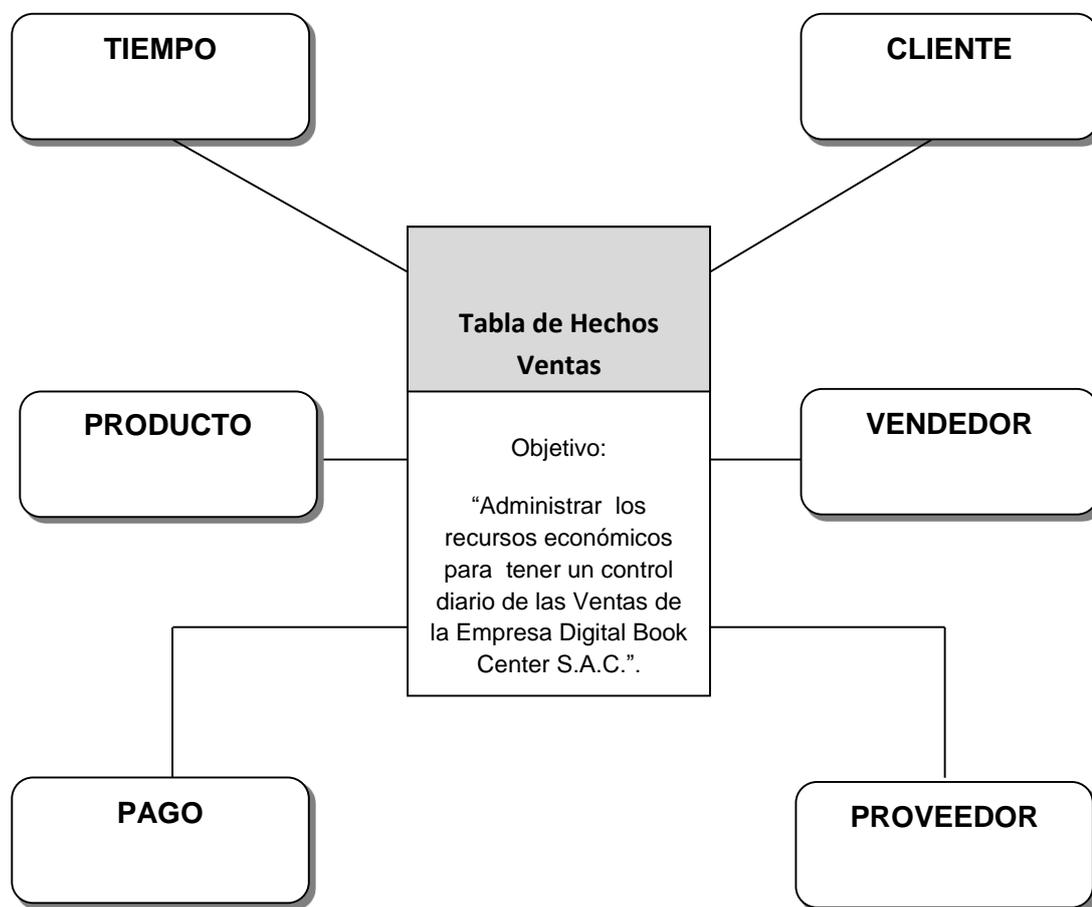
En nuestro caso tenemos:

**Tabla N° 9:** Elección de Hechos

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

#### 2.4.2. DIAGRAMA DE LA TABLA DE HECHOS

Para esta etapa se prepara un diagrama lógico para cada Tabla de Hechos completada. Cada Diagrama nombra a la Tabla de Hechos, establece su Objetivo y muestra todas las dimensiones conectadas a la Tabla de Hechos.

**Diagrama de la Tabla de Hechos Ventas****Figura N° 29:** Diagrama de Tabla de Hechos Ventas

- **Detalle de las Tablas de Hechos**

**Tabla N° 10:** Detalle de las claves de las dimensiones

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

**Tabla N° 11:** Detalle de las Medidas de las Tablas de Hecho

Nombre de la Tabla	Nombre de la Columna	Descripción de la Columna
Hecho 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

Por consiguiente, de las dos tablas completadas anteriormente se obtienen las siguientes Tablas de Hechos:

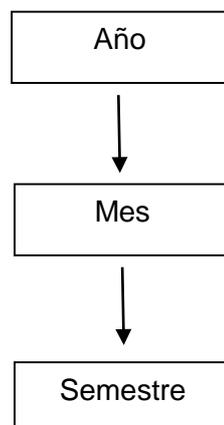
**Figura N° 35:** Tabla de Hechos Venats

TABLA DE HECHOS VENTAS
Tiempo_Id
Producto_Id
Vendedor_Id
Cliente_Id
Monto en Soles
Cantidad de Proveedores
Numero de Vendedores
Cantidad de Ventas
Cantidad de Productos
% Descuento

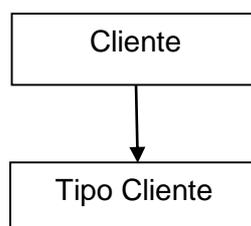
**Figura N° 31:** Tabla de Hechos del Data Mart

- **Detalle de las Tablas Dimensión**

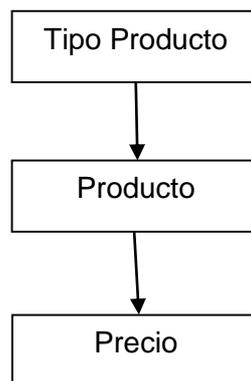
De acuerdo a nuestras Dimensiones elegidas anteriormente construimos los detalles y sus jerarquías de cada una de ellas.

**Dimensión Tiempo:**

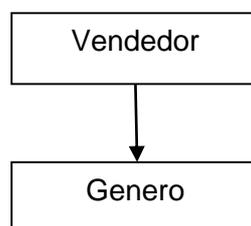
Nombre del Atributo	Descripción del Atributo	Valores de Muestra
<b>Año</b>	Representa el año.	2012, 2013,...
<b>Mes</b>	Representa los meses que posee un año.	Abril, Mayo,...
<b>Semestre</b>	Representa el semestre	Enero-Marzo, Marzo-Junio

**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

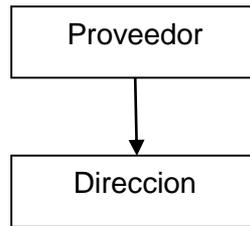
**Dimensión Producto:**

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

#### 2.4.2.1. 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.

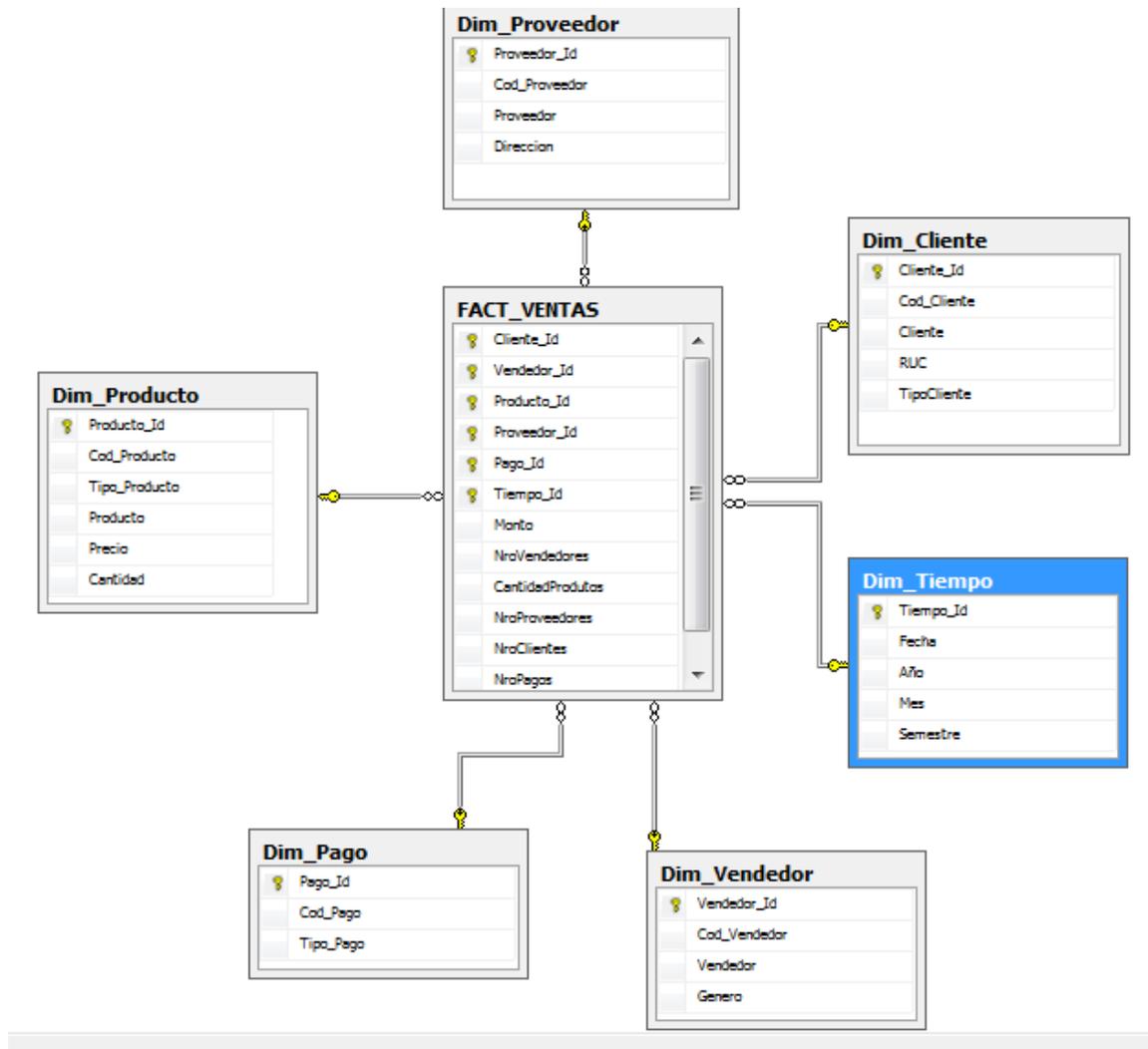
Para una mejor comprensión, listamos primero los componentes que intervendrán el esquema, luego dividimos el Data Mart en dos gráficos: El esquema estrella de cada Tabla de Hechos y el Diseño Lógico del Data Mart.

Componente: TABLA DE HECHOS VENTAS

Componente: TABLA DE DIMENSIÓN TIEMPO

Componente: TABLA DE DIMENSIÓN VENDEDOR

- Componente: TABLA DE DIMENSIÓN CLIENTE  
 Componente: TABLA DE DIMENSIÓN PROVEEDOR  
 Componente: TABLA DE DIMENSIÓN PRODUCTO  
 Componente: TABLA DE DIMENSIÓN PAGO



**Figura N° 37:** Esquema Estrella del Data Mart

## 2.5. 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:

**Tabla N°18:** Nombres estándares para las Tablas Hechos y Dimensiones

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° 19:** Nombres estándares para los atributos de las 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° 20:** Nombres estándares para los atributos de las Tablas de Hechos

NOMBRE DE LA TABLA	DISEÑO LÓGICO	DISEÑO FÍSICO
Tabla de Hecho 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

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		DESCRIPCION
	DATOS	NULOS	LONGITUD	
<b>Cliente_Id</b>	int	no	12	Dimensión que contiene toda la información detallada de los Clientes de Book Center S.A.C
<b>Cod_Cliente</b>	nchar	no	12	
<b>Cliente</b>	varchar	no	50	
<b>RUC</b>	char	no	11	
<b>TipoCliente</b>	varchar	no	50	

**2. Dim\_Vendedor**

ATRIBUTOS	TIPOS DE	VALORES		DESCRIPCION
	DATOS	NULOS	LONGITUD	
<b>Vendedor_Id</b>	int	no	12	Dimensión que contiene toda la información detallada de los Vendedores de Book Center S.A.C
<b>Cod_Vendedor</b>	char	no	10	
<b>Vendedor</b>	varchar	no	50	
<b>Genero</b>	char	no	10	

**3. Dim\_Producto**

ATRIBUTOS	TIPOS DE	VALORES		DESCRIPCION
	DATOS	NULOS	LONGITUD	
<b>Producto_Id</b>	int	no	12	

<b>Cod_Producto</b>	nchar	no	10	Dimensión que contiene toda la información detallada de los Productos de Book Center S.A.C
<b>Tipo_Producto</b>	nchar	no	10	
<b>Producto</b>	varchar	no	50	
<b>Precio</b>	nchar	no	10	
<b>Cantidad</b>	int	no	4	

#### 4. Dim\_Proveedor

ATRIBUTOS	TIPOS DE		VALORES		DESCRIPCION
	DATOS		NULOS	LONGITUD	
<b>Proveedor_Id</b>	int		no	12	Dimensión que contiene toda la información detallada de los Proveedores de Book Center S.A.C
<b>Cod_Proveedor</b>	char		no	10	
<b>Proveedor</b>	varchar		no	50	
<b>Direccion</b>	varchar		no	50	

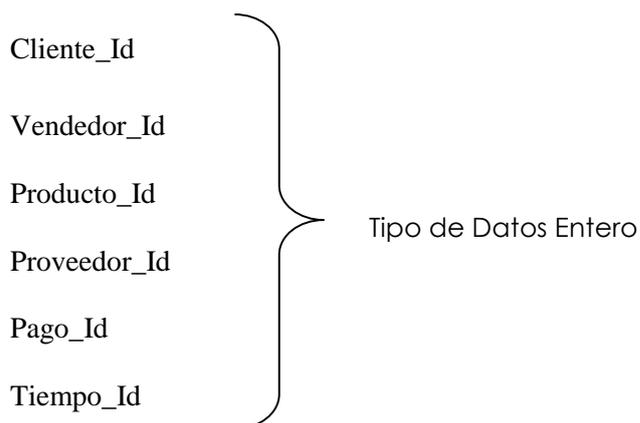
#### 5. Dim\_Pago

ATRIBUTOS	TIPOS DE		VALORES		DESCRIPCION
	DATOS		NULOS	LONGITUD	
<b>Pago_Id</b>	int		no	12	Dimensión que contiene la información de los Pagos de Book Center S.A.C.
<b>Cod_Pago</b>	char		no	10	
<b>Tipo_Pago</b>	char		no	10	

## 6. Dim\_Tiempo

ATRIBUTOS	TIPOS DE		VALORES		DESCRIPCION
	DATOS		NULOS	LONGITUD	
<b>Tiempo_Id</b>	int		no	12	Dimensión que contiene detalle del tiempo de las ventas en Book Center S.A.C.
<b>Fecha</b>	date		no	10	
<b>Año</b>	int		no	4	
<b>Mes</b>	varchar		no	50	
<b>Semestre</b>	varchar		no	50	

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



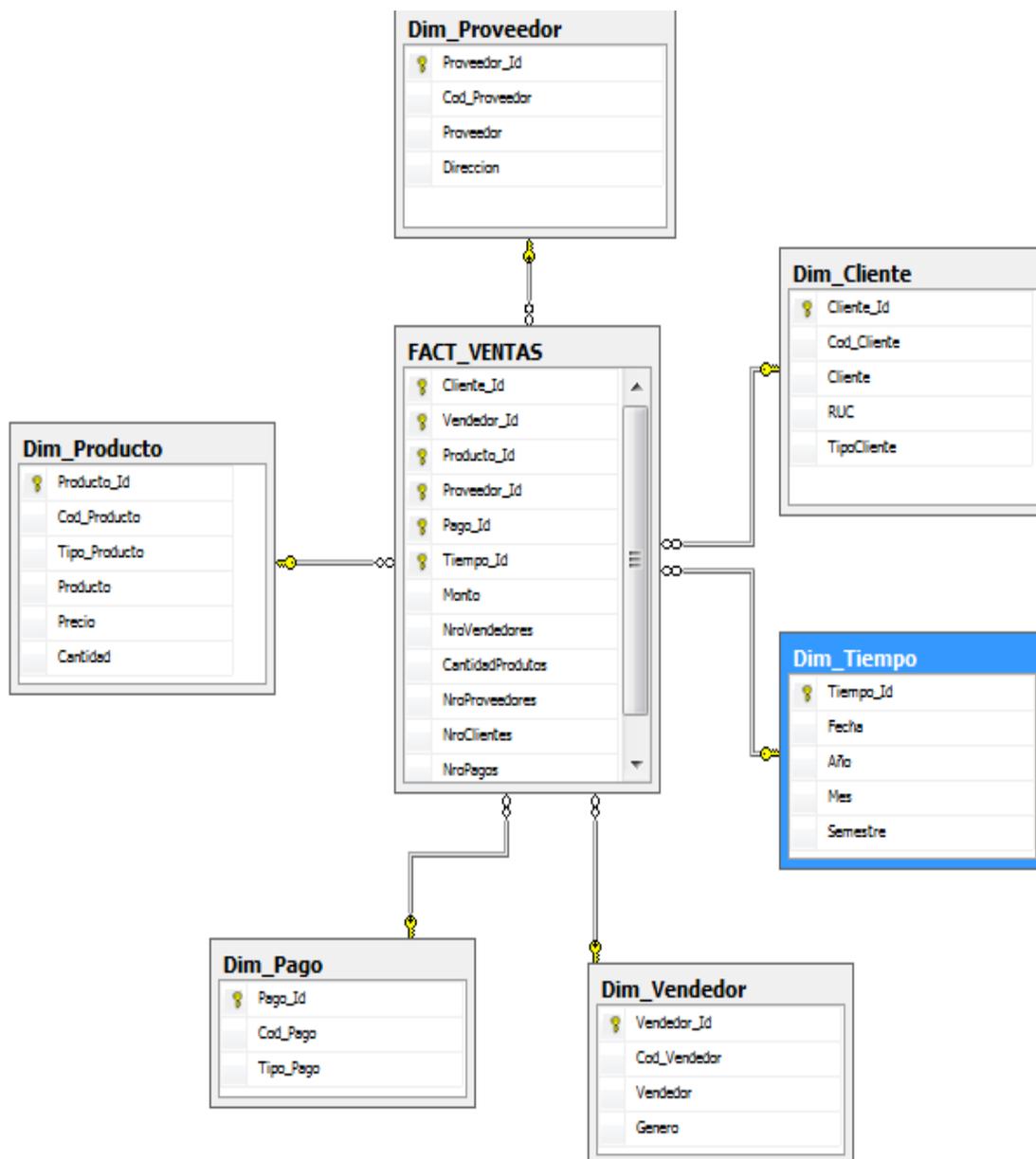
**Tabla:** Claves foraneas de la tabla de Hechos

**Hechos Ventas**

<b>FACT_VENTAS</b>
<b>Cliente_Id (FK)</b>
<b>Vendedor_Id (FK)</b>
<b>Producto_Id (FK)</b>
<b>Proveedor_Id (FK)</b>
<b>Pago_Id (FK)</b>
<b>Tiempo_Id (FK)</b>

**Figura N° 38:** Claves Foráneas de las Tablas de Hechos

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 Data Mart.



*Figura N° 39: Diseño Físico de la Base de Datos del Data Mart*

### 2.5.1. Determinación de las agregaciones

Determinamos las agregaciones por defecto que tendrán cada Hecho o medidas en la Tabla de Hechos. La mayoría de las reglas de agregación son sumas, como veremos a continuación:

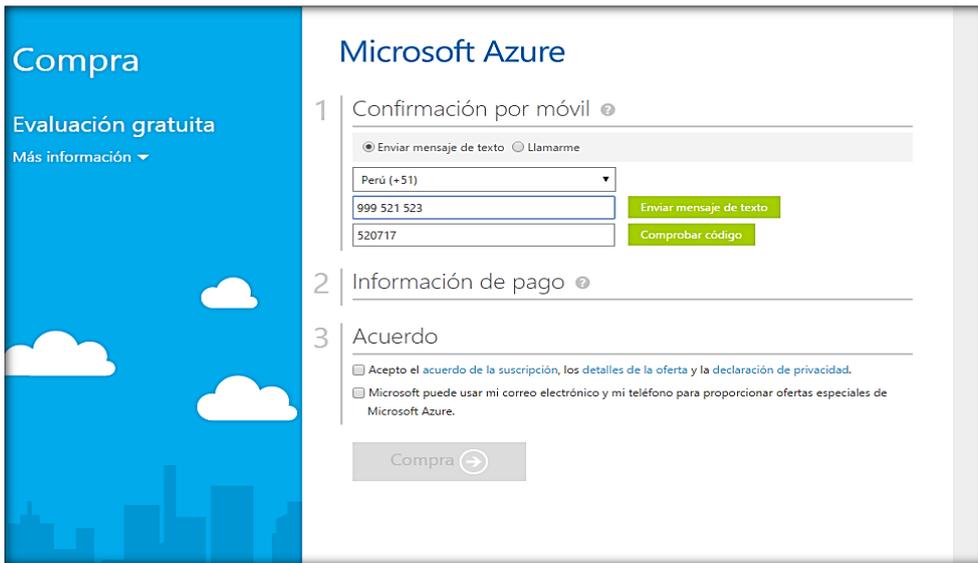
TABLA DE HECHOS	MEDIDAS	REGLA DE LA AGREGACION	FORMULA MDX
FACT_VENTAS	Monto	La suma del precio por la cantidad de productos	SUM(PRECIO*CANTIDAD)
	Cantidad de proveedores	La de los proveedores que colaboran en la empresa	SUM(CANTIDAD)
	Cantidad de vendedores	La de los vendedores de laboran en la empresa	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)

**Tabla N° 27:** Determinación de las Agregaciones

## 2.5.2. Construcción de las Tablas y la Base de Datos en Microsoft Azure

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

### 1° CREAMOS UNA CUENTA EN MICROSOFT AZURE



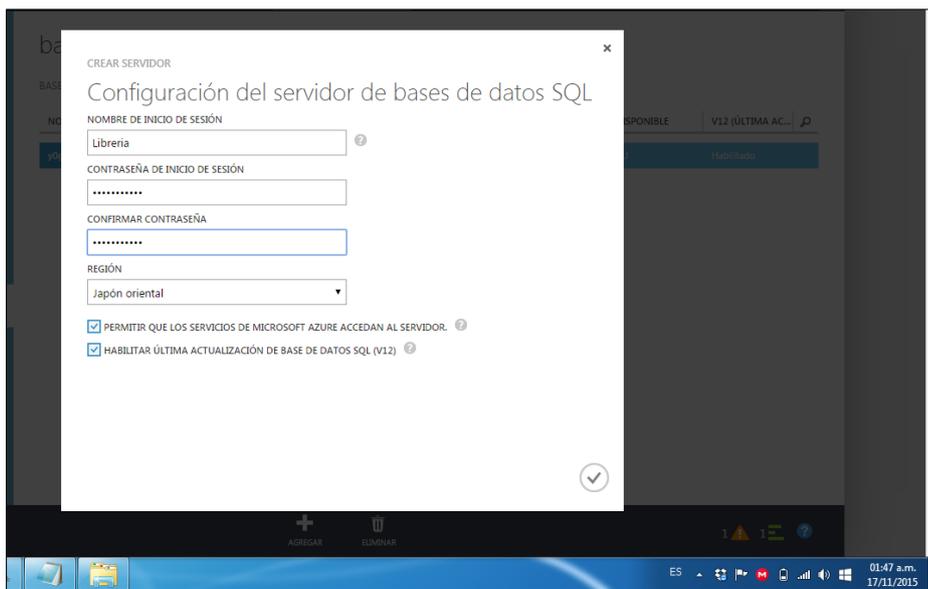
**Compra**

Evaluación gratuita  
Más información ▾

### Microsoft Azure

- 1 Confirmación por móvil ⓘ  
 Enviar mensaje de texto  Llámarme  
Perú (+51) ▾  
999 521 523    
520717
- 2 Información de pago ⓘ
- 3 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.

### 2° CREAMOS LA BASE DE DATOS “BookCenter DW”



CREAR SERVIDOR

### Configuración del servidor de bases de datos SQL

NOMBRE DE INICIO DE SESIÓN  
Libreria ⓘ

CONTRASEÑA DE INICIO DE SESIÓN  
.....

CONFIRMAR CONTRASEÑA  
.....

REGIÓN  
Japón oriental ▾

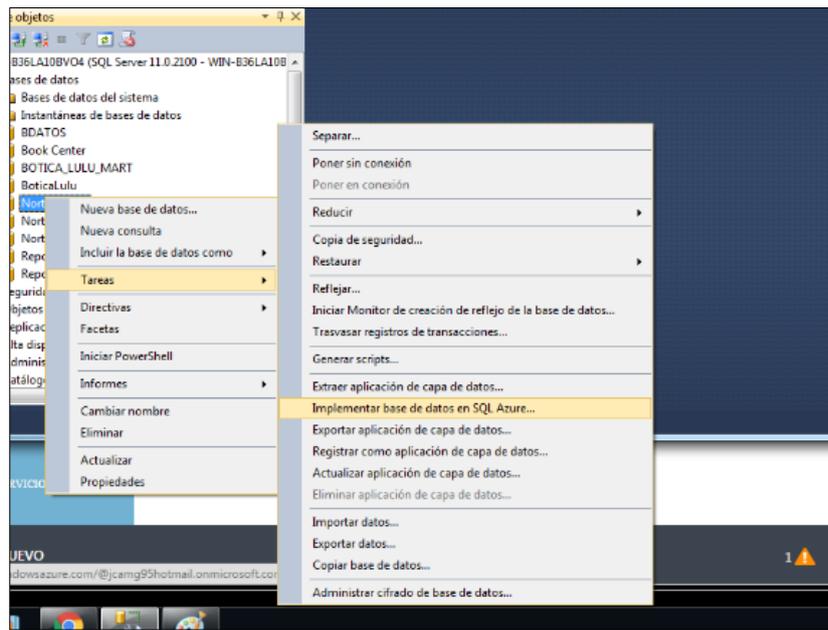
PERMITIR QUE LOS SERVICIOS DE MICROSOFT AZURE ACCEDAN AL SERVIDOR. ⓘ

HABILITAR ÚLTIMA ACTUALIZACIÓN DE BASE DE DATOS SQL (V12) ⓘ

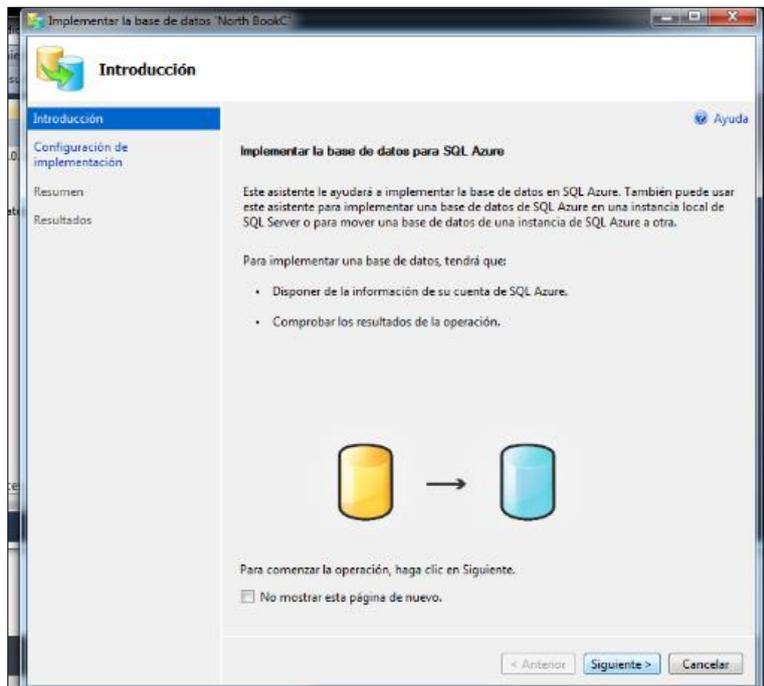
✓

AGREGAR ELIMINAR

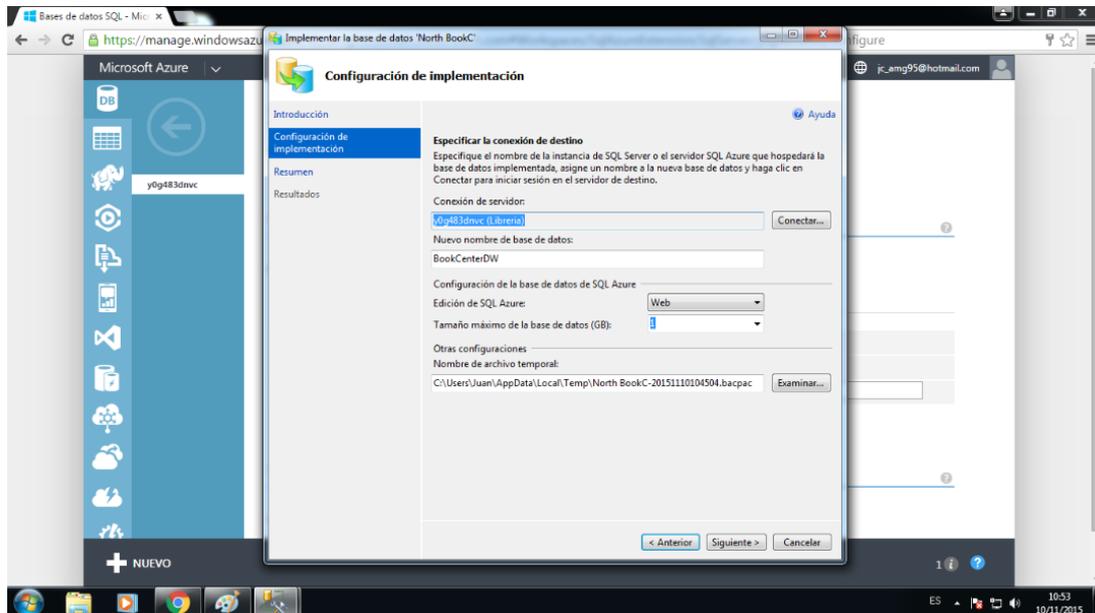
ES 01:47 a.m. 17/11/2015



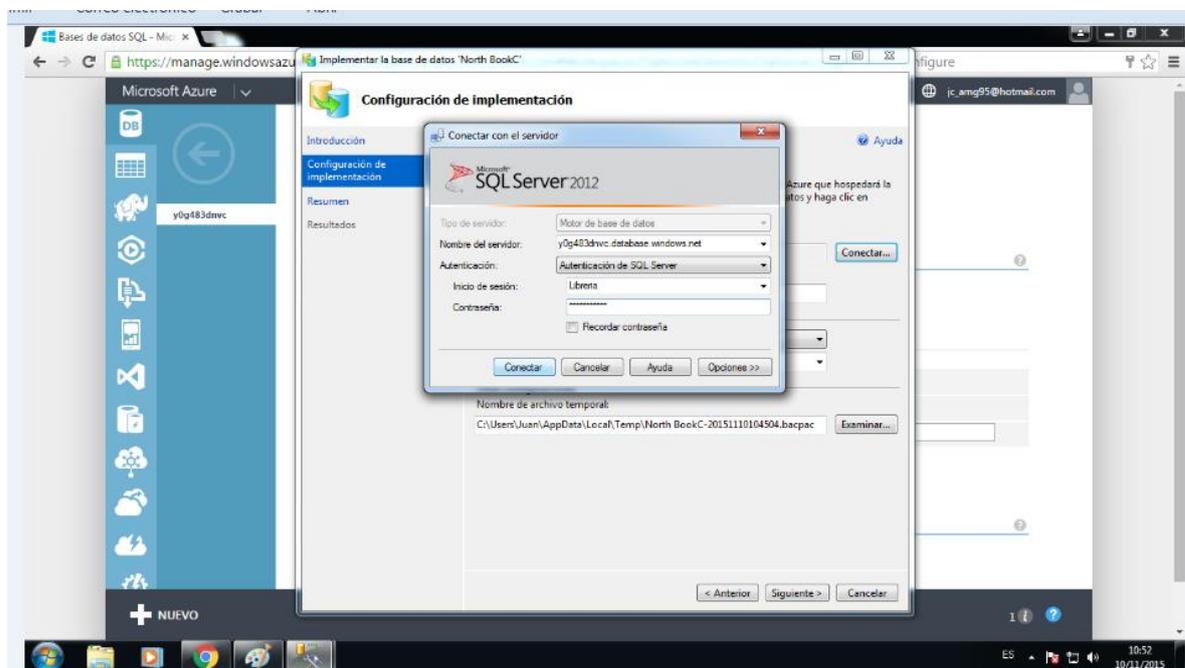
Aparece la Interfaz de Inicio de Implementación de Azure



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



### 3° SE ESTABLECE CONEXIÓN CON MICROSOFT AZURE DESDE SQL



### 4° SE CREAN LAS TABLAS

Tabla 1:Dim\_Tiempo\_Ventas

```
CREATE TABLE Dim_Tiempo_Ventas
(
    Tiempo_Key int not null Primary key,
    Fecha date not null,
    Mes int not null,
    Año int not null
)
```

**Figura N° 40:** Tabla de la Dimensión Tiempo Ventas

### Tabla 3:Dim\_Proveedor

```
CREATE TABLE Dim_Proveedor
(
    Proveedor_Key int not null Primary key,
    ProveedorID nchar(10) not null,
    Nombre_Proveedor nvarchar(60) not null
)
GO
```

**Figura N° 42:** Tabla de la Dimensión Proveedor

### Tabla 4:Dim\_Cliente

```
CREATE TABLE Dim_Cliente
(
    Cliente_Key int not null Primary key,
    ClienteID nchar(10) not null,
    Tipo_Cliente nchar(2) not null,
    Nombre_Cliente nvarchar(60) not null
)
```

**Figura N° 43:** Tabla de la Dimensión Cliente

### Tabla 5:Dim\_Vendedor

```
CREATE TABLE Dim_Vendedor
(
    Vendedor_Key int not null Primary key,
    Nombre_Vendedor nvarchar(60) not null
)
GO
```

**Figura N° 44:** Tabla de la Dimensión Vendedor

#### Tabla 6:Dim\_Pago

```
CREATE TABLE Dim_Transferencia
(
    Transferencia_Key int not null Primary key,
    TransferenciaID nchar(10) not null,
    Motivo nchar(80) not null
)
GO
```

**Figura N° 45:** Tabla de la Dimensión Pago

#### Tabla 7:Dim\_Producto

```
GO
CREATE TABLE Dim_Producto
(
    Producto_Key int not null Primary key,
    ProductoID nchar(10) not null,
    Categoria nchar(60) not null,
    Tipo_Producto nchar(60) not null,
    Nombre_Producto nchar(70) not null,
)
GO
```

**Figura N° 46:** Tabla de la Dimensión Producto

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

## 2.6. 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 (Flujos de trabajo).
- 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.

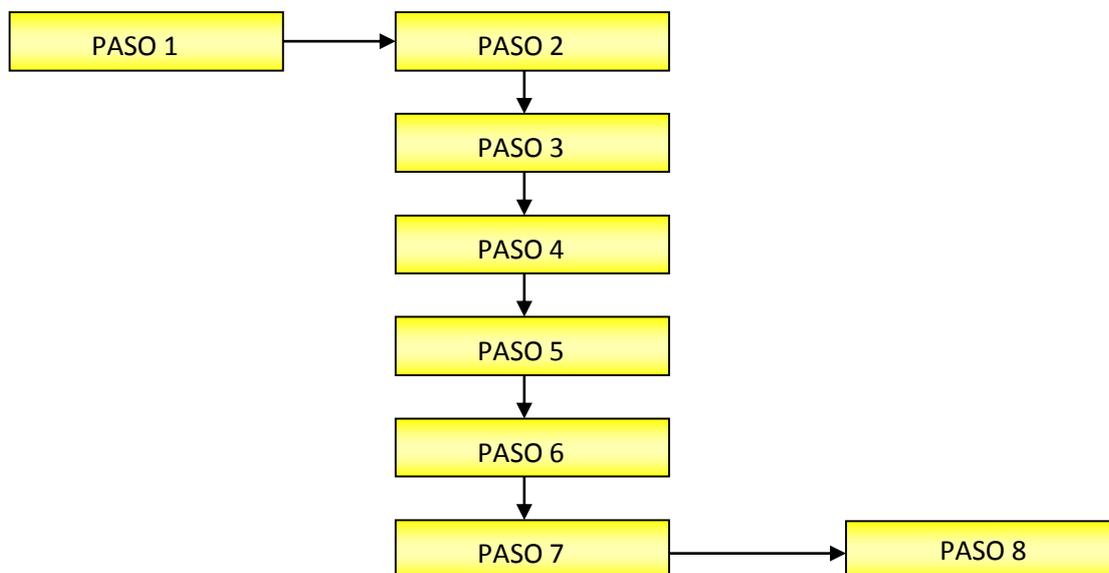
**Paso 4:** Poblamiento de la Dimensión Tiempo, consiste en ejecutar una sentencia SQL, para transferir datos a la Dimensión Tiempo

**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**

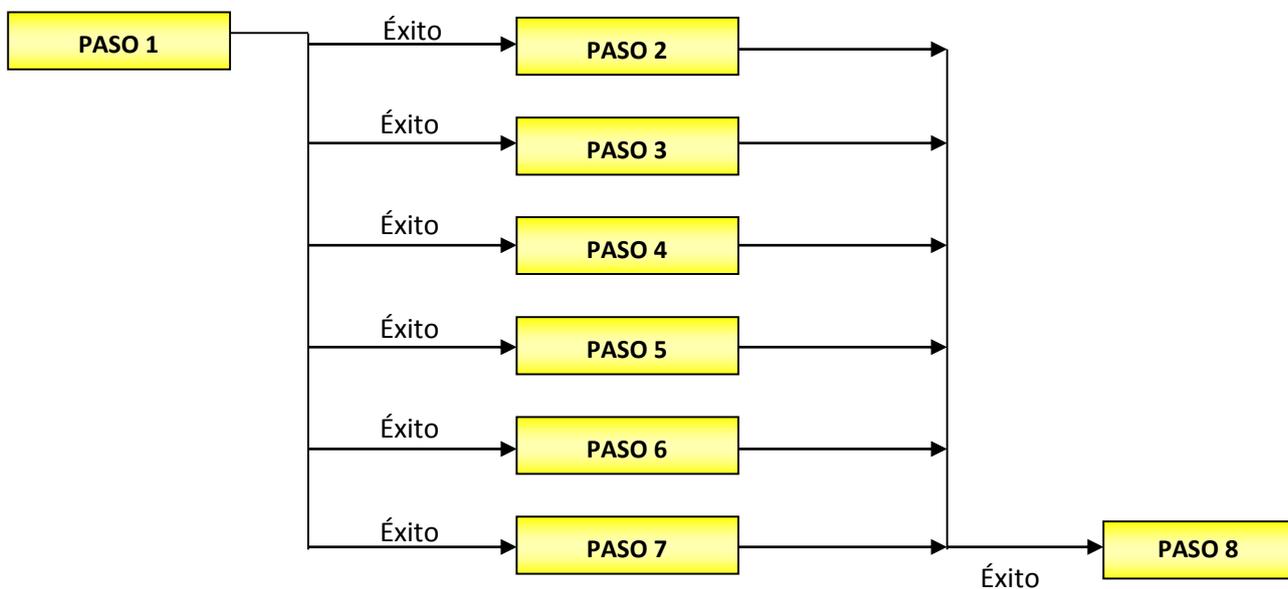
**Figura N° 48:** Workflow de los Pasos de Transformación

**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 Book Center S.A.C.



**Figura Nº 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 2012, 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 General Data Mart Book Center”, con los pasos previos definidos así como todas sus restricciones.

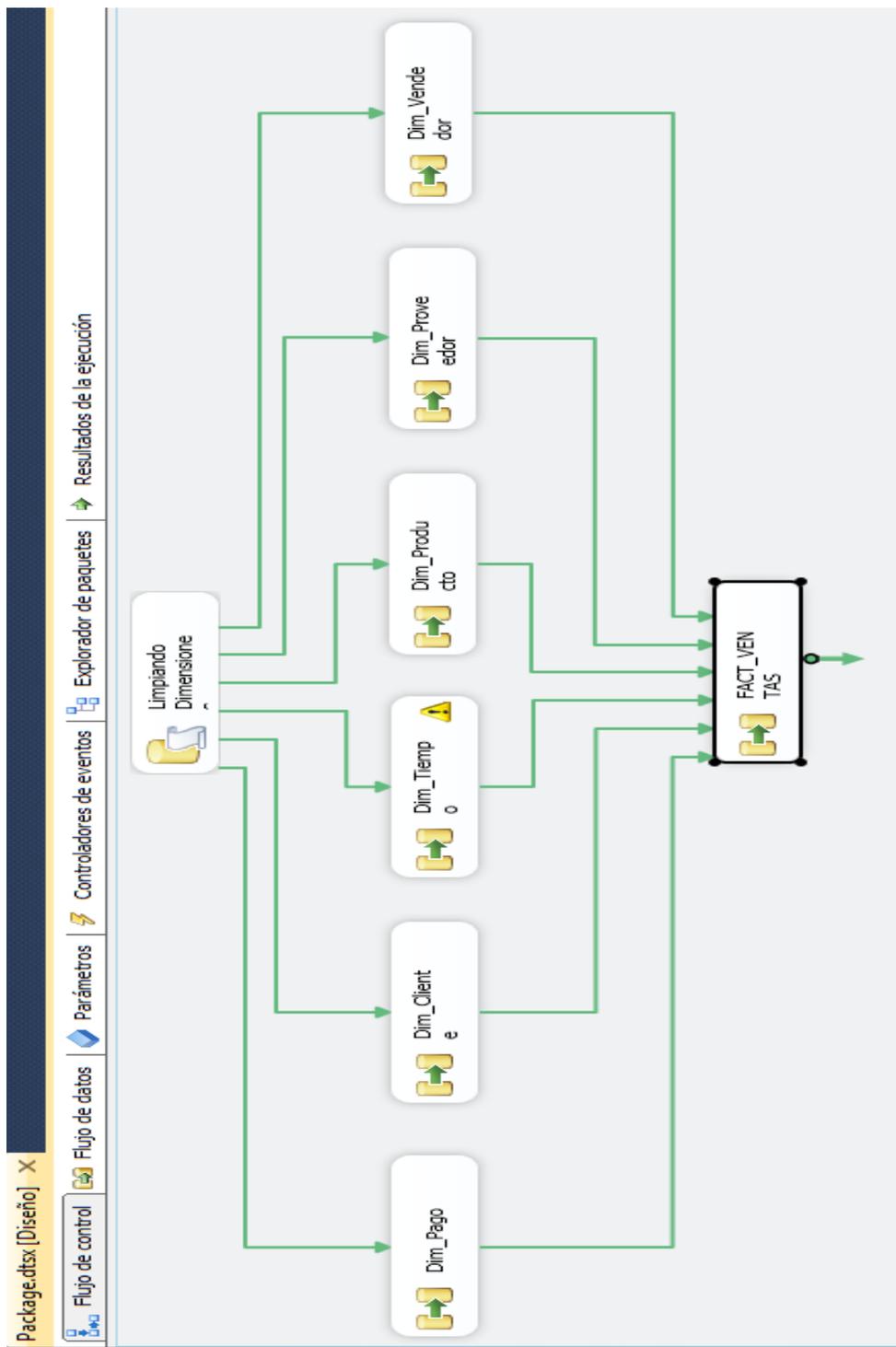


Figura N° 50: ETL del DataMart

**SENTENCIAS DE CADA PASO DE WORKFLOW****Paso 1: Limpiando Dimensiones**

Añadimos una tarea al paquete al cual tendrá por nombre “Limpieza Total”, con la siguiente sentencia:

```
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)

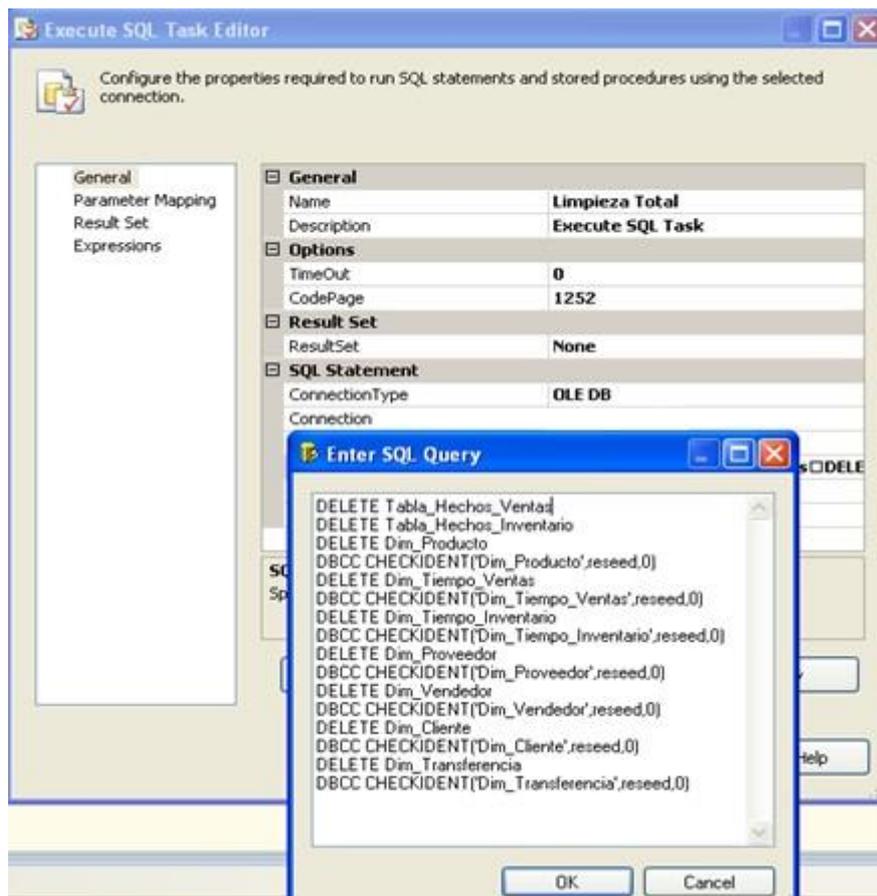
DELETE FROM Dim_Vendedor

DBCC CHECKIDENT('Dim_Vendedor',reseed,0)

DELETE FROM Dim_Proveedor

DBCC CHECKIDENT('Dim_Proveedor',reseed,0)
```

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



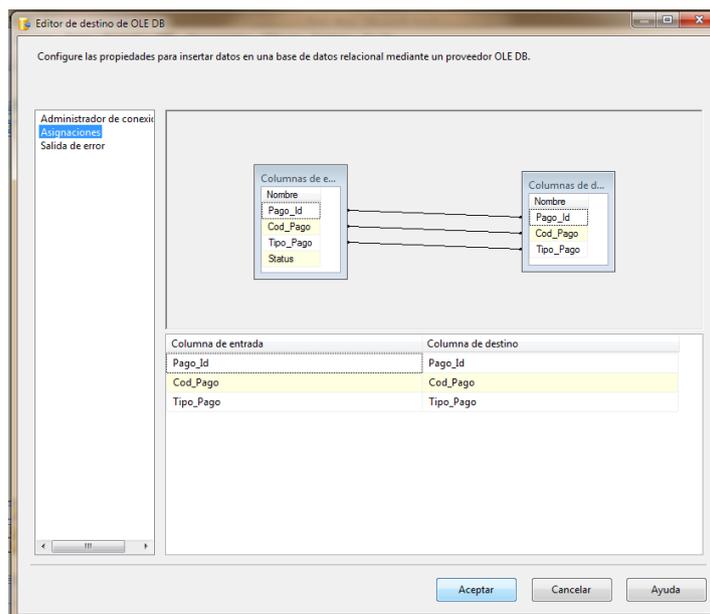
**Figura N° 51:** Ventana de Código para el Paso Limpieza Total

## **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.



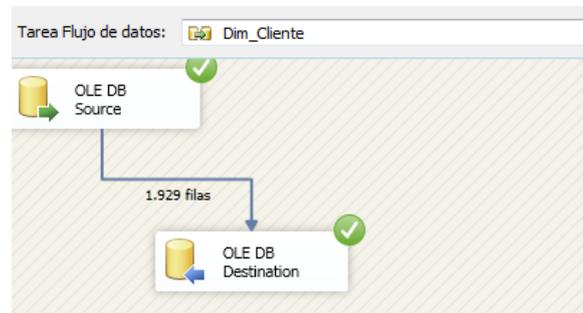
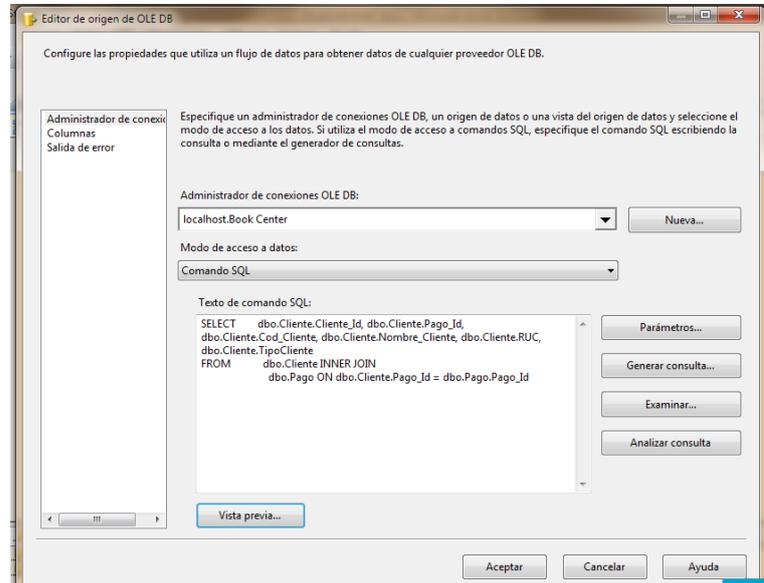
### Paso 3: Poblamiento de la Dimensión Cliente

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

- a. 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
```

- b. Determinamos la Dimensión Cliente como destino de datos.
- c. 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,
```

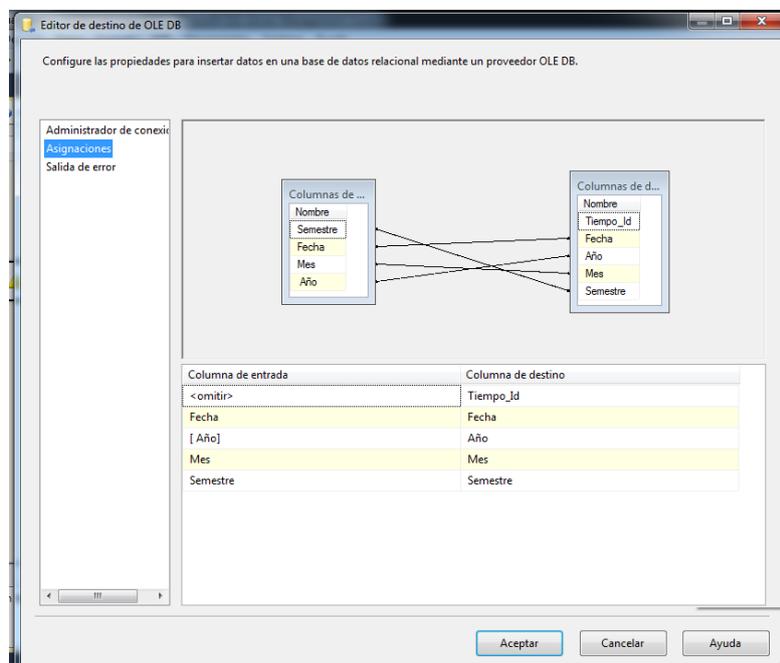
```
DatePart(mm, T.TheDate) AS [Mes], DATENAME(MM,T.TheDate) AS  
[Nombre Mes],
```

```
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.



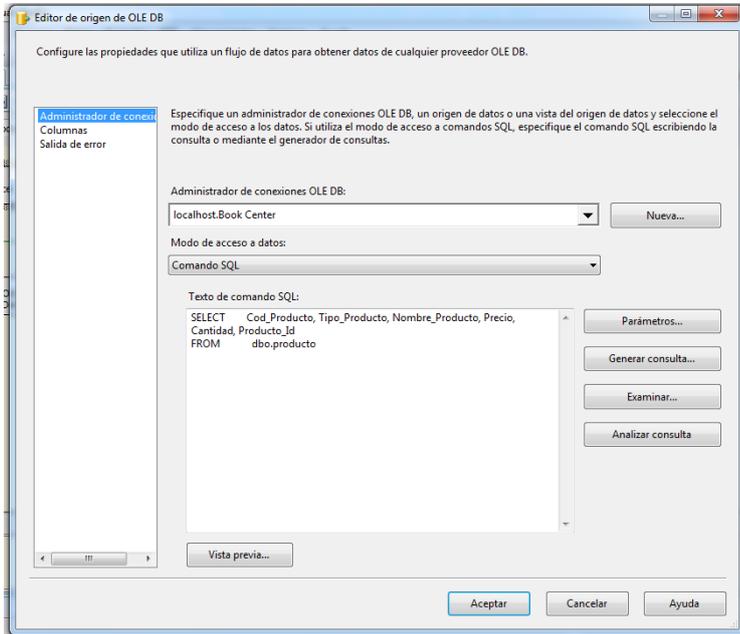
### Paso 5: Poblamiento de la Dimensión Producto

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

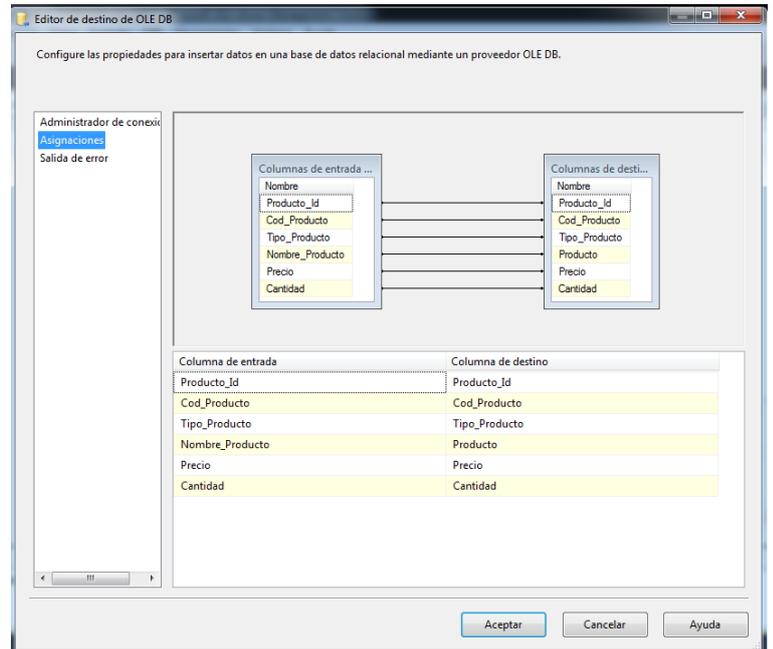
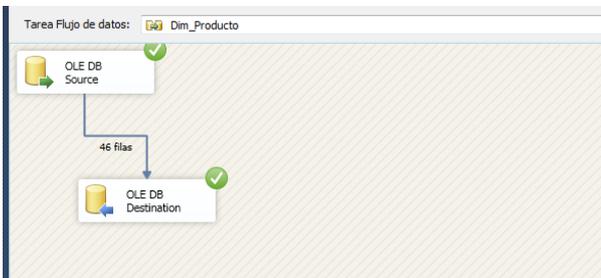
- Definimos una sentencia SQL como origen de datos.
 

```
SELECT    Cod_Producto, Tipo_Producto,
           Nombre_Producto, Precio, Cantidad
FROM      dbo.producto
```
- Determinamos la Dimensión Producto como destino de datos.

f. 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
```

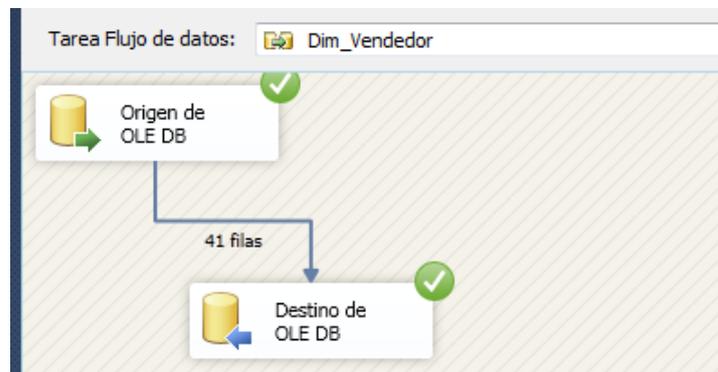
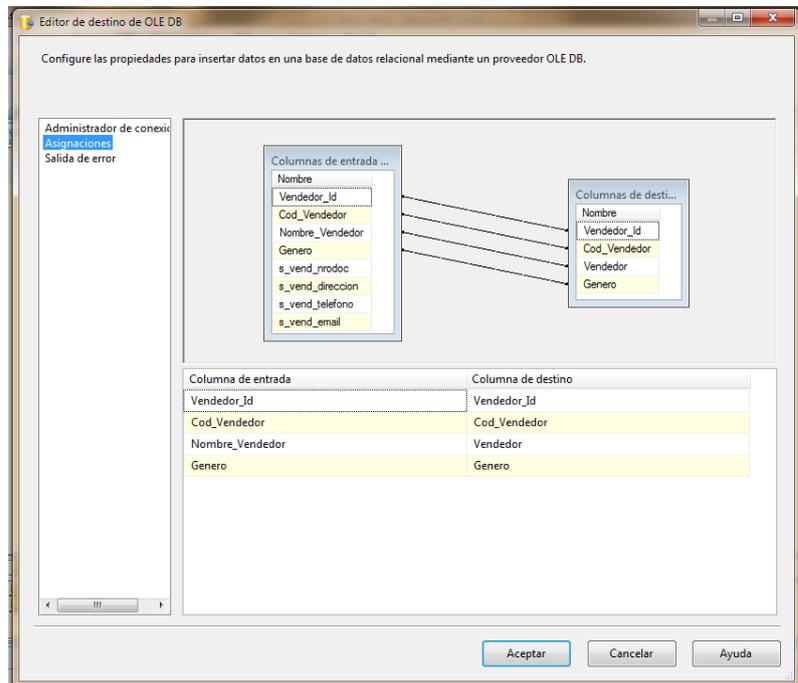


## Paso 6: Poblamiento de la Dimensión Vendedor

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

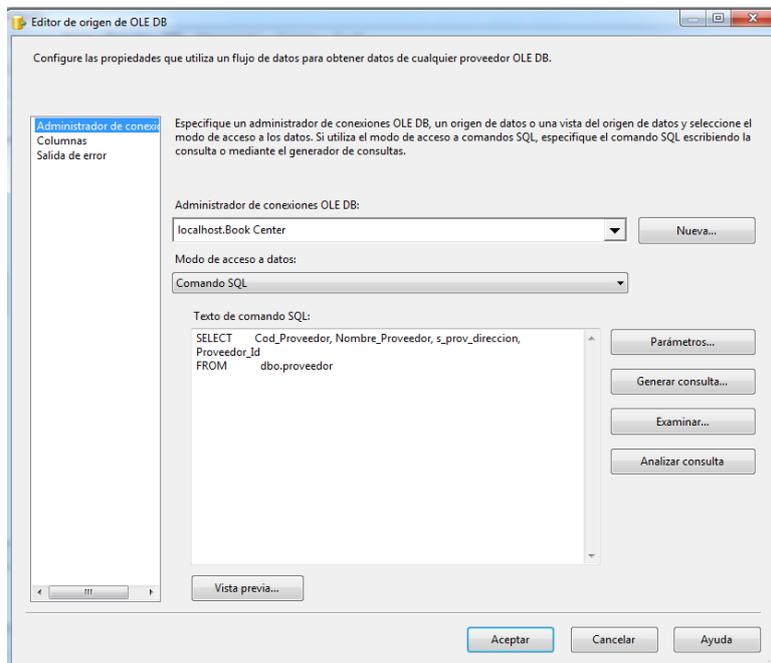
Para transferir los datos a la dimensión Proveedor, se efectúan los siguientes pasos:

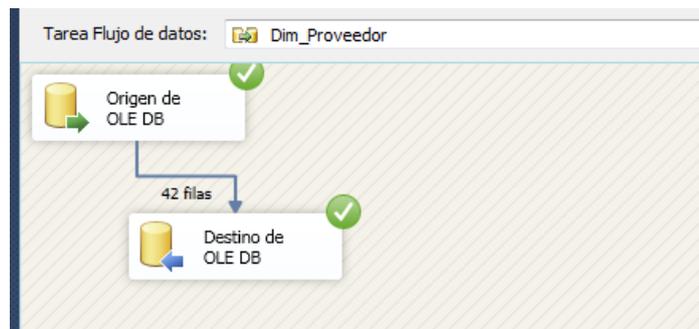
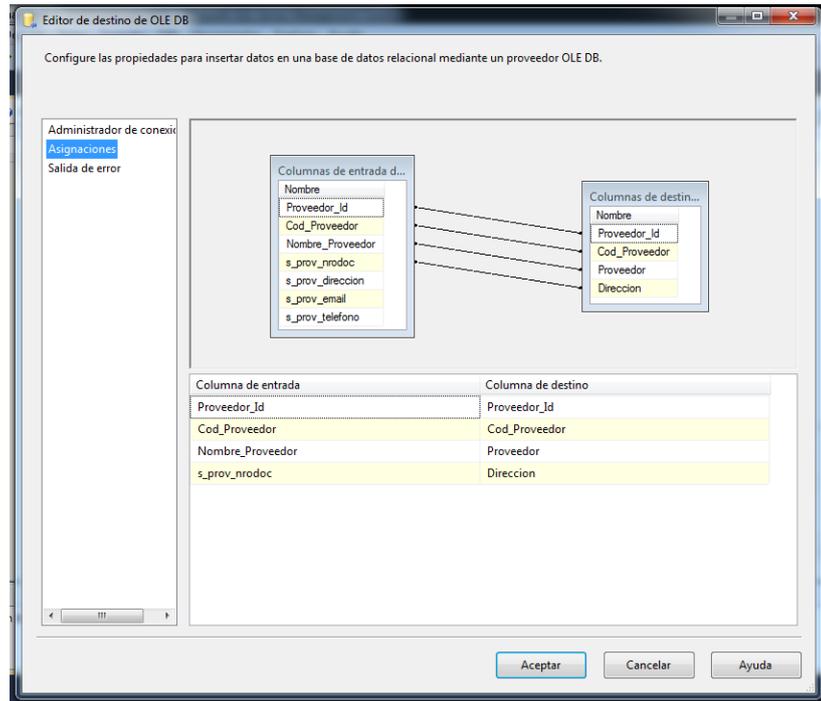
- g. Definimos una sentencia SQL como origen de datos.

```
SELECT          Cod_Proveedor, Nombre_Proveedor, s_prov_direccion,
Proveedor_Id
FROM            dbo.proveedor
```

- h. Determinamos la Dimensión Proveedor como destino de datos.

- i. En las transformaciones se determina que columnas del origen corresponderán con las de la Dimensión Proveedor





### Paso 8: Poblamiento de las Tablas de Hechos

Para poblar la **TABLA DE HECHOS VENTAS**, se efectúan los siguientes pasos:

- Definimos una sentencia SQL como origen de datos.

```
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 [Book Center].dbo.Cliente CL INNER JOIN [Book
Center].dbo.Pedido PS ON
    CL.Pedido_Id = PS.Pedido_Id INNER JOIN [Book Center].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.

```



**Figura N° 60:** Poblamiento de la Tabla de Hechos

## 2.7 Selección de Productos

### 2.7.1 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.

### 2.7.2 Software

Para la implementación se trabajará con ERwin Data Modeler r7, MS SQL Server 2012. En la tabla N°32 se detalla la herramienta a utilizar y los procesos a ejecutar.

**Tabla N° 32:** Software para el proceso de Diseño del DataMart

Componente del Proceso	Herramienta
<b>Construcción de Interfaces</b>	MS Excel 2013, PowerBI
<b>Cubos</b>	AnalysServices
<b>ETL</b>	SQL Server Business Intelligence
<b>DataMart</b>	Microsof Azure
<b>Modelo Lógico y Físico</b>	ERwin Data Modeler r7 y SQLServer

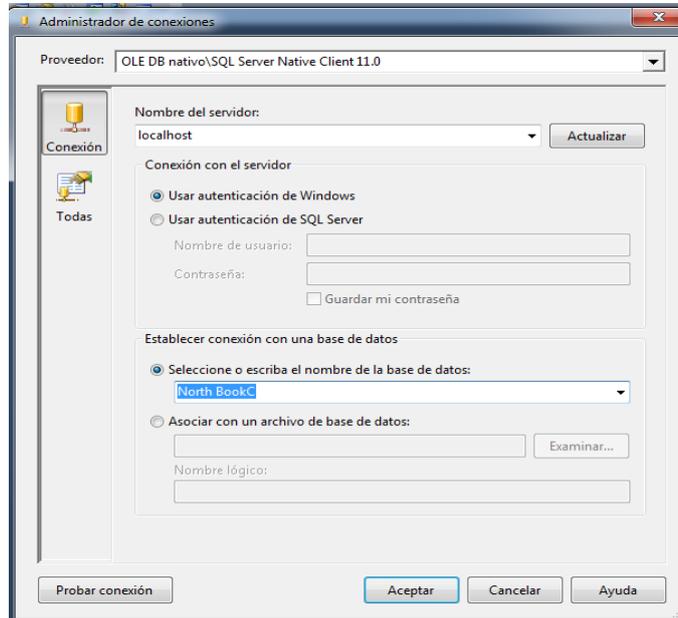
## 2.8 Especificación de la Aplicación del Usuario Final

### 2.8.1 Estructura de Cubo

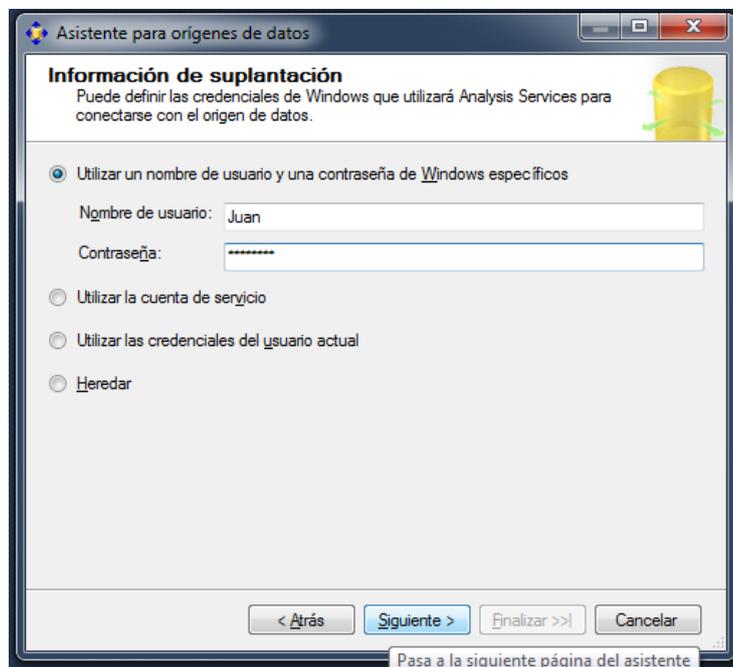
La herramienta utilizada para el diseño del cubo es el SSAS (SQL Server Analysis Services) que es el diseñador de cubos de Business Intelligence Development Studio para crear el cubo, incluido las medidas, las dimensiones y sus respectivas relaciones.

El cubo de Ventas se llama “CUBO\_VENTAS”. Las medidas y dimensiones de este cubo son originadas por su equivalente a las tablas del modelo dimensional como se muestra en la Figura N° 61

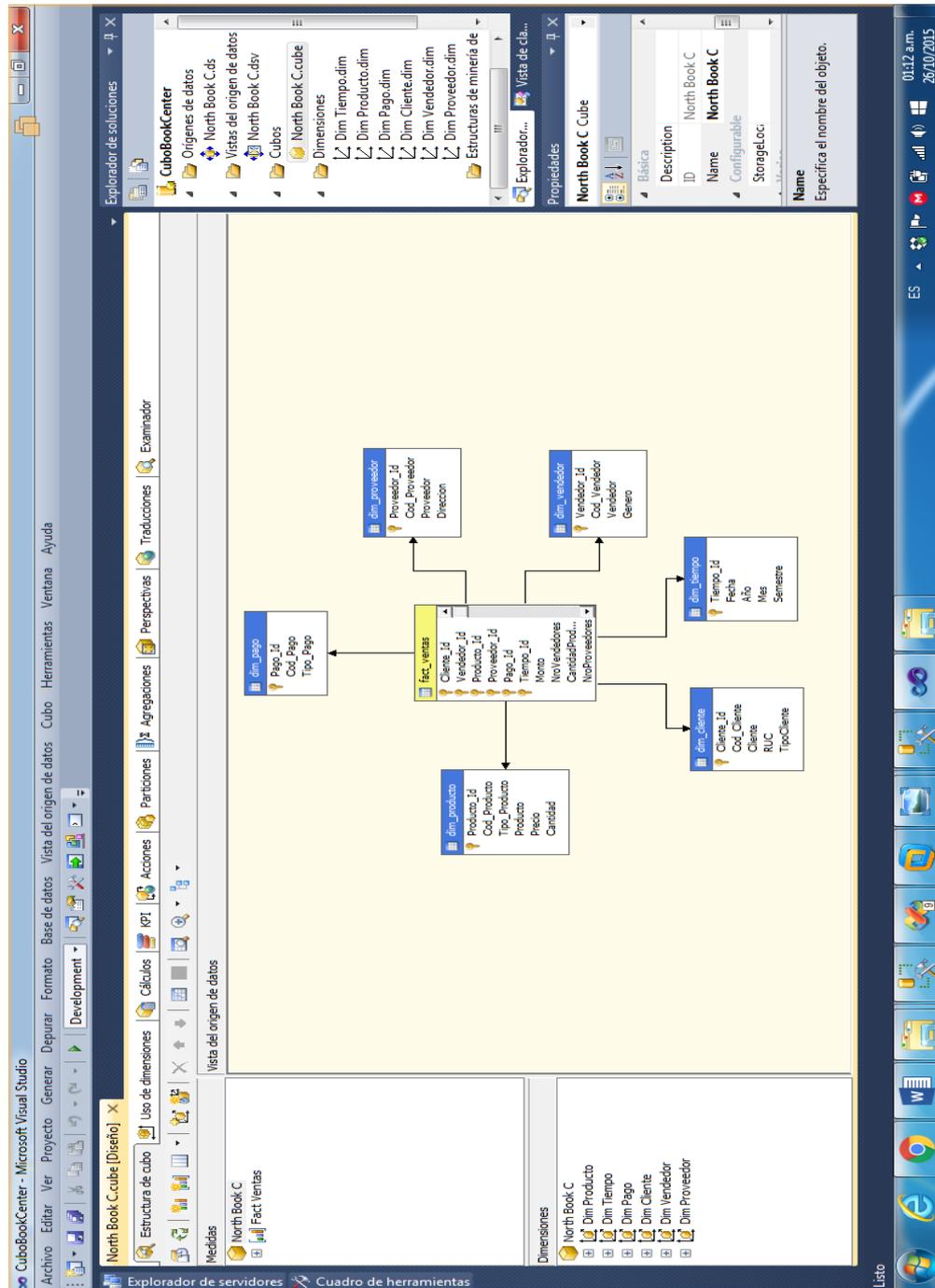
Primero definimos el nombre del servidor en este caso “localhost” y luego seleccionamos la base de dato (datawarehouse)



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 NORTH BOOK C en el Data Source Views, seleccionando las tablas a usar.



**Figura N° 61:** Diseño de Cubo de Ventas

- b) Se procede a la creación del cubo definiendo primero la dimensión tiempo y posteriormente las demás dimensiones y las medidas.

### Definiendo Dimensión Tiempo

Asistente para dimensiones

**Especificar información de origen**  
Seleccione un origen de datos y especifique cómo se enlaza la dimensión al mismo.

Vista del origen de datos:  
North Book C

Tabla principal:  
dim\_tiempo

Columnas de clave:  
Tiempo\_Id  
(Agregar columna de clave)

Columna de nombre:  
Tiempo\_Id

< Atrás    **Siguiente >**    Finalizar >>    Cancelar

Nota: Una barra de herramientas inferior muestra un botón 'Pasa a la siguiente página del asistente'.

Asistente para dimensiones

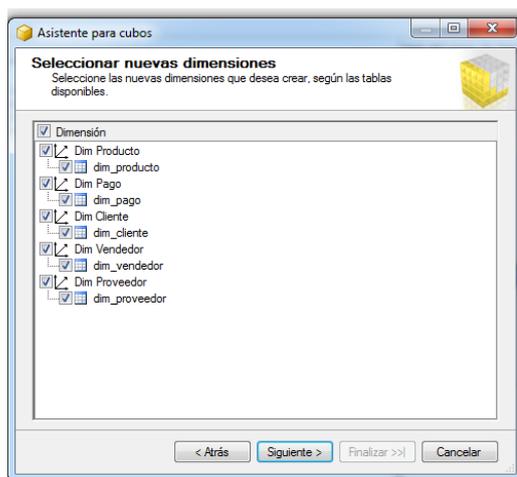
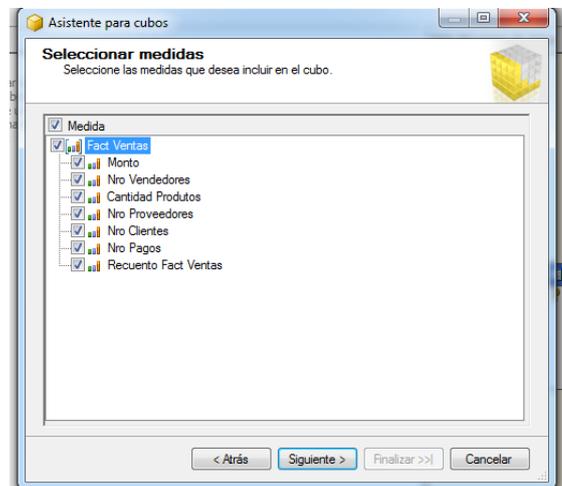
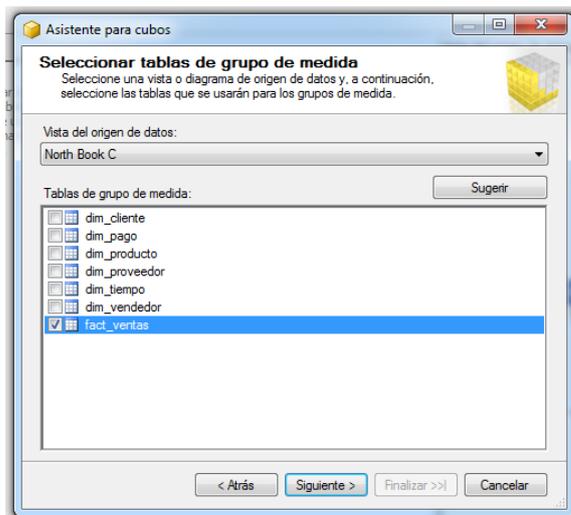
**Seleccionar los atributos de la dimensión**  
Especifique los atributos de dimensión y seleccione Habilitar exploración para mostrarlos como jerarquías.

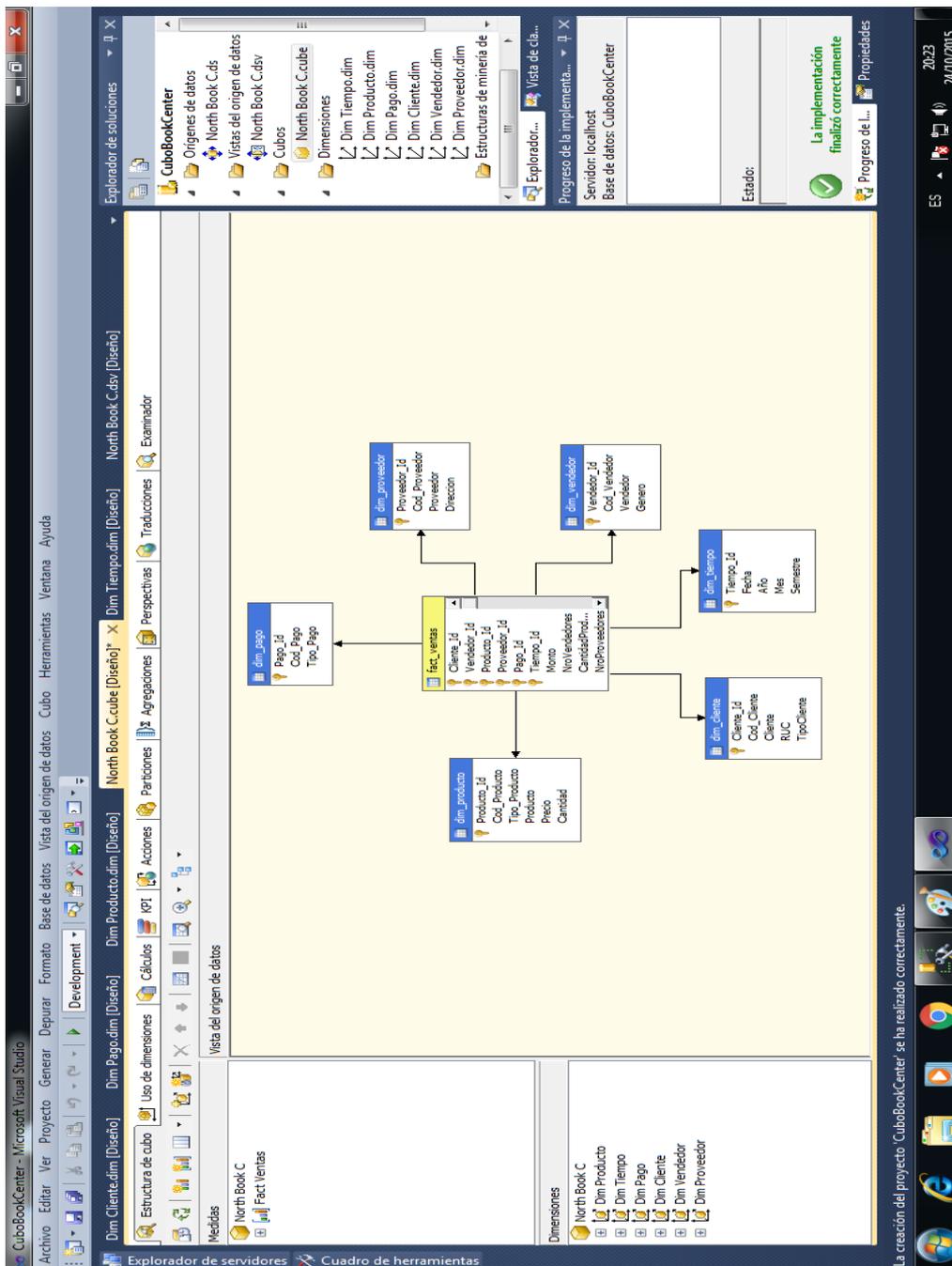
Atributos disponibles:

<input checked="" type="checkbox"/> Nombre del atributo	<input checked="" type="checkbox"/> Habilitar explor...	Tipo de atributo
<input checked="" type="checkbox"/> Tiempo Id	<input checked="" type="checkbox"/>	Normal
<input checked="" type="checkbox"/> Fecha	<input checked="" type="checkbox"/>	Fecha
<input checked="" type="checkbox"/> Año	<input checked="" type="checkbox"/>	Año
<input checked="" type="checkbox"/> Mes	<input checked="" type="checkbox"/>	Mes
<input checked="" type="checkbox"/> Semestre	<input checked="" type="checkbox"/>	<b>Semestre</b>

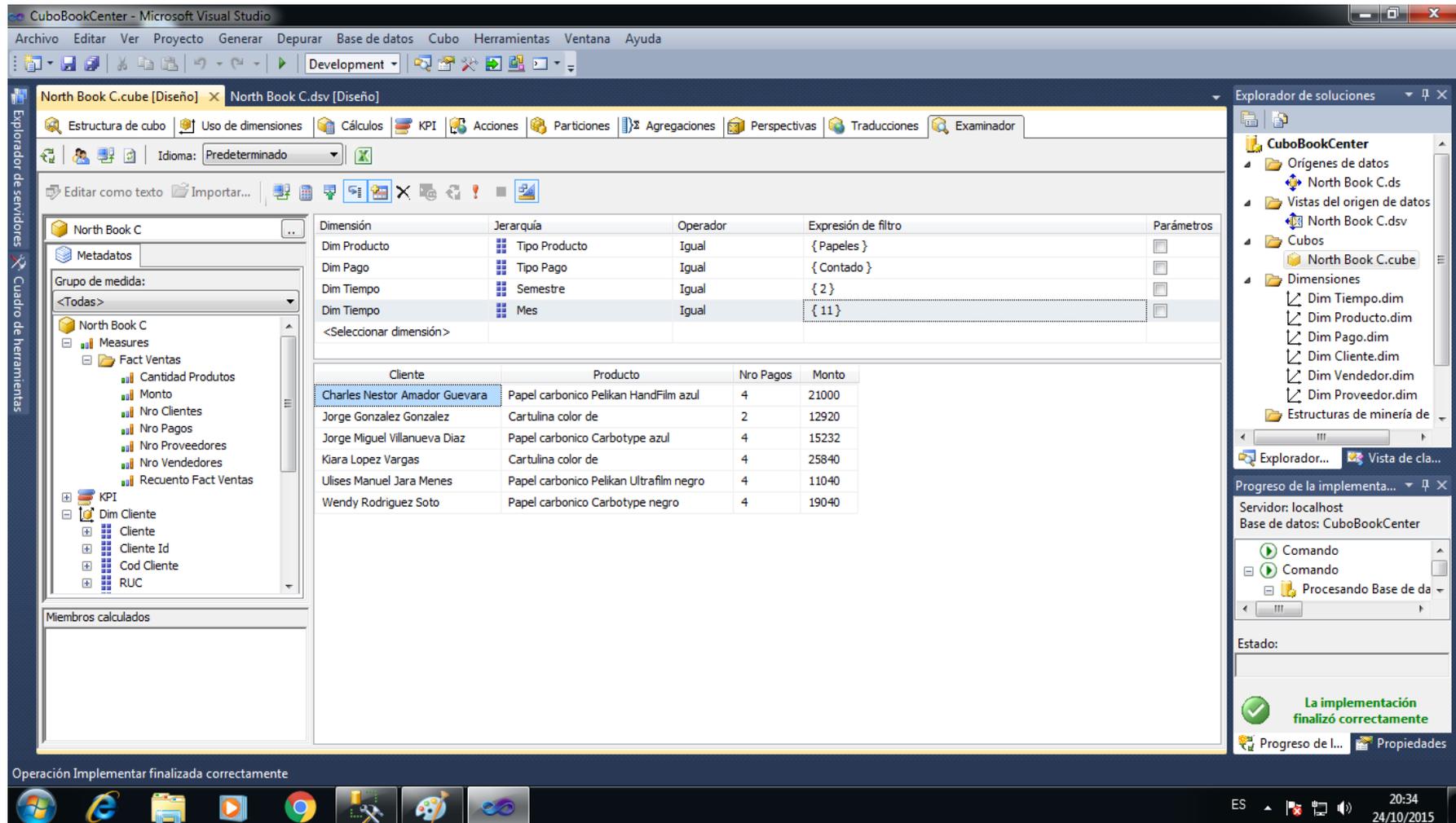
< Atrás    **Siguiente >**    Finalizar >>    Cancelar

Creación Cubo





Mostrando Implementación



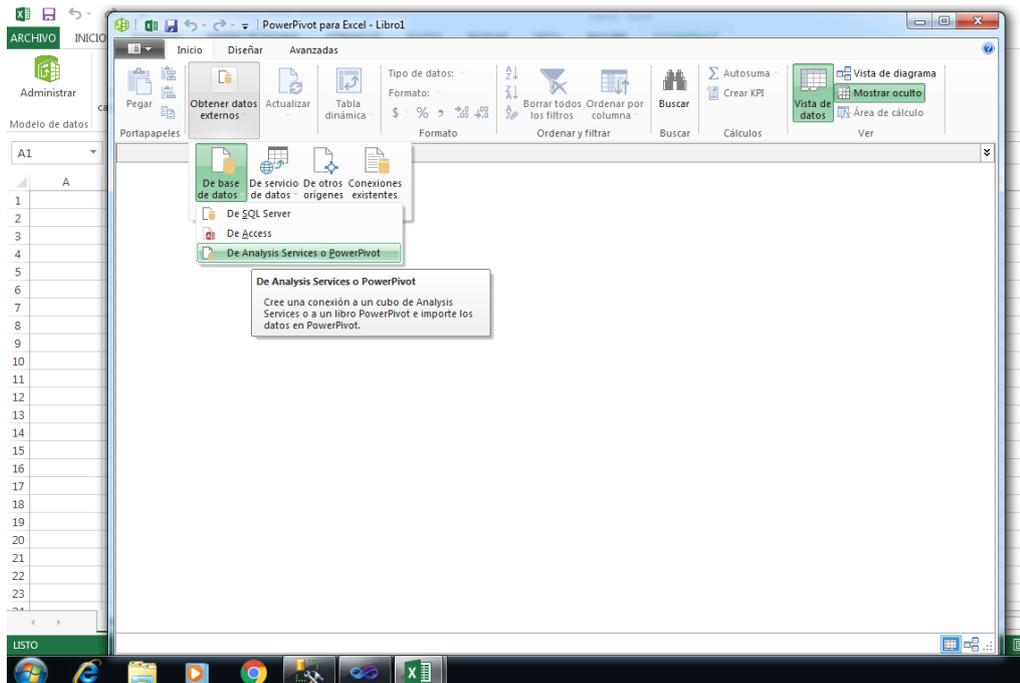
## 2.9 Desarrollo de la Aplicación del Usuario Final

El desarrollo de las aplicaciones de los usuarios finales involucra configuraciones del metadata y construcción de reportes específicos. Las Figuras que se van a mostrar a continuación son interfaces personalizadas de cada Requerimiento creadas en Excel con herramientas comunes, utilizando los cubos Ventas e Inventario anteriormente descritos.

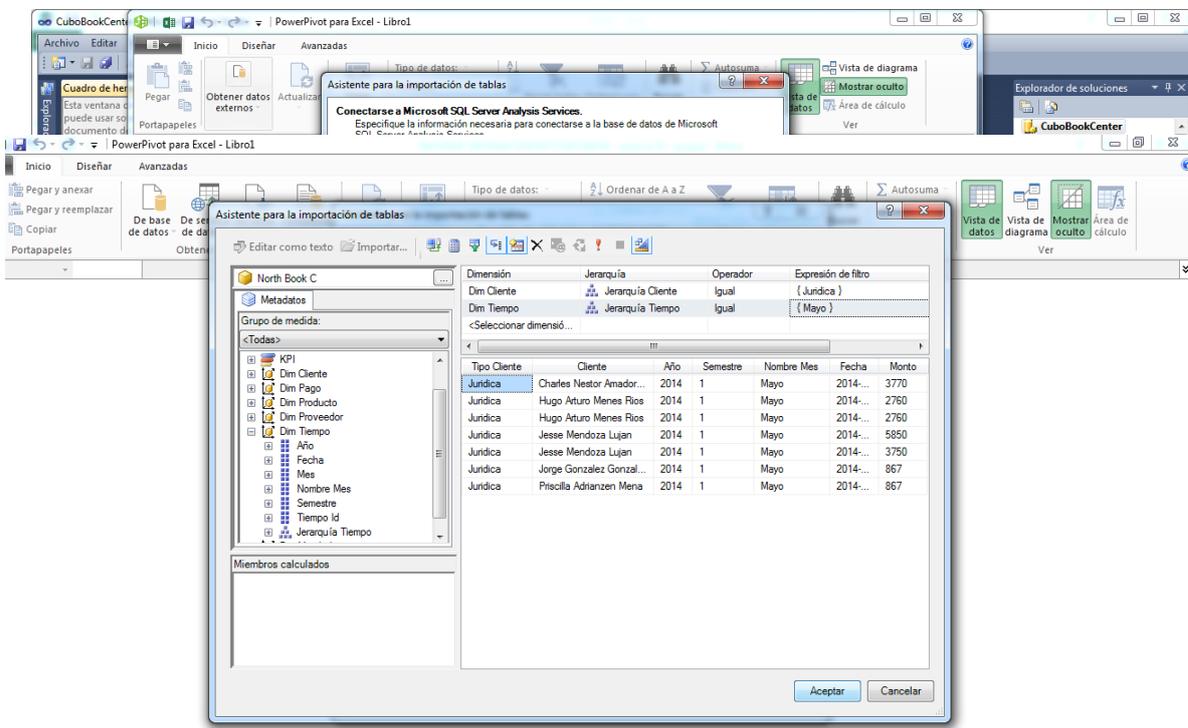
### 1. Elegir la herramienta para procesamiento analítico.

Se han usado Microsoft Visual (Business Intelligence) y Microsoft PowerBI(Power Pivot)

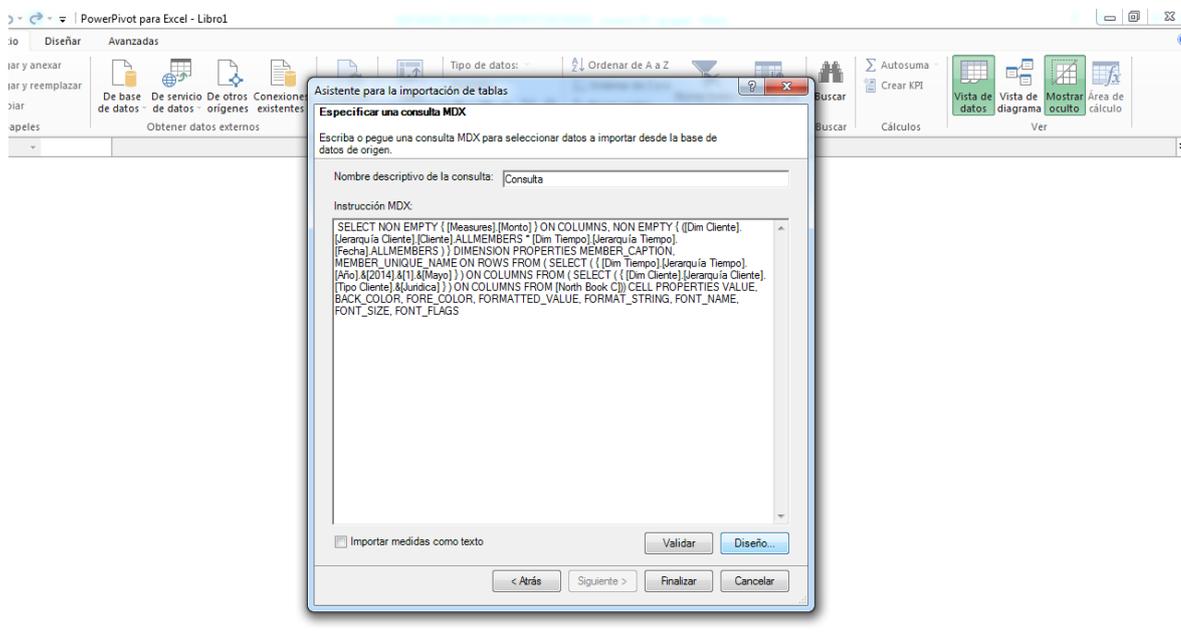
### 2. Determinar los cubos.



3. Diseñar los reportes a mostrar.

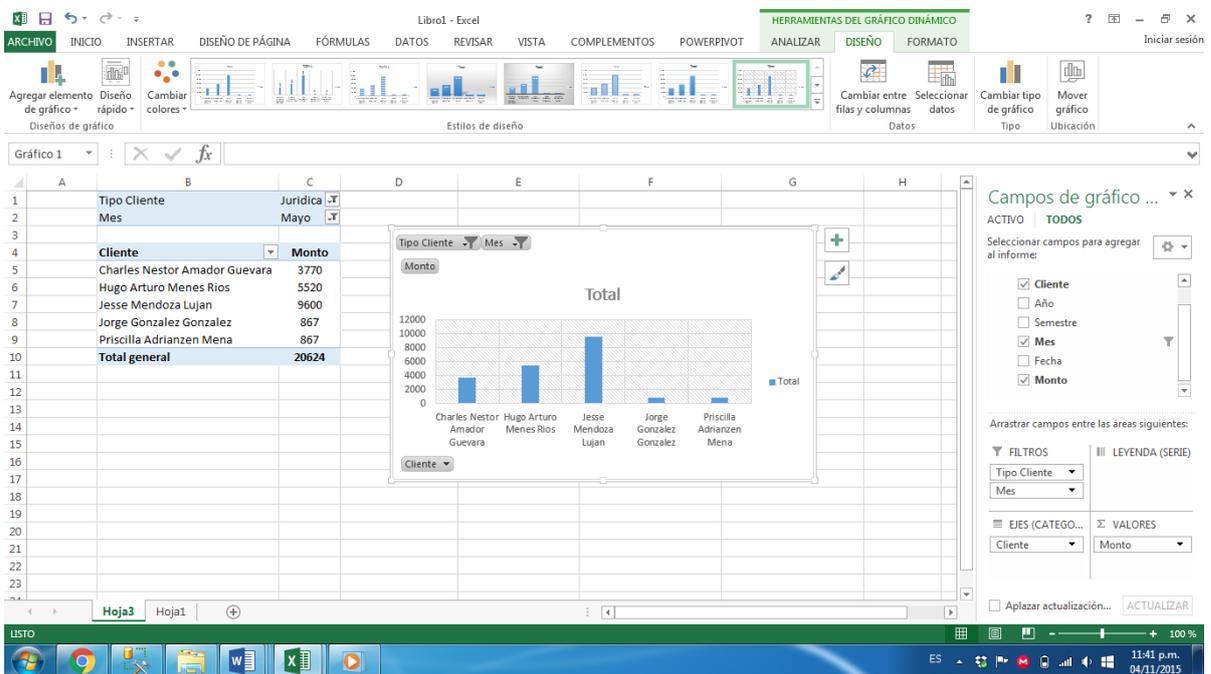


4. Selección de herramienta de aplicación para usuario final: se usara Ms. Excel. Power Pivot



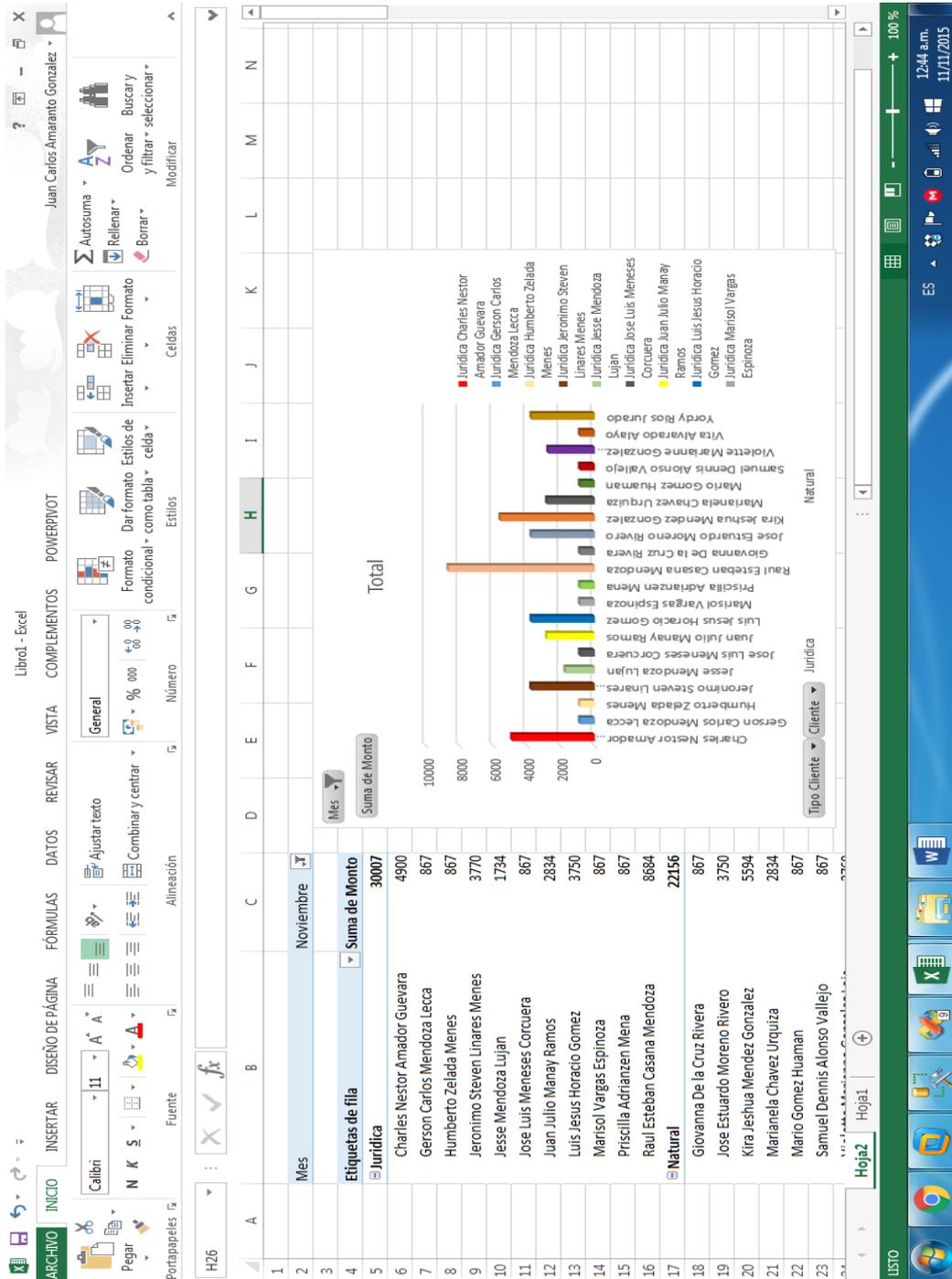
### 5. Implementación de la herramienta Power Pivot de Ms Excel.

Tipo Cliente	Cliente	Año	Semestre	Mes	Fecha	Monto
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

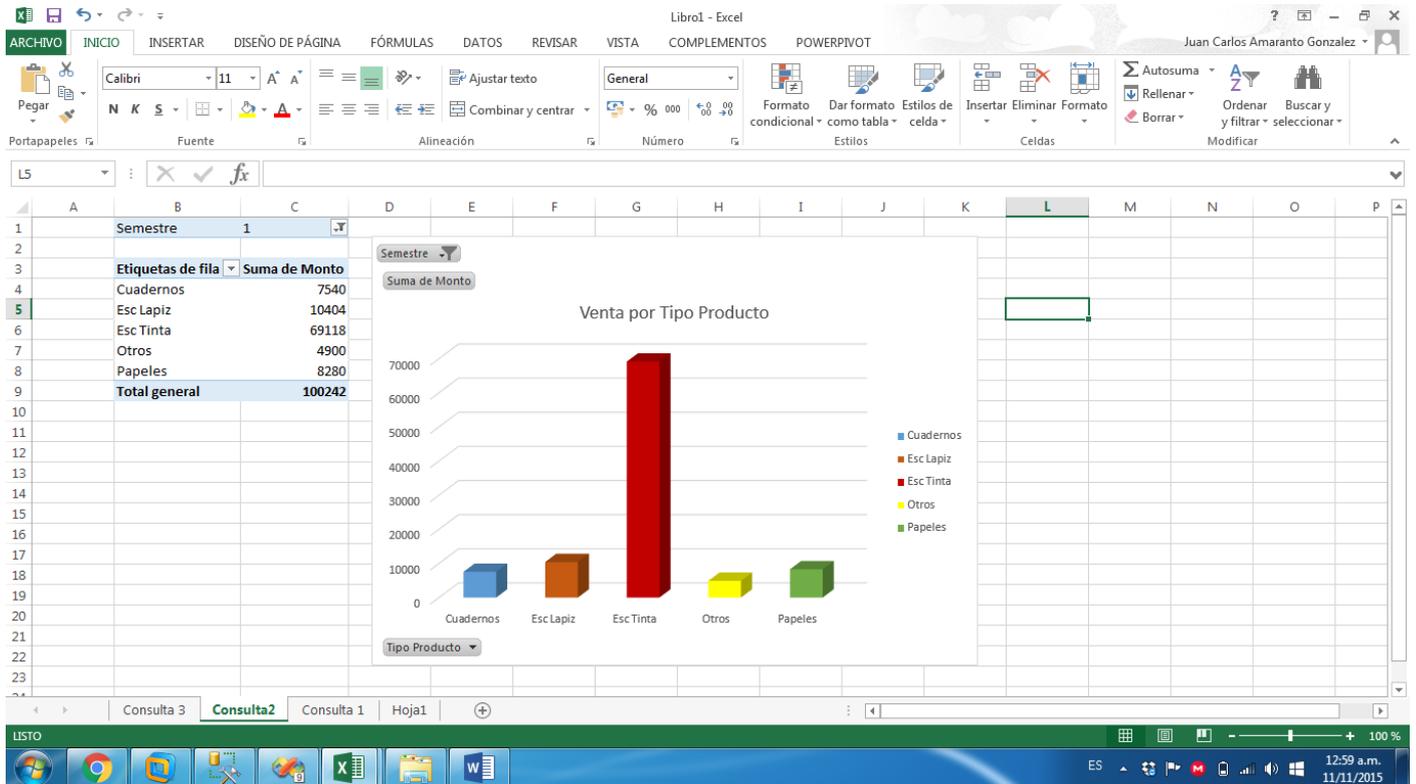


REPORTES EN POWERPIVOT

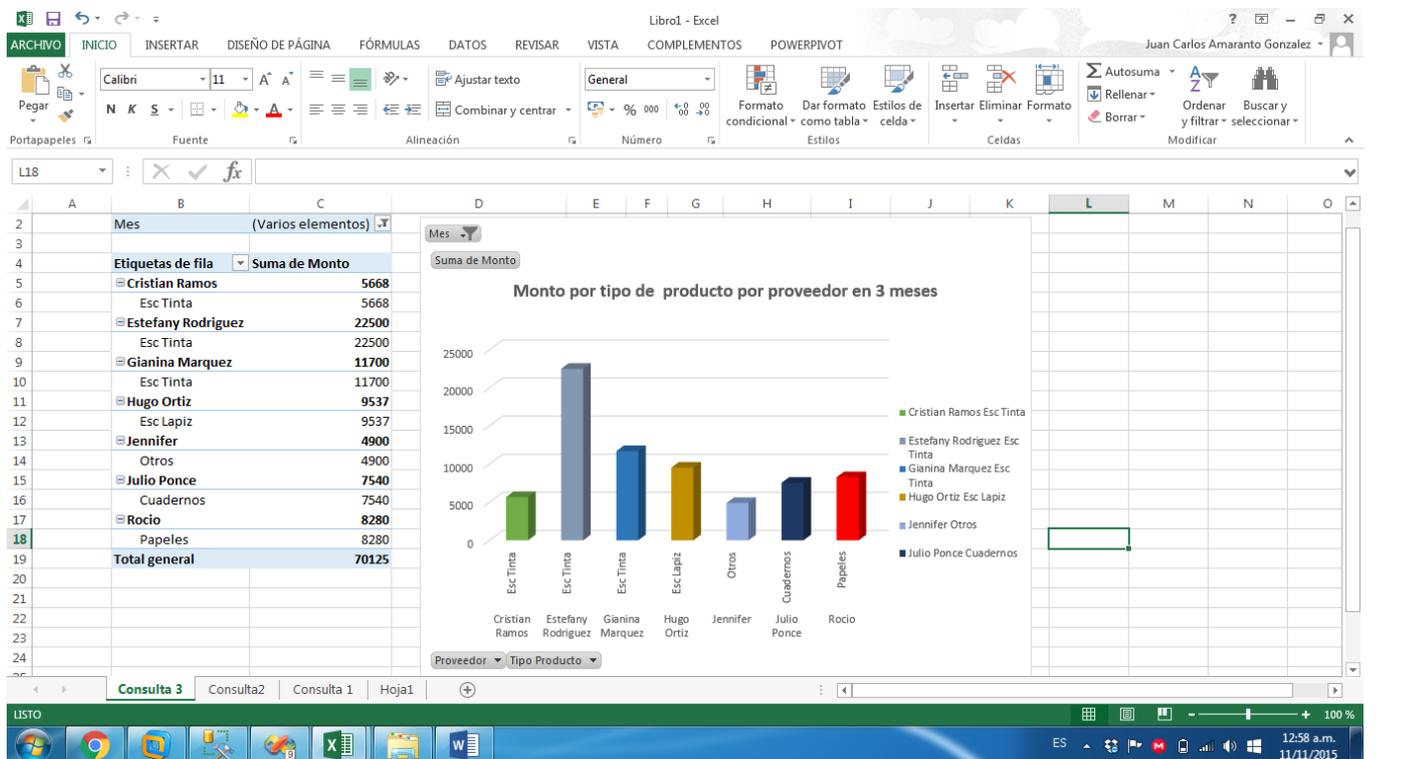
a. **Requerimiento 1: ¿Cuál es el monto de un determinado cliente en compras en un mes?**



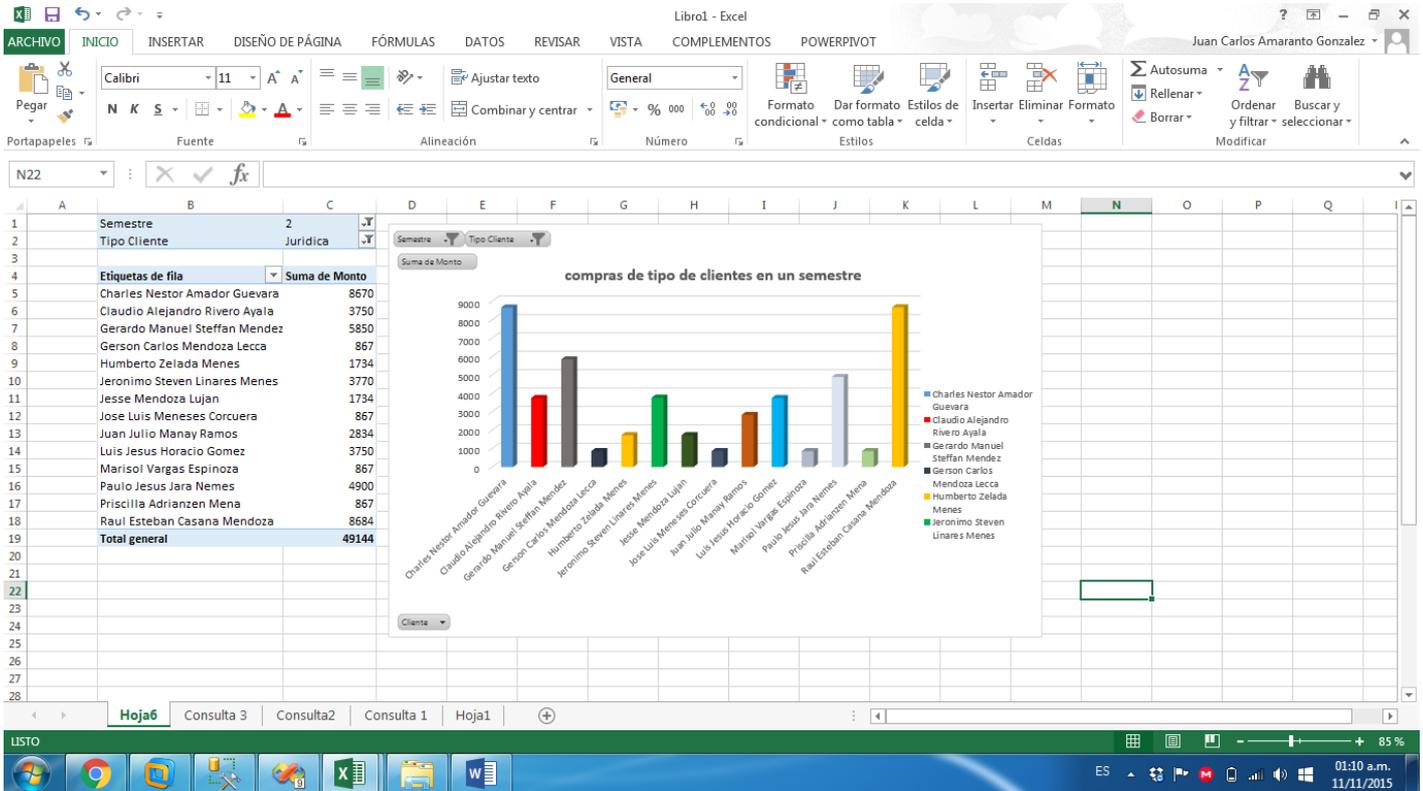
**b. Requerimiento 2: ¿Cuál es el Monto de ventas por tipo de producto?**



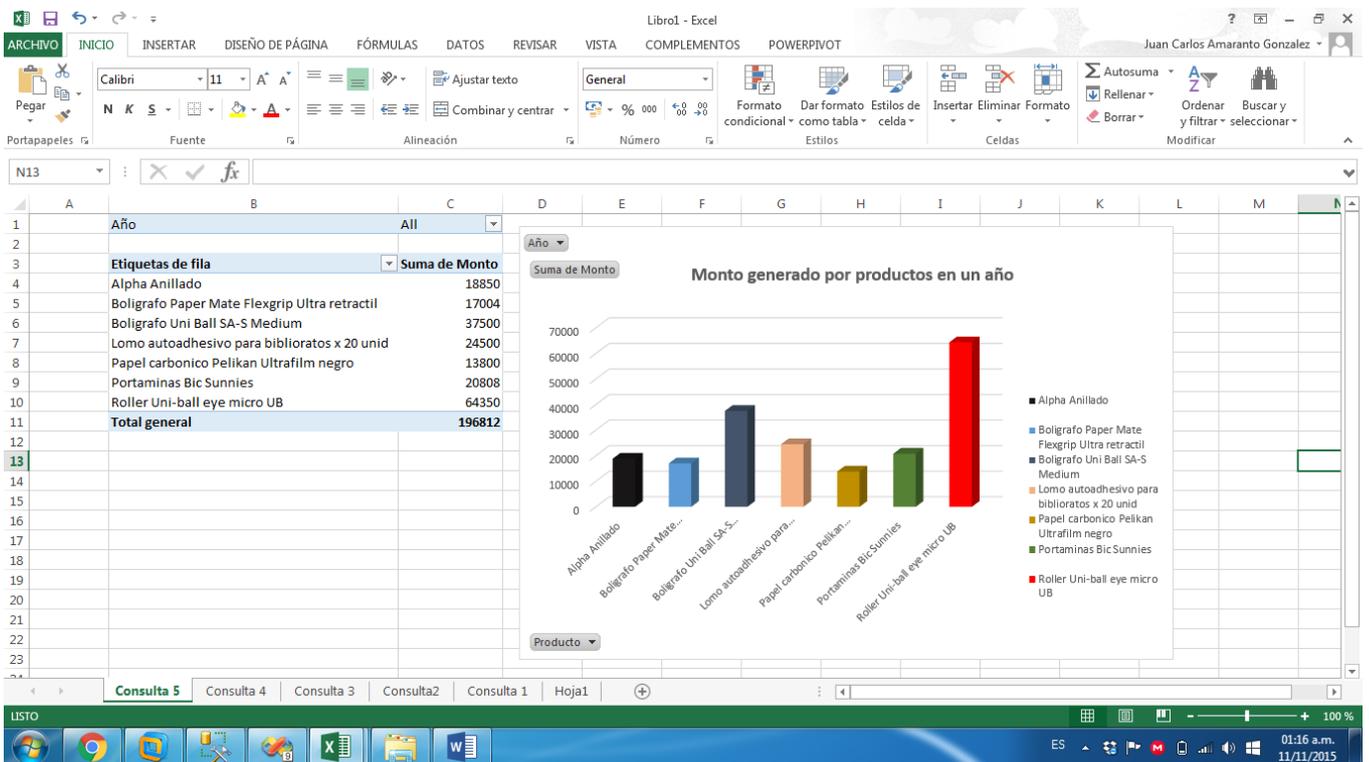
**c. Requerimiento 3: ¿Cuál es el monto de un tipo de producto por proveedor en 3 meses?**



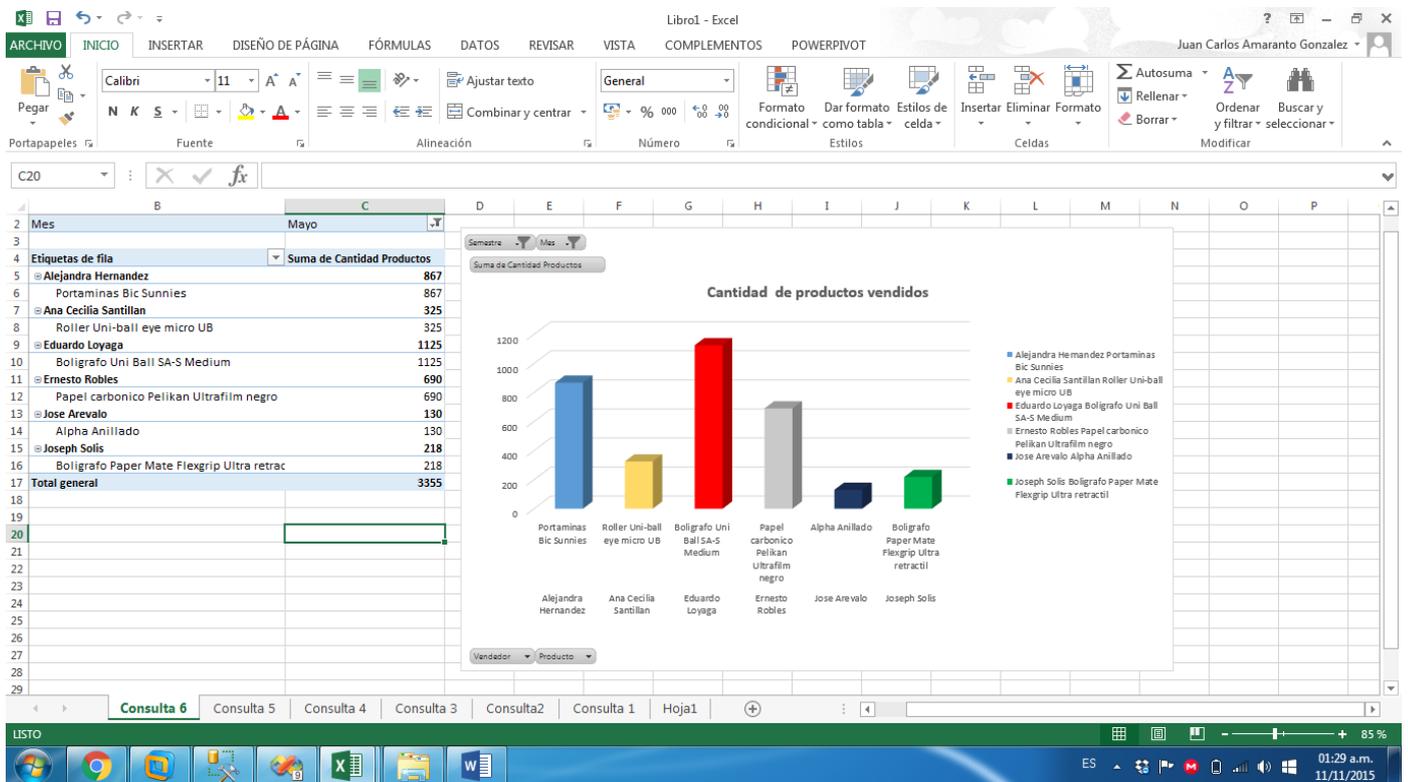
**d. Requerimiento 4: ¿Cuál es el monto de compras de tipos de clientes en un semestre?**



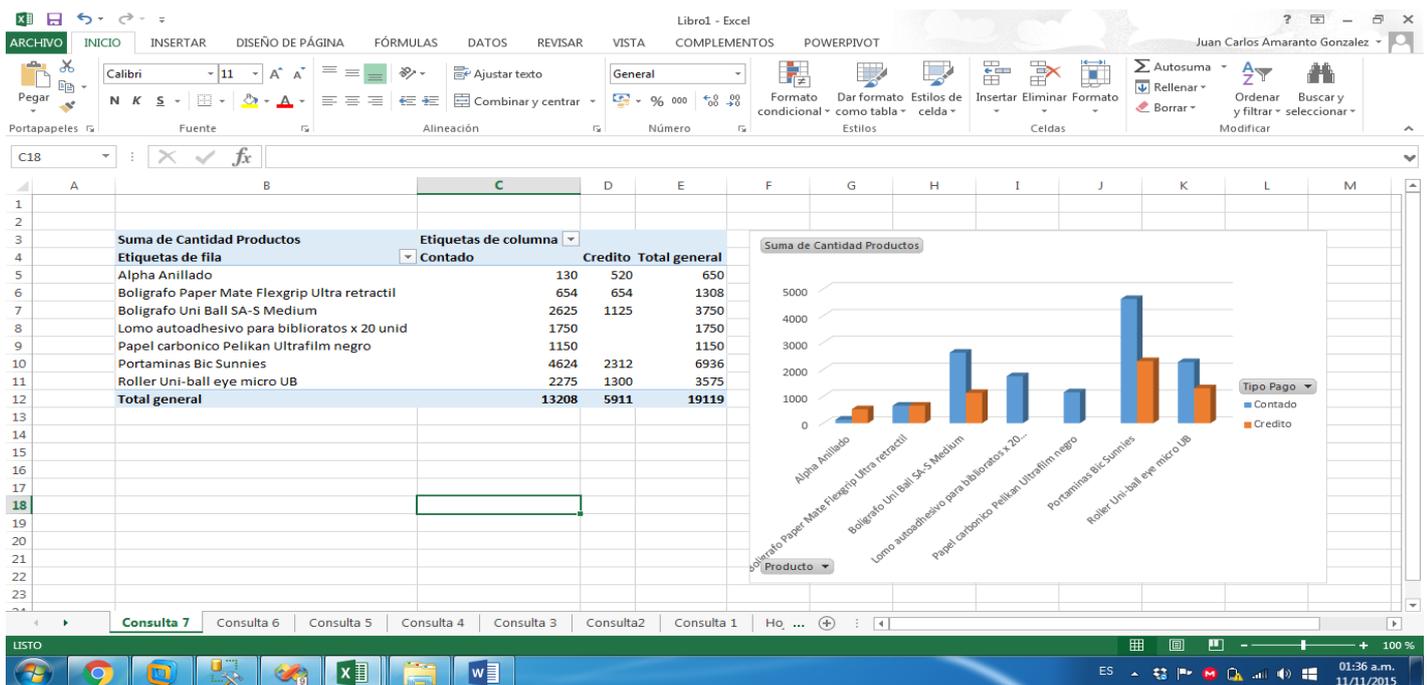
**e. Requerimiento 5: ¿Cuál es el monto que generan los productos en un año?**



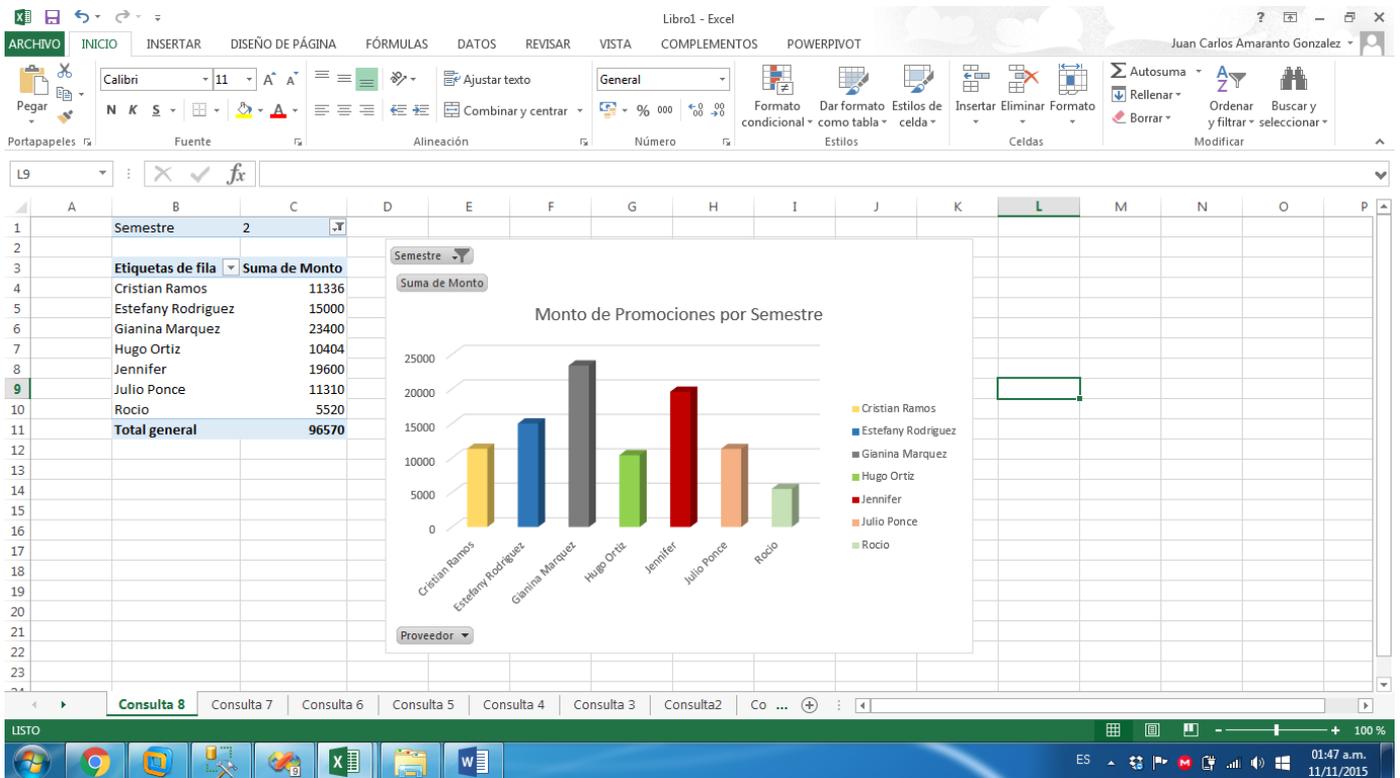
**f. Requerimiento 6: ¿Cuál es la cantidad de productos vendidos por Tipo, vendedor, mes y semestre? Cantidad de Productos**



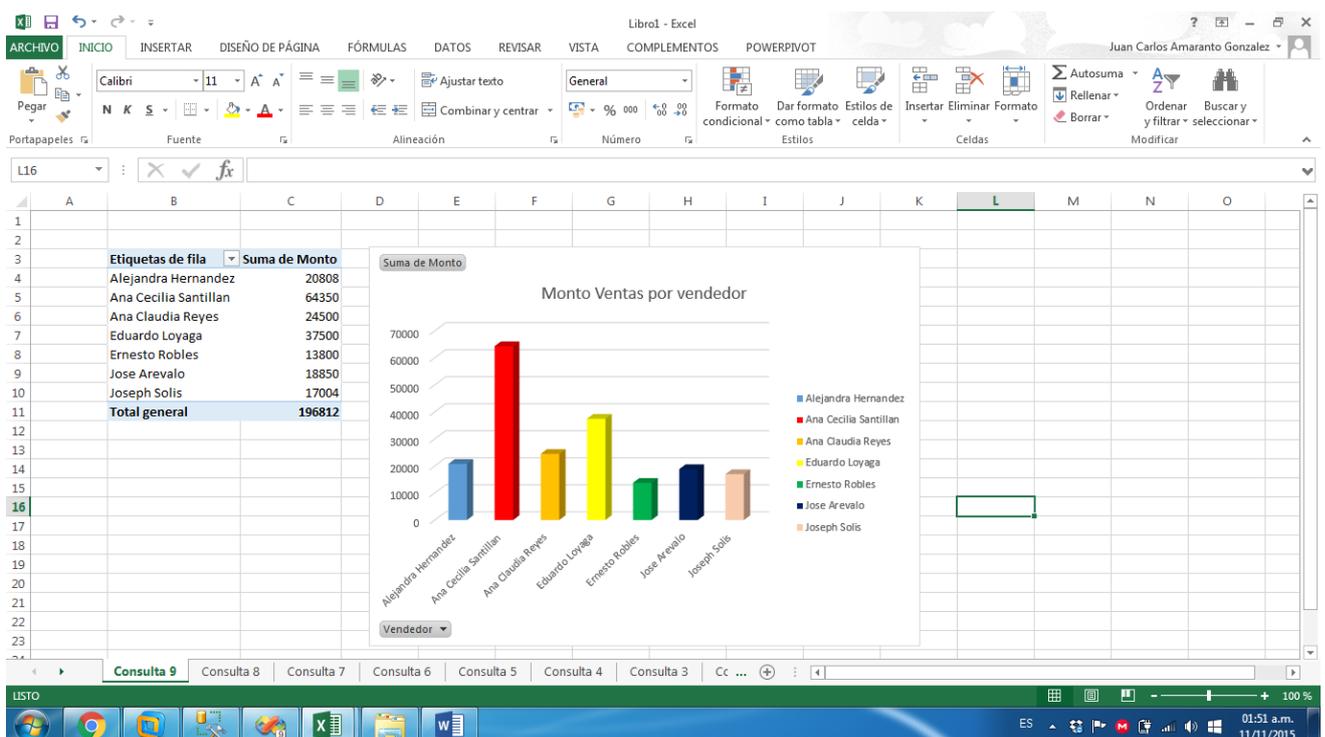
**g. Requerimiento 7: ¿Cuál es la cantidad de productos que se han pagado al crédito y al contado?**



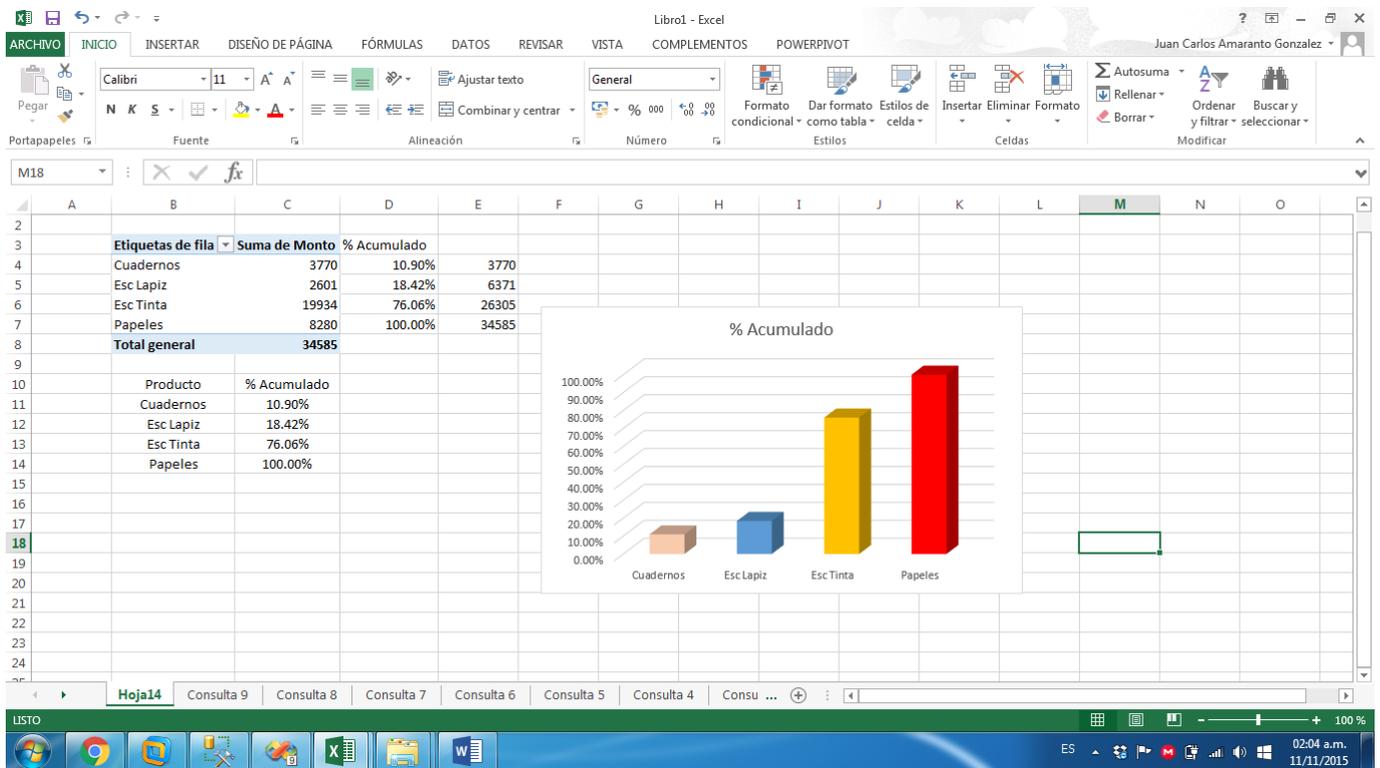
**h. Requerimiento 8: ¿Cuál es el monto que generan las promociones que da un proveedor en un semestre?**



**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?**



## CAPITULO III

### DISCUSIÓN

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

#### Formulación del Problema

¿Cómo mejorar el soporte online en la toma de decisiones para el área de ventas de la empresa Book Center S.A.C, utilizando tecnologías de información?

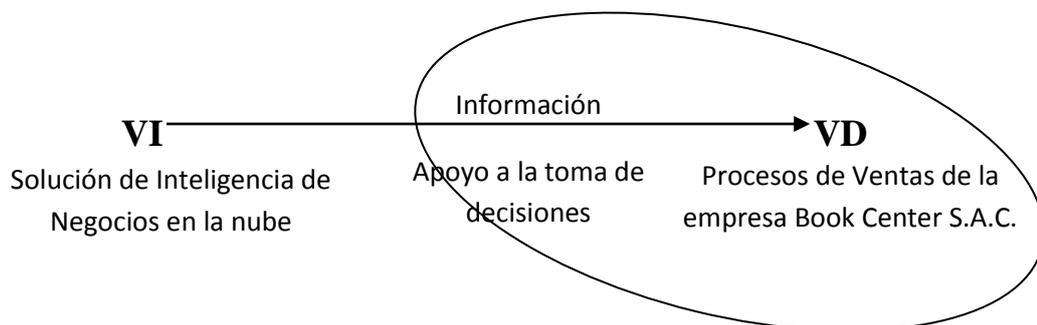
#### Hipótesis

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 Book Center S.A.C.

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): Mejor apoyo en la Toma de Decisiones en el área de Ventas de la empresa Book Center S.A.C.”.

#### 3.1. MANERA PRESENCIAL



### 3.1.1. DISEÑO PREEXPERIMENTAL 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

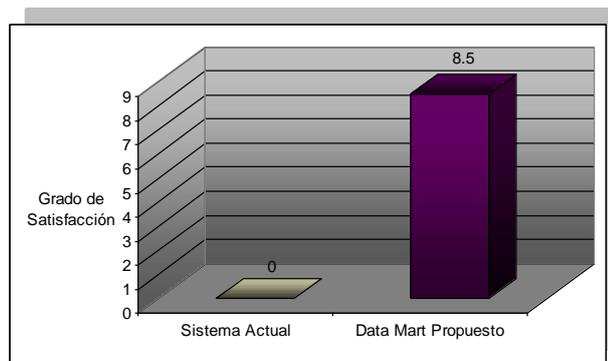
Para el cálculo de los indicadores de la hipótesis en el Data Mart Propuesto (DMP) propuesto y el Sistema Actual (SA), se realizó un cuestionario (Ver Anexo B) donde se evaluó a los usuarios luego de haber interactuado con el Data Mart.

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

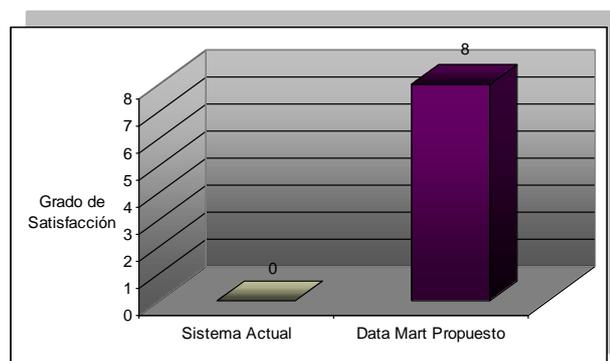
Tabla 3.1. Rango de grado de satisfacción

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



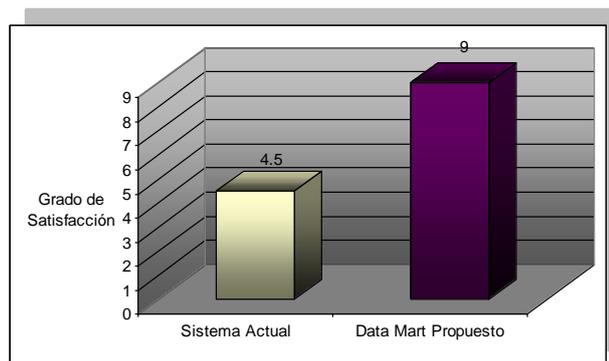
**Resultado:** Los valores obtenidos son : 0 (S.A.) y 8.5 (D.M.P). El Data Mart muestra información sobre la cantidad total de ventas de varios productos con comparativas de meses mediante tablas y gráficos

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



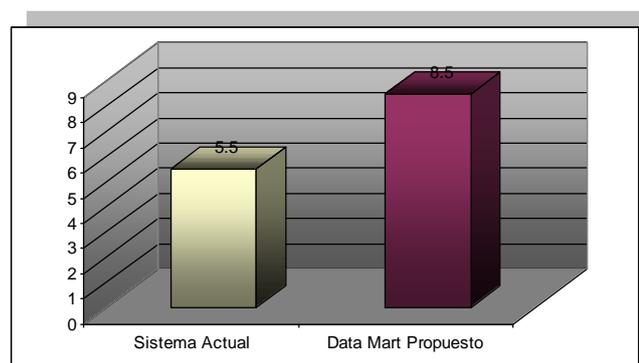
**Resultado:** Los valores obtenidos son : 0 (S.A.) y 8 (D.M.P). El Data Mart muestra información sobre la cantidad de ventas por tipo haciendo una comparación de ventas de meses anteriores.

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



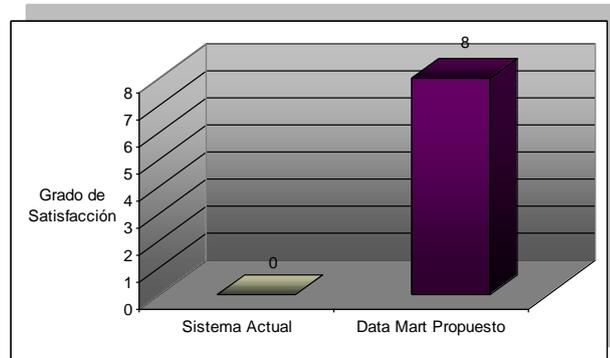
**Resultado:** Los valores obtenidos son : 4.5 (S.A.) y 9 (D.M.P). El Sistema Actual emite un resumen de productos por proveedor pero no se puede hacer comparaciones entre periodos. El Data Mart presenta dinámicamente y con gráficos para un mejor entendimiento.

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



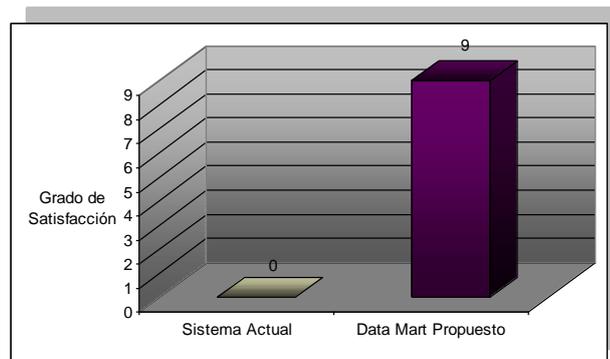
**Resultado:** Los valores obtenidos son : 5.5 (S.A.) y 8.5 (D.M.P). El Sistema Actual permite conocer la compras de clientes, pero no lo presenta en forma dinámica. El Data Mart presenta dinámicamente y con gráficos para un mejor entendimiento.

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



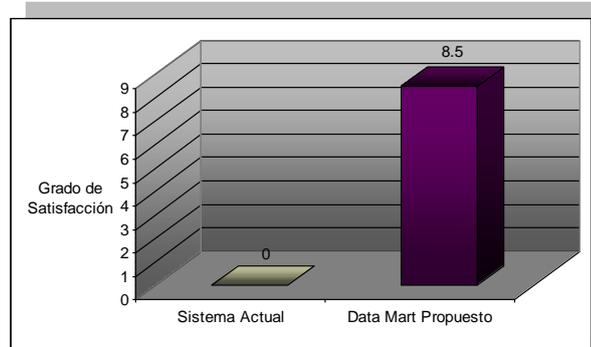
**Resultado:** Los valores obtenidos son: 0 (S.A.) y 8 (D.M.P). El Data Mart muestra información sobre las ventas totales que se hizo a un cliente. El Sistema Actual no presenta esta consulta.

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



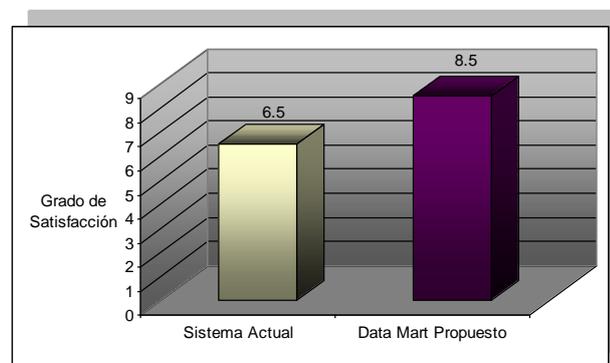
**Resultado:** Los valores obtenidos son: 0 (S.A.) y 8 (D.M.P). El Data Mart muestra información sobre los productos en diferentes periodos de tiempo. El Sistema Actual no presenta esta consulta

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



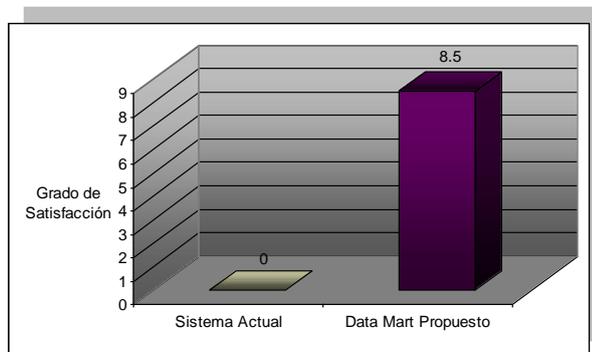
**Resultado:** Los valores obtenidos son: 0 (S.A.) y 8.5 (D.M.P). El Data Mart muestra información sobre los tipos de pago en forma dinámica. El Sistema Actual no presenta esta consulta

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



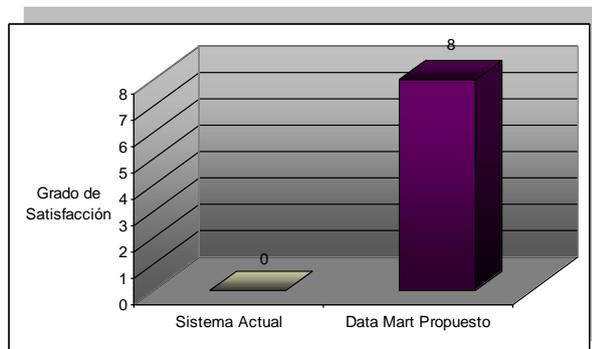
**Resultado:** Los valores obtenidos son: 6.5 (S.A.) y 8.5 (D.M.P). El Sistema Actual permite conocer esta información, pero no lo presenta en forma dinámica. El Data Mart presenta dinámicamente y con gráficos para un mejor entendimiento

**Pregunta N° 9: ¿Se puede conocer la Lista de Vendedores y el monto total de cada una de las ventas que han generado?**



**Resultado:** Los valores obtenidos son : 0 (S.A.) y 8.5 (D.M.P). El Data Mart muestra información sobre los vendedores y su monto de ventas de forma dinámica. El Sistema Actual no presenta esta consulta

**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 (D.M.P). El Data Mart muestra información sobre el los porcentaje de las ventas de los productos con comparativas de tiempo 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
¿Se puede conocer Cuál es el monto de un determinado cliente en compras en un mes comparando con otros?	0	8.5
¿Se puede conocer Cuál es el Monto de ventas por tipo de producto?	0	8
¿Se puede conocer Cuál es el monto de un tipo de producto por proveedor en 3 meses?	4.5	9
¿Se puede conocer Cuál es el monto de ventas clientes en un semestre?	5.5	8.5
¿Se puede conocer Cuál es el monto que generan las productos en un año?	0	8
¿Se puede saber Cuál es la cantidad de productos vendidos por Tipo, vendedor, mes y semestre?	0	9
¿Se puede conocer Cuál es la cantidad de Productos que se han pagado al crédito y al contado?	0	8.5
¿Se puede conocer Cuál es el monto que generan las promociones que da un proveedor en un semestre?	6.5	8.5
¿Se puede conocer la Lista de Vendedores y el monto total de cada una de las ventas que han generado?	0	8.5
¿Se puede conocer Cuál es el porcentaje de las ventas de los productos según su tipo por mes?	0	8
<b>PROMEDIO</b>	<b>1.65</b>	<b>8.45</b>

Tabla 3.2. Evaluación de los indicadores de la hipótesis.

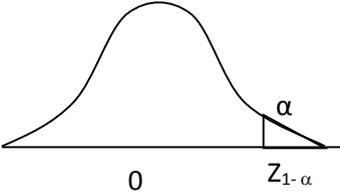
### 3.1.4. ANÁLISIS ESTADÍSTICO PARA LA PRUEBA PRESENCIAL DE LA HIPÓTESIS

**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 H <sub>0</sub>	VALOR CRÍTICO
Hipótesis Nula H <sub>0</sub> : $\mu_1 - \mu_2 = 0$ 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 H<sub>0</sub> 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*

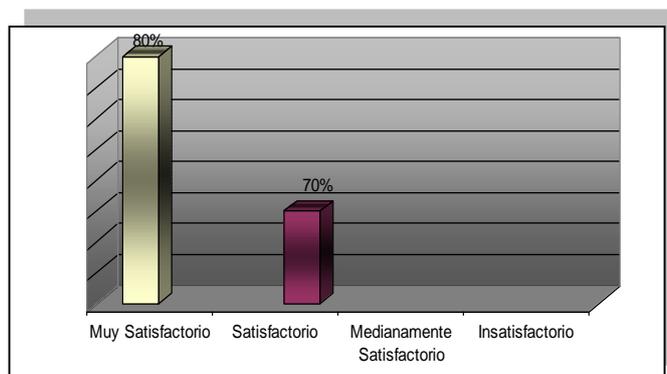
**CONCLUSIÓN:** La implementación de una Solución de Inteligencia de Negocios en la nube mejorará el proceso de Toma de Decisiones en el área de Ventas para la empresa Book Center S.A.C..

### 3.2. MANERA NO PRESENCIAL

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

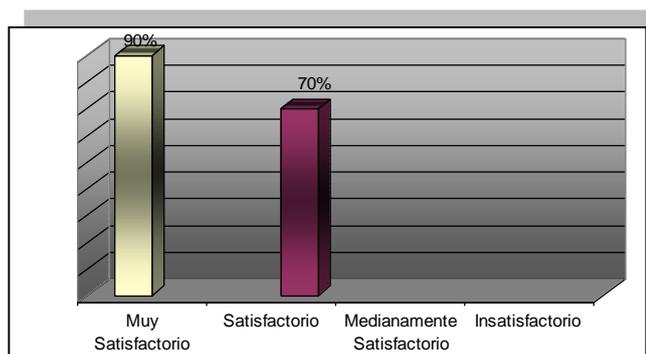
**CUESTIONARIO DIRIGIDO AL JEFE DE VENTAS E INVENTARIO y AL JEFE DE AREA INFORMATICA:**

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



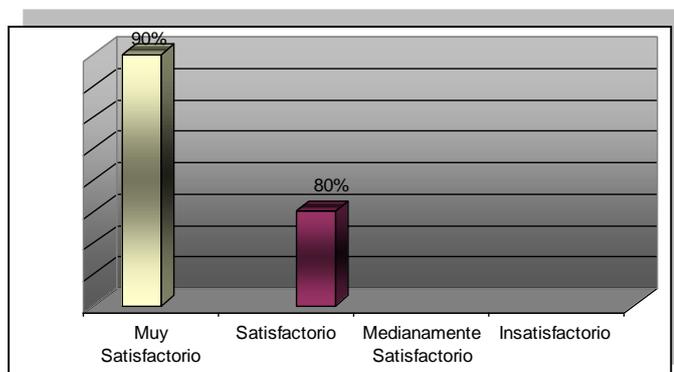
**Resultado:** El Jefe del Área de Ventas respondió que el ingreso a los datos que contiene el Data Mart se realizó muy satisfactoriamente (80%). El Jefe del Área de Informática respondió el ingreso a los datos que contiene el Data Mart se realizó satisfactoriamente (70%).

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



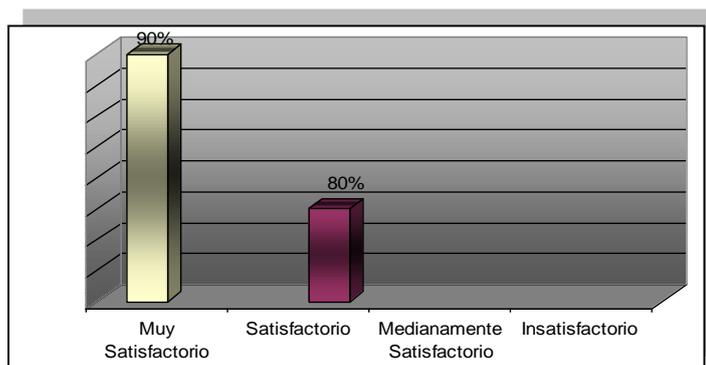
**Resultado:** El Jefe del Área de Ventas respondió que el Data Mart permitió realizar consultas que se consideran valiosas para el Área de Ventas de una manera muy satisfactoriamente (90%). El Jefe del Área de Informática respondió que el Data Mart permitió realizar consultas que se consideran valiosas 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 muy satisfactoriamente (80%).

**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 Data Mart presenta información valiosa y de forma dinámica para un mejor análisis de los datos que pueda dar soporte en la toma de decisiones.

## CONCLUSIONES

- Se elaboró el marco teórico respecto al trabajo evolutivo sobre la base de datos en la nube de Microsoft basada en soluciones de inteligencia de negocios.
- Se determinó que el alcance del proyecto está en el área de Ventas obteniendo 10 requerimientos que la gerencia necesita, 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.(Ver pág. 55 y 56)
- Al realizar el diseño se identificaron 1 tablas hecho y 6 dimensiones.
- La construcción del Datamart se realizó utilizando las herramientas de SQL Server 2012, como son SQL Server Data Tools (Business Intelligence Development) y Microsoft Azure.
- Se realizaron las pruebas sobre los requerimientos a fin de garantizar el correcto funcionamiento del Datamart, estas pruebas se realizaron utilizando el Power BI en Excel 2013 (Ver Pág. 107)

## RECOMENDACIONES

- 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.
- Es necesario realizar coordinaciones con los directivos y el personal para realizar capacitaciones sobre la importancia del manejo de las tecnologías y cultura analítica dentro de la empresa.
- 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.
- Para proyectos futuros se recomienda el uso de la metodología de Ralph Kimball y de SQL Server 2012 como herramientas de trabajo para el diseño de DataMart.
- Para la fase de análisis de requisitos se recomienda utilizar el análisis de molde de consulta para una mejor detección de las dimensiones y medidas candidatas a intervenir en el datamart.
- Para próximas versiones se recomienda el uso de dashboard que ofrece la Herramienta de Qlikview para un mejor control de los indicadores.

## REFERENCIAS

- Jesús Rodríguez-Gómez, 2011. *Toma de decisiones. Más allá de la intuición*, Primera edición. CreateSpace Independent Publishing Platform.
- Kendall, K. y Kendall, J., 2012, *Análisis y Diseño de Sistemas*, 8° Edición. Prentice-Hall. Hispanoamericana.
- Klein S., Roggero H. , 2010. *Pro SQL Azure*. Apress; 1 edition. Estados Unidos.
- Kimball, R. y Ross, M., 2013, *The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling*. 3° Edition, Editorial Wiley. Canadá.
- Laudon, K.C. y Laudon, J.P., 2013, *Management Information Systems: Managing the Digital Firm*. 13° Edición. Prentice-Hall PTR. Estados Unidos.
- Root R. y Mason C., 2012, *Pro SQL Server 2012 BI Solutions*. 1° Edición. Apress. Estados Unidos
- Brunetti Roberto, 2011. *Windows Azure Step by Step*. Microsoft Press; 1 edition. Estados Unidos.
- Villalobos E. y Constenia J., 2010, *Tecnologías de la Información y la Comunicación*. Instituto Politécnico Nacional. [En línea]. Disponible en <http://www.dcyt.ipn.mx/dcyt/quesonlastics.aspx#> [accesado el 25 de Marzo 2014]

# ANEXOS

## ANEXO A: CUESTIONARIOS

### CUESTIONARIO PARA ENTREVISTA N° 1: Administrador

#### 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: Área 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 INFORMATICA

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 las 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 Informática*

## ANEXO C

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

PREGUNTAS	VALORES										
	0	1	2	3	4	5	6	7	8	9	10
1. ¿De que 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 semestres 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 Informática*