

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



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

---

**SOLUCIÓN DE INTELIGENCIA DE NEGOCIOS QUE PERMITA UNA**  
**MEJOR INTEGRACIÓN DE DATOS Y SOPORTE EN LA TOMA DE**  
**DECISIONES EN EL PROCESO DE VENTAS DE LA EMPRESA**  
**LIBRERÍA CRISOL S.A.C. SUCURSAL TRUJILLO**

---

**Línea de Investigación:**

Gestión de Datos y de Información.

**Autores:**

**Br. LIONARD LEYVA HURTADO**

**Br. DENIS ANDERSON BERNABÉ BARRIOS**

**Asesor:**

**ING. AGUSTIN EDUARDO ULLON RAMIREZ**

**2019**

**“SOLUCIÓN DE INTELIGENCIA DE NEGOCIOS QUE PERMITA UNA MEJOR INTEGRACIÓN DE DATOS Y SOPORTE EN LA TOMA DE DECISIONES EN EL PROCESO DE VENTAS DE LA EMPRESA LIBRERÍA CRISOL S.A.C. SUCURSAL TRUJILLO”**

**Elaborado por:**

Br. Lionard Leyva Hurtado

Br. Denis Anderson Bernabé Barrios

**Aprobada por:**

---

**Ing. Edward Fernando Castillo Robles**  
**Presidente**  
**CIP: 192352**

---

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

---

**Ing. Percy Lucio Carranza Medina**  
**Vocal**  
**CIP: 149877**

---

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

## PRESENTACIÓN

Señores Miembros del Jurado:

De acuerdo a los requisitos descritos en el reglamento de grados y Títulos de la Universidad Privada Antenor Orrego y del Reglamento Interno de la Escuela Profesional de Ingeniería de Computación y Sistemas, se pone a disposición el presente Trabajo de Tesis titulado: **“SOLUCIÓN DE INTELIGENCIA DE NEGOCIOS QUE PERMITA UNA MEJOR INTEGRACIÓN DE DATOS Y SOPORTE EN LA TOMA DE DECISIONES EN EL PROCESO DE VENTAS DE LA EMPRESA LIBRERÍA CRISOL S.A.C. SUCURSAL TRUJILLO”** con el fin de alcanzar el título profesional de Ing. de Computación y Sistemas.

El trabajo de tesis se desarrolló bajo los lineamientos implementados por la Facultad de Ingeniería, la Escuela Profesional de Ingeniería de Computación y Sistemas y la información recibida y adquirida en nuestra formación académica dentro de la universidad, y además de las consultas realizadas a las diferentes fuentes bibliográficas.

Br. Lionard Leyva Hurtado

Br. Denis Anderson Bernabé Barrios

## **DEDICATORIA**

La presente Tesis está *dedicada* a Dios, ya que gracias a Él he logrado concluir mi carrera.

A mi madre Tilha Hurtado Arbildo, porque siempre estuvo a mi lado brindándome su apoyo y sus consejos para hacer de mí una mejor persona.

A mi hermano Jherson Lee Leyva Hurtado, tío Goodvasolenco Hurtado Arbildo, Abuelo Fabio Hurtado Crisóstomo QEPD.

**Br. Lionard Leyva Hurtado**

A mis padres por siempre apoyarme en todo momento, sin ellos no hubiera logrado lo que soy, a mi pareja porque siempre estuvo motivándome para sobresalir y ser mejor cada día y a mi amada hija quien me da muchas alegrías en todo momento, a ellos les debo todo.

**Br. Denis Anderson Bernabé Barrios**

## **AGRADECIMIENTO**

Un Agradecimiento muy especial a las personas que nos apoyaron dentro de la empresa Librería Crisol S.A.C. sucursal Trujillo, y nos brindaron la información necesaria para conocer más sobre la problemática dentro del proceso de ventas, brindándonos acceso a los datos que se necesitaron para el desarrollo del presente trabajo de tesis.

Un agradecimiento a nuestro asesor que nos dio el tiempo necesario con su asesoría para el desarrollo del trabajo de tesis.

También un agradecimiento a los amigos y familiares que siempre estuvieron pendiente de nuestro avance para el desarrollo del trabajo.

Muy agradecidos con todos.

**Los autores.**

**RESUMEN**  
**“SOLUCIÓN DE INTELIGENCIA DE NEGOCIOS QUE PERMITA UNA  
MEJOR INTEGRACIÓN DE DATOS Y SOPORTE EN LA TOMA DE  
DECISIONES EN EL PROCESO DE VENTAS DE LA EMPRESA  
LIBRERÍA CRISOL S.A.C. SUCURSAL TRUJILLO”**

**Por:**

Br. Lionard Leyva Hurtado

Br. Denis Anderson Bernabé Barrios

La competitividad en las empresas están estrechamente ligadas a la explotación y análisis de sus datos, es decir están enfocados dentro de proyectos de Inteligencia de Negocios, y así crear un ambiente propicio para tener una mejor transferencia de la información, que genere el conocimiento adecuado para mejorar todas las actividades permitiendo obtener, manipular y usar la información para generar conocimiento y obtener un soporte en la toma de decisiones dando un beneficios económico para las empresas.

En el presente trabajo se realiza un análisis la realidad en la que se encuentra el área de ventas de Crisol S.A.C. de la sucursal de Trujillo para mejorar la integración de datos y tener un mejor soporte en la toma de decisiones del área del problema. En este contexto se propone implementar una solución de inteligencia de negocios como una herramienta que se ajuste a los requerimientos reales de la empresa añadiendo el análisis dimensional y los indicadores necesarios para el aprovechamiento de la información y dando un mejor soporte a la toma de decisiones. Teniendo como objetivo general “Mejorar la integración de datos y dar un mejor soporte en la toma de decisiones para el proceso de ventas de la empresa Crisol S.A.C. sucursal Trujillo, usando la arquitectura tecnológica MS SQL Server, la metodología de Ralph Kimball y Qlik Sense”.

Para alcanzar este objetivo en el proyecto se realizó lo siguiente: Primero se realizó un análisis de todos los requerimientos para luego llegar a elaborar un modelo de datos dimensional, luego se implementó el modelo dimensional en MS SQL Server integrando las distintas fuentes de datos de la empresa con una solución ETL, para después implementar los cubos OLAP de acuerdo al esquema dimensional en SQL Server permitiendo tener una mejor vista de análisis y navegabilidad de la información. Finalmente se pasó a construir los reportes dinámicos en Qlik sense de acuerdo a las necesidades del usuario final.

## **ABSTRACT**

### **“SOLUCIÓN DE INTELIGENCIA DE NEGOCIOS QUE PERMITA UNA MEJOR INTEGRACIÓN DE DATOS Y SOPORTE EN LA TOMA DE DECISIONES EN EL PROCESO DE VENTAS DE LA EMPRESA LIBRERÍA CRISOL S.A.C. SUCURSAL TRUJILLO”**

**By:**

Br. Lionard Leyva Hurtado

Br. Denis Anderson Bernabé Barrios

The competitiveness in companies are closely linked to the exploitation and analysis of their data, that is, they are focused within Business Intelligence projects, and thus create an environment conducive to have a better transfer of information, which generates the right knowledge for improve all the activities allowing to obtain, manipulate and use the information to generate knowledge and obtain a support in the decision making giving an economic benefits for the companies.

In the present work an analysis is made of the reality in which the sales process of the company Crisol S.A.C. from the Trujillo branch to improve the integration of data and have better support in the decision making of the area of the problem. In this context, it is proposed to implement a business intelligence solution as a tool that adjusts to the real requirements of the company by adding dimensional analysis and the necessary indicators for the use of information and giving better support to decision making. With the general objective of "Improving data integration and giving better support in decision-making for the sales process of the company Crisol S.A.C. Trujillo branch, using the MS SQL Server technology architecture, the Ralph Kimball methodology and Qlik sense. "

To achieve this goal in the project, the following was done: First, the requirements analysis for the elaboration of a Data Modeling, then the dimensional model was implemented in MS SQL Server, integrating the different data sources of the company with an ETL solution, to then implement the OLAP cubes according to the dimensional schema in SQL Server allowing a better view of analysis and navigability of the information. Finally, dynamic reports were built in Qlik Sense according to the needs of the end user.

## ÍNDICE DE CONTENIDO

<b>PRESENTACIÓN</b> .....	ii
<b>DEDICATORIA</b> .....	iii
<b>AGRADECIMIENTO</b> .....	iv
<b>RESUMEN</b> .....	v
<b>ABSTRACT</b> .....	vi
<b>ÍNDICE DE CONTENIDO</b> .....	vii
<b>ÍNDICE DE FIGURAS</b> .....	x
<b>ÍNDICE DE TABLAS</b> .....	xi
<b>1. INTRODUCCIÓN</b> .....	12
1.1. Planteamiento del problema .....	12
1.2. Delimitación del problema .....	13
1.3. Características problemáticas.....	13
1.4. Definición del problema .....	15
1.5. Formulación del problema.....	15
1.6. Formulación de la hipótesis.....	15
1.7. Objetivos del estudio .....	15
1.8. Justificación del estudio.....	16
1.8.1. Importancia de la investigación.....	16
1.8.2. Viabilidad de la investigación.....	16
1.8.3. Aportes.....	17
<b>2. MARCO TEÓRICO</b> .....	18
2.1. ANTECEDENTES.....	18
2.2. BASES TEÓRICAS.....	20
2.3. METODOLOGÍA PARA EL DESARROLLO DEL PROYECTO.....	34
<b>3. MATERIALES Y MÉTODOS</b> .....	39
3.1. Material.....	39
3.1.1. Población.....	39
3.1.2. Muestra.....	39
3.1.3. Unidad de análisis.....	39
3.2. Método.....	39
3.2.1. Tipo de investigación.....	39

3.2.2. Diseño de Investigación.....	39
3.2.3. Variables de estudio y Operacionalización.....	40
3.2.4. Técnicas e instrumentos de recolección de datos.....	41
3.2.5. Técnicas de procesamiento y análisis de datos.....	42
<b>4. RESULTADOS: APLICACIÓN DE LA METODOLOGÍA.....</b>	<b>43</b>
ETAPA 1: PLANEACIÓN Y ADMINISTRACIÓN DEL PROYECTO.....	43
4.1.1. El negocio.....	43
4.1.2. Selección de la Estrategia de Implementación.....	45
4.1.3. Selección de la Metodología de Desarrollo .....	45
4.1.4. Selección del Ámbito de Implementación .....	45
4.1.5. Selección del enfoque arquitectónico.....	46
4.1.6. Desarrollo de un Programa y del Presupuesto del Proyecto .....	47
4.1.7. Desarrollo del escenario del uso empresarial .....	50
ETAPA 2: DETERMINACIÓN DE LOS REQUERIMIENTOS.....	51
4.2.1. Requerimientos del Propietario .....	51
4.2.2. Requerimientos del Usuario Final .....	52
4.2.3. Requerimientos no Funcionales .....	53
4.2.4. Análisis de los requerimientos .....	54
ETAPA 3: DISEÑO TÉCNICO DE LA ARQUITECTURA.....	59
4.3.1. Nivel de Datos (Base de Datos Transaccional) .....	59
4.3.2. Nivel Técnico .....	60
ETAPA 4: MODELADO DIMENSIONAL .....	60
4.4.1. Identificación de los Componentes del Modelo .....	60
4.4.2. Diagrama de la Tabla de Hechos .....	63
4.4.3. Esquema Estrella .....	68
ETAPA 5: DISEÑO FÍSICO .....	70
4.5.1. Determinación de las agregaciones .....	74
4.5.2. Construcción de las Tablas y la Base de Datos .....	74
ETAPA 6: PROCESO DE EXTRACCIÓN, TRANSFORMACIÓN Y CARGA DE DATOS .....	78
ETAPA 7: SELECCIÓN DE PRODUCTOS .....	95
7.1. Hardware .....	95
7.2. Software .....	95

ETAPA 8: ESPECIFICACIÓN DE LA APLICACIÓN DEL USUARIO FINAL	95
8.1. Estructura de Cubo .....	95
ETAPA 9: DESARROLLO DE LA APLICACIÓN DEL USUARIO FINAL	98
9.1. Elegir la herramienta para procesamiento analítico. ....	98
9.2. Lista los resúmenes de información requeridos por los usuarios. ....	98
9.3. Implementación de los reportes en Qlik Sense.....	99
<b>4. DISCUSIÓN DE RESULTADOS</b> .....	105
<b>5. CONCLUSIONES</b> .....	114
<b>6. RECOMENDACIONES</b> .....	115
<b>7. REFERENCIAS BIBLIOGRÁFICAS</b> .....	116
<b>ANEXOS</b> .....	118

## ÍNDICE DE FIGURAS

Figura N° 1. Esquema estrella	14
Figura N° 2. Esquema copo de nieve	15
Figura N° 3. Proceso ETL	17
Figura N° 4: Organigrama por tienda	20
Figura N° 5: Cronograma de actividades	39
Figura N° 6: Diagrama E-R de la base de datos	48
Figura N° 7: Enfoque arquitectónico	49
Figura N° 8: Diagrama de Tabla de Hechos	52
Figura N° 9: Tabla de Hechos	53
Figura N° 10: Jerarquía de los Atributos de la Dimensión Tiempo	54
Figura N°11: Jerarquía de los Atributos de la Dimensión Categoría	55
Figura N°12: Jerarquía de los Atributos de la Dimensión Cliente	55
Figura N° 13: Jerarquía de los Atributos de la Dimensión Presentación	56
Figura N° 14: Jerarquía de los Atributos de la Dimensión Autor	57
Figura N° 15: Jerarquía de los Atributos de la Dimensión Vendedor	57
Figura N° 16: Jerarquía de los Atributos de la Dimensión Libro	58
Figura N° 17: Esquema estrella del BI	58
Figura N° 18: Diseño Físico de la Base de Datos del BI	66
Figura 19: Workflow de los Pasos de Transformación	68
Figura 20: Diagrama Workflow con Restricciones de Precedencia.	69
Figura 21: ETL de la solución de BI	70

## ÍNDICE DE TABLAS

Tabla N° 1. Diagrama de investigación.....	29
Tabla N° 2: Operacionalización de las variables.....	29
Tabla N° 3: Cronograma de actividades.....	39
Tabla N° 4: Consultas y medidas .....	42
Tabla N° 5: Objetivos de las tablas de Hecho .....	54
Tabla N° 6: Dimensiones de la tabla de Hechos Ventas .....	51
Tabla N° 7: Elección de los hechos.....	51
Tabla N° 8: Tabla de dimensiones.....	53
Tabla N° 9: Tabla de Hechos ventas.....	53
Tabla N° 10: Detalle de Dimensión Tiempo Ventas.....	54
Tabla N° 11: Detalle de Dimensión Categoría.....	55
Tabla N° 12: Detalle de Dimensión Cliente.....	55
Tabla N° 13: Detalle de Dimensión Presentación .....	56
Tabla N° 14: Detalle de Dimensión Autor.....	56
Tabla N° 15: Detalle de Dimensión Libro.....	57
Tabla N° 16: Nombres estándares para la Tabla de Hecho y Dimensiones.....	59
Tabla N° 17: Nombres estándares para los atributos de las dimensiones.....	59
Tabla N° 18: Nombres estándares para los atributos de las Tablas de Hechos.....	60
Tabla 19: Tipo de Dato para dim_tiempo.....	60
Tabla 20: Tipo de Dato para dim_cliente.....	60

# **1. INTRODUCCIÓN**

## **1.1. PLANTEAMIENTO DEL PROBLEMA**

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

El valor fundamental de una empresa reside en su capacidad para superar a sus competidores en términos de costos y calidad de los productos y servicios que ofrece. Es por ello que las empresas dependen considerablemente de la habilidad para adaptarse a las exigencias de su mercado y de su entorno. Una de las principales variables que considera cualquier empresa exitosa es la innovación y la gestión de su conocimiento e información, de esta manera podría garantizar su continuidad en el mercado.

Para las empresas la competitividad está estrechamente ligada a la explotación de sus datos es decir dentro de proyectos de Inteligencia de Negocios y, es así que en una empresa moderna debe crear un ambiente propicio para tener una mejor transferencia de la información y que genere el conocimiento adecuado para mejorar todas las actividades de sus trabajadores y también en todas las instancias de la estructura de la empresa. Esto permitirá obtener, manipular y usar la información para generar conocimiento y luego este sea el factor que determine si la empresa obtiene o no beneficios económicos.

Gracias a la capacidad de explotar la información, los sistemas de soporte de decisiones enfocados en la inteligencia de negocios van a permitir a las empresas a adquirir un mejor entendimiento de ellas mismas. Permitiendo, con la integración y análisis de datos, una manera más sencilla y entender el desempeño de la empresa, plantear escenarios a futuro, dando de esta manera un mejor apoyo en la toma de decisiones.

La Empresa Crisol es una empresa que se dedica a la venta minorista de libros y

artículos de entretenimiento cultural. También se puede encontrar gran variedad de productos como CD y DVD musicales, películas, rompecabezas, juegos educativos, entre otros. En la actualidad, Librerías Crisol cuenta con 29 tiendas a nivel nacional entre ellas la sucursal de Trujillo. La empresa tiene ya 16 años en este rubro y se ha convertido en un espejo para otras muchas empresas por su sistema de gestión de las necesidades del cliente, otorgándole a éste, la oportunidad de conseguir cultura al mejor precio.

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

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

## **1.2. DELIMITACIÓN DEL PROBLEMA**

El proyecto se realizará analizando la realidad en la que se encuentra el proceso de ventas de la empresa Crisol S.A.C. de la sucursal de Trujillo para mejorar la integración de datos y tener un mejor soporte en la toma de decisiones del área del problema.

## **1.3. CARACTERÍSTICAS PROBLEMÁTICAS**

- ✓ **Tiempo Perdido en realizar reportes.** Hay demasiada demora al realizar reportes esto incluye el tiempo que toma recopilar información, construir formulas, gráficos, verificar información y compartir el archivo.

- ✓ **Estándares no definidos y Errores Humanos en reportes.** Hay redundancia al hacer uso de diferentes fórmulas o conceptos para explicar una misma situación, al no haber un estándar definido los diseños de algunas planillas son difíciles de leer o contienen errores humanos.
- ✓ **Existe un reprocesamiento de la información.** Se reprocesa información proporcionada por los sistemas actuales y esto se ha convertido en un problema para los tomadores de decisiones ocasionando que se organice la información por separado y no tengan una sola visión de lo que sucede.
- ✓ **Riesgos de Dependencia.** Existe información que solo algunos Individuos específicos en la empresa realizan estas tareas, haciéndolas en hojas de cálculo que solo ellos saben usar.
- ✓ **Duplicación de Información.** Existen reportes que tienen un mismo fin y suelen ser creados más de una vez por diferentes usuarios.
- ✓ **Toma de Decisiones erróneas.** El área de sistemas de la empresa no explota adecuadamente la información que provienen de sus fuentes transaccionales, por lo que solo entregan reportes difíciles de entender para un proceso de toma de decisiones y muchas veces estas decisiones son tomadas sin haber hecho los cálculos correctos.
- ✓ **Costos altos por problemas de tomar decisiones fuera de tiempo.** La empresa no cuenta con soluciones tecnológicas que den estadísticas y reportes en línea y al no contar con estas soluciones generan una pérdida económica por no saber si están funcionando adecuadamente.
- ✓ **Falta de integración de datos de clientes:** Existe datos de los clientes guardados en hojas de Excel que no son llevados y consultados con los registrados en la base de datos, y en algunos casos estos clientes han realizado compras en la tienda.

En este contexto se propone implementar una solución de inteligencia de negocios como una herramienta que se ajuste a los requerimientos reales de la empresa añadiendo el análisis dimensional e integración de datos, dando un mejor aprovechamiento de la información y un mejor soporte a la toma de decisiones.

#### **1.4. DEFINICIÓN DEL PROBLEMA**

La falta de un mejor análisis de información e integración de datos en el proceso de ventas de la Librería Crisol S.A.C. Sucursal Trujillo no le permite a la empresa tener un mejor soporte en la toma de decisiones.

#### **1.5.FORMULACIÓN DEL PROBLEMA**

¿Cómo se podría mejorar la integración de datos y el soporte en la toma de decisiones para el proceso de ventas de la empresa Crisol S.A.C. sucursal Trujillo?, utilizando tecnologías de información?

#### **1.6. FORMULACIÓN DEL HIPÓTESIS**

El Desarrollo de una Solución de Inteligencia de negocios usando la arquitectura tecnológica de Microsoft y la metodología de Kimball mejora la Integración de Datos y el análisis de información para un mejor soporte en la toma de decisiones en el proceso de Ventas de la empresa Librería Crisol S.A.C. sucursal Trujillo.

#### **1.7. OBJETIVOS DEL ESTUDIO**

El **Objetivo general** es:

Mejorar la integración de datos y dar un mejor soporte en la toma de decisiones para el proceso de ventas de la empresa Crisol S.A.C. sucursal Trujillo, usando la arquitectura tecnológica MS SQL Server, la metodología de Ralph Kimball y Qlik Sense.

Los **objetivos específicos** son los siguientes:

- ✓ Definir los Stakeholders y obtener los requerimientos a través de entrevistas al área crítica del negocio (Ventas).
- ✓ Realizar el análisis de requerimientos para la elaboración de un Modelado de Datos que de soporte a la toma de decisiones en la organización.
- ✓ Implementar el modelo dimensional en Microsoft Azure
- ✓ Integrar las distintas fuentes de datos de la empresa en una solución ETL usando SQL Server 2014.
- ✓ Implementar los cubos OLAP de acuerdo al esquema dimensional e

permitiendo tener una mejor vista de análisis y navegabilidad de la información

- ✓ Construir los reportes dinámicos en Qlik Sense de acuerdo a las necesidades del usuario final.

## **1.8. JUSTIFICACIÓN DEL ESTUDIO**

### **1.8.1. Importancia de la investigación**

- ✓ El presente trabajo de inteligencia de negocios representará un gran apoyo al momento analizar la información de los informes y reportes que genere a fin de que la gerencia pueda definir mejores estrategias y decisiones.
- ✓ La solución de inteligencia de negocios que se realizará permitirá recolectar, integrar, depurar y transformar los datos que se encuentren desestructurados de la base de datos transaccional y en otros archivos de la organización evitando la redundancia de datos.
- ✓ Tener una buena consistencia en los datos es un factor muy importante para cualquier empresa sea empresa grande o pequeña ya que de acuerdo a esto se podrá tener una mejor confiabilidad y disponibilidad de la información que se procesara para la toma de decisiones.
- ✓ Los datos procesados en el BI podrán brindar una realidad más clara a la situación actual del negocio.
- ✓ Los datos estadísticos que se generaran de la herramienta de inteligencia de negocios permitirá determinar medidas de cambio y mejora provenientes de los sistemas transaccionales.

### **1.8.2. Viabilidad de la investigación**

- ✓ La investigación es factible porque se cuenta con el acceso a los datos y a la información que maneja la empresa y también brindará un mejor entendimiento de los procesos siendo primordial esto para el desarrollo de la empresa.
- ✓ Es viable porque el equipo de trabajo cuenta con el conocimiento para poder realizar el proyecto que se ha propuesto contando con la experiencia en conocimientos del asesor en el tema.

- ✓ El proyecto tiene como objetivo implementar un modelo multidimensional que ayude en el soporte de las tomas de decisiones generando la reducción del tiempo en obtener informes, también se podrá tener una mejor personalización de realizar los informes de acuerdo a los requerimientos de la empresa.
- ✓ Es viable porque el trabajo de BI permitirá tener la información integrada y disponible para el tomador de decisiones, generando la optimización de costos y crecimiento de la empresa por ello la inversión que se realizará para el proyecto es justificada.

### **1.8.3. Aportes**

El desarrollo de esta investigación generará aportes a la empresa, entre los cuales tenemos:

- ✓ Información correcta y oportuna para que la empresa pueda tomar decisiones acertadas.
- ✓ Confianza en los tomadores de decisiones por tener un mejor análisis de información.
- ✓ Aporte en conocimientos adquiridos en este estudio que permitirá desarrollar proyectos similares en alguna otra institución, facilitando la implementación de un sistema de soporte de decisiones basados en Inteligencia de negocios.

## 2. MARCO TEÓRICO

### 2.1. ANTECEDENTES

- Sánchez (2014); en su investigación titulada “Modelo de inteligencia de negocio para la toma de decisiones en la empresa San Roque S.A.”, El presente trabajo presenta una propuesta de Modelo de Inteligencia de Negocios que asegure una mejora en el proceso de ventas y que permita transformar información clave del negocio en acciones concretas traduciéndose en beneficios tangibles que les proporcione una ventaja competitiva a los tomadores de decisiones. Para lograrlo se identificaron y compararon las características más relevantes de las principales metodologías, luego se propuso un Modelo de Inteligencia de Negocios que cuenta con las siguientes fases: planificación, modelo del negocio, análisis, diseño, construcción e implementación. Para el despliegue del modelo se utilizó la plataforma BI Pentaho EE versión 4.8 y los componentes de Big Data: Hadoop y Hive. Así se implementó el Modelo de Inteligencia de Negocio en el área de ventas de la empresa San Roque S.A. Se analizaron los datos obtenidos a través de la implementación del modelo, con lo que se demostró mejoras en los indicadores definidos para su evaluación. (Sánchez Guevara, 2014)
- Medina y Calzado (2008), en su investigación titulada “Data Mart para el área de ventas para la empresa Mi Mercado S.A.C. utilizando la metodología de Ralph Kimball y SQL Server como herramienta tecnológica”, para el desarrollo de este trabajo se propuso la generación de reportes de manera más eficiente y rápida apoyados en el data mart. Llegando a los siguientes resultados el Data Mart construyéndose para brindar un mejor análisis de los datos en el área de Ventas. Este análisis de la información se da de forma dinámica, permitiendo que los datos sean accesibles al usuario final sin tener que volver a analizarlo. El principal aporte de este trabajo de investigación es que solución construida mostró un adecuado análisis de datos en la cual permitió apreciar información de forma dinámica mejorando el análisis del usuario final. (Medina & Calzado, 2008)

- Benites y López (2015); en su investigación titulada “Solución de Inteligencia de Negocio para Empresas de Servicio de Importación y exportación de Calzado del departamento La Libertad”, el objetivo que persigues este trabajo es de Implementar una solución de Inteligencia de Negocio para la empresa Carubi SAC dedicada al servicio de Importación y Exportación de Calzado del departamento La Libertad en donde existe muchas fallas y demoras en la toma de decisiones, Esto ha ocasionado pérdidas considerables en sus ganancias. Para elaborar la solución se propuso: Elaborar un marco ontológico respecto al trabajo evolutivo sobre la base de datos basada en la Refactorización de Base de Datos, definieron los usuarios de la solución obteniendo sus requerimientos, para luego realizar un análisis de cada requerimiento para obtener un Refactoring del modelo dimensional que se tiene del área de estudio. Dar solución a los problemas de Integración de datos usando las herramientas ETL. Implementar los cubos y reportes dinámicos de acuerdo a las necesidades del cliente. Concluyendo en la implementación de 2 tablas de hechos y 6 dimensiones en SQL Server mejorando de esta manera la rapidez en la obtención de reportes. (Benites Robles & López Girón, 2015)
- Rodríguez y Pereda (2013) en su investigación titulada “Implementación de un Dashboard para la toma de decisiones estratégicas en la unidad de negocio de producción de huevo incubable de la empresa Avícola Santa Fe S.A.C. Usando Tecnologías Oracle Business Intelligence”, en el presente trabajo se tiene como objetivo implementar un cuadro de mando, que será usado por la Empresa y para lograr esta solución se definió cada indicador necesario basado en el análisis de cada requerimiento y de su base de datos de donde obtuvo información relevante. Para la implementación se basó en la metodología de Ralph Kimball y el software de Business Intelligence de Oracle, que permitió implementar indicadores. La solución logró eliminar muchos de los trabajos manuales que perjudicaban la toma de decisiones en la empresa. (Rodriguez Torres & Pereda Morales, 2013)
- Gastañaudi y Tamayo (2016), en su investigación titulada “Inteligencia de negocios para la dinamización en la toma de decisiones en la gestión contable y

presupuestal de la Municipalidad de Víctor Larco Herrera”, tiene como finalidad dinamizar la toma de decisiones en la institución, reduciendo el tiempo en la generación de reportes en la gestión Contable y Presupuestal, reduciendo el costo de Horas -Hombre en la generación de reportes e Incrementar el nivel de satisfacción de los usuarios de la solución. Propuesta que concluye con una mejora de 85.07% en el tiempo promedio que se emplea para la generación de reportes además de un decremento en los costos en la generación de reportes. Con un nivel de satisfacción de usuario del 93.4% dinamización de esta forma la toma de decisiones en el área. (Gastañaudi Caballero & Tamayo Avalos, 2016)

## **2.2. BASES TEÓRICAS**

### **2.2.1. INTELIGENCIA DE NEGOCIOS:**

Una interesante definición para inteligencia de negocios o BI, por sus siglas en inglés, según el Data Warehouse Institute, lo define como la combinación de tecnología, herramientas y procesos que me permiten transformar mis datos almacenados en información, esta información en conocimiento y este conocimiento dirigido a un plan o una estrategia comercial. La inteligencia de negocios debe ser parte de la estrategia empresarial, esta le permite optimizar la utilización de recursos, monitorear el cumplimiento de los objetivos de la empresa y la capacidad de tomar buenas decisiones para así obtener mejores resultados. (Oracle, 2016)

#### **Problemas que enfrentan las empresas hoy en día:**

- Tenemos datos, pero carecemos de información – Es importante almacenar los datos de clientes, empleados, departamentos, compras, ventas, entre otros en aplicaciones, sistemas financieros o fuentes de datos. Si queremos que nuestra empresa tenga mayor ventaja sobre la competencia esta gestión no es suficiente. Necesitamos profundizar el nivel de conocimiento de nuestros clientes, empleados, operaciones para así, tener la capacidad de encontrar patrones de comportamiento, monitorear, rastrear, entender, administrar y contestar aquellas

interrogantes que me permitan maximizar el rendimiento de nuestra empresa (Oracle, 2016)

- Fragmentación – Poseen aplicaciones independientes a través de todos los departamentos, pero se carece de una visión global de la empresa. Tal vez por la incapacidad de las herramientas de BI de integrar fuentes de datos heterogéneas. Esto limita a la empresa a tomar decisiones importantes sin tener todos los elementos imprescindibles a la mano. Esta fragmentación conduce a lo que se llama diferentes versiones de la verdad. Los gerenciales solicitan informes a los distintos departamentos obteniendo diferentes resultados del mismo informe. La tarea ya no es solo crear el informe sino justificar de donde y en qué condiciones se utilizaron para la creación de este informe. Si el gerencial decide agregar una nueva variable a esta ecuación, recrear este informe puede conllevar un esfuerzo de semanas. (Oracle, 2016)
- Manipulación manual – La necesidad de generar análisis de negocios e informes nos ha llevado a utilizar herramientas de BI y/o de reportes que no son las más confiables. Esta práctica conlleva la exportación de datos a distintas herramientas que resultan en un proceso lento, costoso, duplicación de trabajo, poca confiabilidad en los informes, propenso a errores y sujetos a la interpretación individual. (Oracle, 2016)
- Poca agilidad – Debido a la carencia de información, la fragmentación y la manipulación manual me mantiene en un nivel de rendimiento bajo. Como dice el dicho: “Justo cuando me aprendí las respuestas me cambiaron las preguntas...”. Necesitamos de una herramienta lo suficientemente ágil que se ajuste a las necesidades del negocio. (Oracle, 2016)

Según el autor (Cano, 2007) El objetivo básico de la Business Intelligence es apoyar de forma sostenible y continuada a las organizaciones para mejorar su competitividad, facilitando la información necesaria para la toma de decisiones. El primero que acuñó el término fue Howard Dresner que, cuando era consultor de Gartner, popularizó Business Intelligence o BI

como un término paraguas para describir un conjunto de conceptos y métodos que mejoraran la toma de decisiones, utilizando información sobre qué había sucedido (hechos). Mediante el uso de tecnologías y las metodologías de Business Intelligence pretendemos convertir datos en información y a partir de la información ser capaces de descubrir conocimiento. (Cano, 2007)

### **2.2.2. VENTAJAS DE LA INTELIGENCIA DE NEGOCIOS:**

La adquisición de un sistema de BI debe proporcionar las siguientes ventajas: (Rio, 2006)

- ✓ Una plataforma de tecnología integrada que se añadiría a las inversiones ya realizadas por una organización, para proporcionar información de alta calidad a cada computadora o servidor de cada departamento de una empresa, añadiendo valor en cada paso del proceso y proporcionando una versión única de la realidad.
- ✓ Acceso amplio y potenciado a las capacidades de análisis ya conocidas que ayudan a conocer el pasado de una organización para controlar y comunicar el presente y predecir el futuro con fiabilidad.
- ✓ Interfaces de usuario personalizadas que “se adapten a cada tarea”, concebida y diseñadas para todos los niveles de experiencia y patrones de uso de los usuarios de la información (como expertos, ejecutivos, gerentes, tecnólogos, etc.).
- ✓ Una gama de soluciones para satisfacer la demanda de información y generación de informes de diferentes sectores de actividad (servicios financieros, fabricación, telecomunicaciones, salud, etc.) y de toda la empresa (gestión del rendimiento empresarial, inteligencia de clientes, inteligencia financiera, gestión del capital humano, inteligencia de la cadena de suministro, etc.)

### **2.2.3. OLAP**

Los sistemas OLAP son bases de datos orientadas al procesamiento analítico. Este análisis suele implicar, generalmente, la lectura de grandes cantidades de

datos para llegar a extraer algún tipo de información útil: tendencias de ventas, patrones de comportamiento de los consumidores, elaboración de informes complejos, etc. Este sistema es típico de los datamarts. (Sinnexus, 2014)

- ✓ El acceso a los datos suele ser de sólo lectura. La acción más común es la consulta, con muy pocas inserciones, actualizaciones o eliminaciones. (Sinnexus, 2014)
- ✓ Los datos se estructuran según las áreas de negocio, y los formatos de los datos están integrados de manera uniforme en toda la organización. (Sinnexus, 2014)
- ✓ El historial de datos es a largo plazo, normalmente de dos a cinco años. (Sinnexus, 2014)
- ✓ Las bases de datos OLAP se suelen alimentar de información procedente de los sistemas operacionales existentes, mediante un proceso de extracción, transformación y carga (ETL). (Sinnexus, 2014)

El procesamiento analítico en línea permite obtener acceso a datos organizados y agregados de orígenes de datos empresariales, organiza subconjuntos de datos con una estructura multidimensional de manera que represente un significado especial o responda a una pregunta en particular (Roussel, 2006)

Estas herramientas soportan el análisis interactivo de la información de resumen, soportando muchas tareas de agrupación de datos que no pueden realizarse empleando las facilidades básicas de agregación y agrupamiento. (Silberschatz , Korth , & Sudarshan, 2006)

**Tradicionalmente, este sistema se clasifica según las siguientes**

**categorías:**

- ROLAP. Implementación que almacena los datos en un motor relacional. Típicamente, los datos son detallados, evitando las agregaciones y las tablas se encuentran normalizadas. (Sinnexus, 2014)
- MOLAP. Esta implementación almacena los datos en una base de datos multidimensional. Para optimizar los tiempos de respuesta, el resumen de la información es usualmente calculado por adelantado. (Sinnexus, 2014)
- HOLAP. Almacena algunos datos en un motor relacional y otros en una base de datos multidimensional (Sinnexus, 2014)

**A. Esquema estrella:** con este esquema tendremos una visión multidimensional de un proceso que medimos a través de unas métricas. A nivel de diseño, consiste en una tabla de hechos (lo que en los libros encontraremos como **fact table**) en el centro para el hecho objeto de análisis y una o varias tablas de dimensión (dimensión table) por cada dimensión de análisis que participa de la descripción de ese hecho. En la tabla de hecho encontramos los atributos destinados a medir (cuantificar) el hecho: sus métricas.

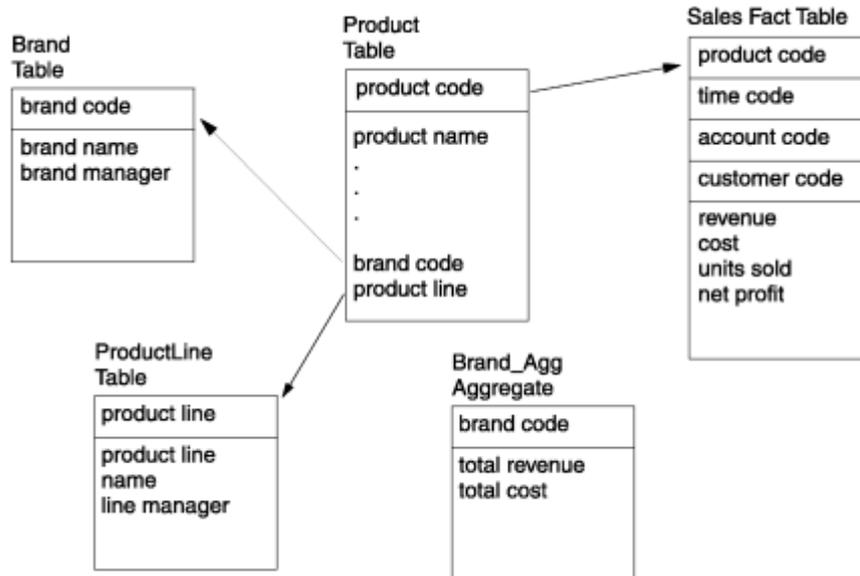
Mientras, en las tablas de dimensión, los atributos se destinan a elementos de nivel (que representan los distintos niveles de las jerarquías de dimensión) y a atributos de dimensión (encargados de la descripción de estos elementos de nivel). En el esquema en estrella la tabla de hechos es la única tabla del esquema que tiene múltiples join que la conectan con otras tablas (foreign keys hacia otras tablas). El resto de tablas del esquema (tablas de dimensión) únicamente hacen join con esta tabla de hechos. Las tablas de dimensión se encuentran además totalmente desnormalizadas, es decir, toda la información referente a una dimensión se almacena en la misma tabla. (Curto, 2007)



**Figura N° 1 . Esquema estrella**

Fuente (IBM, 2017)

**B. Esquema copo de nieve:** El esquema en copo de nieve es un esquema de representación derivado del esquema en estrella, en el que las tablas de dimensión se normalizan en múltiples tablas. Por esta razón, la tabla de hechos deja de ser la única tabla del esquema que se relaciona con otras tablas, y aparecen nuevas join gracias a que las dimensiones de análisis se representan ahora en tablas de dimensión normalizadas. En la estructura dimensional normalizada, la tabla que representa el nivel base de la dimensión es la que hace join directamente con la tabla de hechos. La diferencia entre ambos esquemas (star y snowflake) reside entonces en la estructura de las tablas de dimensión. Para conseguir un esquema en copo de nieve se ha de tomar un esquema en estrella y conservar la tabla de hechos, centrándose únicamente en el modelado de las tablas de dimensión, que, si bien en el esquema en estrella se encontraban totalmente de normalizadas, ahora se dividen en subtablas tras un proceso de normalización. Es posible distinguir dos tipos de esquemas en copo de nieve, un snowflake completo (en el que todas las tablas de dimensión en el esquema en estrella aparecen ahora normalizadas en el snowflake) o un snowflake parcial (sólo se lleva a cabo la normalización de algunas de ellas). (IBM, 2017)



*Figura N° 2. Esquema copo de nieve*

Fuente (IBM, 2017)

#### 2.2.4. DATAWAREHOUSE

Es el proceso de extraer datos de distintas aplicaciones (internas y externas), para que una vez depurados y especialmente estructurados sean almacenados en un depósito de datos consolidado para el análisis del negocio. Requiere una combinación de metodologías, técnicas, hardware y los componentes de software que proporcionan en conjunto la infraestructura para soportar el proceso de información (Stackowiak, Rayman, & Rick, 2007)

La estructura que se defina debe reflejar las necesidades y características del negocio, sus departamentos, equipos de trabajo y directivos, esto permitirá responder a interrogantes generados al tratar de tomar las decisiones (Ian H. witten, 2000) y con el tiempo se va convirtiendo en la memoria corporativa (Wang, 2009) ; describiendo el pasado y el presente de la empresa. Data Warehouse desglosa, resume, ordena y compara, pero no descubre, ni predice. (Flores, 2004) .

Para la construcción de un Data Warehouse se establecen tres etapas; la primera está dedicada a examinar el esquema Entidad Relación de la base de datos operacional, generando los esquemas multidimensionales candidatos.

La segunda etapa, consiste en recoger los requisitos de usuario por medio de

entrevistas, para obtener información acerca de las necesidades de análisis de estos, y la tercera etapa, contrasta la información obtenida en la segunda etapa, con los esquemas multidimensional candidatos formados en la primera etapa generando así, una solución que refleja los requisitos de usuario (sanchez, 2008)

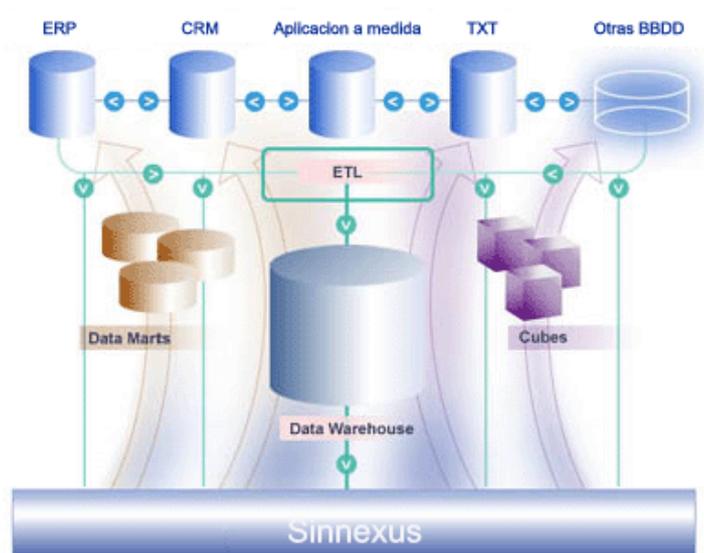
Por otra parte implementar una solución de este tipo, ocasiona un costo que no todas las organizaciones están dispuestas a pagar (debido a sus capacidades de inversión), es por eso que los promotores del proyecto dentro de la empresa deben persuadir a los directivos y compañeros de trabajo, una buena alternativa de hacerlo es mediante el uso de técnicas administrativas, que permitan conocer a los directivos como se puede establecer el retorno de la inversión del proyecto equiparando inversión contra beneficios. (Arturo L, 2001)

#### **2.2.5. DATAMART**

Un Datamart es una base de datos departamental, especializada en el almacenamiento de los datos de un área de negocio específica. Se caracteriza por disponer la estructura óptima de datos para analizar la información al detalle desde todas las perspectivas que afecten a los procesos de dicho departamento. Un datamart puede ser alimentado desde los datos de un Datawarehouse, o integrar por sí mismo un compendio de distintas fuentes de información. Por tanto, para crear el datamart de un área funcional de la empresa es preciso encontrar la estructura óptima para el análisis de su información, estructura que puede estar montada sobre una base de datos OLTP, como el propio Datawarehouse, o sobre una base de datos OLAP. La designación de una u otra dependerá de los datos, los requisitos y las características específicas de cada departamento. (Sinnexus, 2014)

## 2.2.6. ETL

- **Extracción:** En este proceso se extraerán los datos desde múltiples fuentes y se los integra. Primero se identificarán las fuentes, luego se seleccionará las fuentes a usar, y finalmente se unen los datos de las fuentes para ser extraídos. (Sinnexus, 2014)
- **Transformación:** Conjunto de técnicas y herramientas para cambiar los datos extraídos de las diferentes fuentes, es decir quitar y añadir información, corregir si hay errores y uniformizar los datos (asignar un formato estándar); en otras palabras, refinarlo y agregarles calidad. (Sinnexus, 2014)
- En este proceso se Suben los datos transformados al Data Warehouse o almacén de datos a través de métodos, ya que esto controlará la Actualización de datos de acuerdo al tiempo programado para Su elaboración. Finalizado el proceso de Carga se debería tener en el Data Warehouse datos de calidad, los cuales servirán de soporte para la toma de decisiones. (Sinnexus, 2014)



*Figura N° 3. Proceso ETL*

Fuente (Sinnexus, 2014)

### **2.2.7. SISTEMA DE SOPORTE A LAS DECISIONES DSS**

En la empresa no todas las decisiones son de carácter recurrente, sino que algunas se presentan muy pocas veces o incluso una sola vez. Los dss son instrumentos para abordar problemas de definición o estructuración menos precisa, y de carácter más esporádico. Los sistemas de apoyo a la decisión ayudan a los directivos que deben tomar decisiones no estructuradas. Una decisión se considera no estructurada si no existen procedimientos claros para tomarla y tampoco es posible identificar, con antelación, todos los factores que deben considerarse en la decisión. Hay que decir que todos los sistemas de información sirven de apoyo a la toma de decisiones, aunque sea de forma indirecta. Los dss son el tipo de sistemas de información desarrollados expresamente para servir de apoyo en el proceso de toma de decisiones. Estos sistemas facilitan un diálogo con el usuario que está considerando soluciones alternativas a un problema, y el sistema proporciona modelos construidos para la presentación de la información y acceso a bases de datos. Los sistemas de apoyo a la decisión son interactivos y su objetivo es la ampliación del razonamiento humano en la resolución de problemas particulares de toma de decisiones no estructuradas. Este tipo de sistemas se centra en los procesos de decisión y deberá proporcionar de forma fácil, rápida y exacta hechos importantes relacionados con la decisión a tomar y facilitando el acceso interactivo a medios de tratamiento que se utilizan creativamente y que permiten explorar las distintas posibilidades, suministrando las informaciones necesarias para responder a los problemas planteados. Al utilizar un dss, un directivo considerará un número posible de escenarios con el razonamiento: ¿Qué pasaría sí?, por ejemplo, un directivo que estuviera estudiando fijar el precio de un nuevo producto puede utilizar un sistema de apoyo a la decisión en el área marketing. El sistema contendrá un modelo que relacione varios factores, tales como el precio del producto, el coste de los materiales, los gastos de promoción, que afecten a la estimación de beneficios de las ventas del producto para un período de unos cinco años. Variando el precio del

producto en el modelo, el directivo puede comparar resultados previstos y, entonces, seleccionar un precio (Alcami, Carañana, & Herrando, 2009).

### 2.2.8. FUNCIÓN Y CARACTERÍSTICAS DEL DSS

El DSS es una de las herramientas más emblemáticas del Business Intelligence ya que, entre otras propiedades, permiten resolver gran parte de las limitaciones de los programas de gestión. Estas son algunas de sus características principales:

Informes dinámicos, flexibles e interactivos, de manera que el usuario no tenga que ceñirse a los listados predefinidos que se configuraron en el momento de la implantación, y que no siempre responden a sus dudas reales.

No requiere conocimientos técnicos. Un usuario no técnico puede crear nuevos gráficos e informes y navegar entre ellos, haciendo drag&drop o drill through. Por tanto, para examinar la información disponible o crear nuevas métricas no es imprescindible buscar auxilio en el departamento de informática (Alcami, Carañana, & Herrando, 2009).

- **Integración entre todos los sistemas/departamentos de la compañía.**  
El proceso de ETL previo a la implantación de un Sistema de Soporte a la Decisión garantiza la calidad y la integración de los datos entre las diferentes unidades de la empresa. Existe lo que se llama: integridad referencial absoluta.
- **Drag&drop:** Es una expresión informática que se refiere a la acción de mover con el ratón objetos de una ventana a otra o entre partes de una misma ventana.
- **Rapidez en el tiempo de respuesta:** ya que la base de datos subyacente suele ser un Datawarehouse corporativo o un datamart, con modelos de datos en estrella o copo de nieve. Este tipo de bases de datos están optimizadas para el análisis de grandes volúmenes de información.

- **Drill Through:** Es el análisis de los datos que va desde un cubo OLAP en la base de datos relacional.
- **Disponibilidad de información histórica:** En estos sistemas está a la orden del día comparar los datos actuales con información de otros períodos históricos de la compañía, con el fin de analizar tendencias, fijar la evolución de parámetros de negocio... etc.

### 2.2.9. INTEGRACIÓN DE DATOS

La integración de datos es el proceso que permite combinar datos heterogéneos de muchas fuentes diferentes en la forma y estructura de una única aplicación. Este proceso de integración de datos facilita que diferentes tipos de datos, tales como matrices de datos, documentos y tablas, sean fusionados por usuarios, organizaciones y aplicaciones para un uso personal, de procesos de negocio o de funciones. (Powerdata, 2017)

La integración de datos soporta el procesamiento analítico de grandes conjuntos de datos alineando, combinando y presentando cada conjunto de datos de departamentos organizacionales y fuentes de datos remotas y externas, para cumplir con los objetivos del integrador. (Powerdata, 2017)

La integración de datos se implementa generalmente en un data warehouse mediante software especializado que aloja grandes repositorios de datos de recursos internos y externos. Los datos se extraen, se mezclan y se presentan de forma unificada. Por ejemplo, el conjunto completo de datos de un usuario puede incluir datos extraídos y combinados de marketing, ventas y operaciones, que se combinan para formar un informe completo. (Powerdata, 2017)

Un proyecto de integración de datos generalmente implica los siguientes pasos:

- Acceso a los datos desde todas las fuentes y localizaciones tanto si se trata de locales, en la nube o de una combinación de ambos. (Powerdata, 2017)
- Integración de datos de modo que los registros de una fuente de datos mapean registros en otra. Por ejemplo, incluso si un conjunto de datos utilizara “nombre, apellidos” y otro “nom, ape”, el conjunto integrado se asegurará de que en ambos casos los datos van al lugar correcto. Se trata de un tipo de preparación de datos esencial para que las analíticas y otras aplicaciones sean capaces de utilizar los datos con éxito. (Powerdata, 2017)
- Entrega de datos integrados al negocio justo en el momento en que la empresa los necesita, ya sea por lotes, casi en tiempo real o en tiempo real. (Powerdata, 2017)

#### **2.2.10. MS SQL Server 2014**

MS SQL Server es una plataforma de datos de Microsoft que entrega un nivel de desempeño interesante para aplicaciones de misión crítica a través de tecnologías en memoria, perspectivas más rápidas proveniente de cualquier tipo de datos con herramientas familiares como Excel y una plataforma para desarrollar, implementar y administrar soluciones que abarquen tanto implementaciones locales como en la nube. (MSDN Microsoft, 2017)

Esta solución, tiene varias versiones que de adaptan a las necesidades de la empresa. Por ejemplo:

#### **SQL SERVER ENTERPRISE**

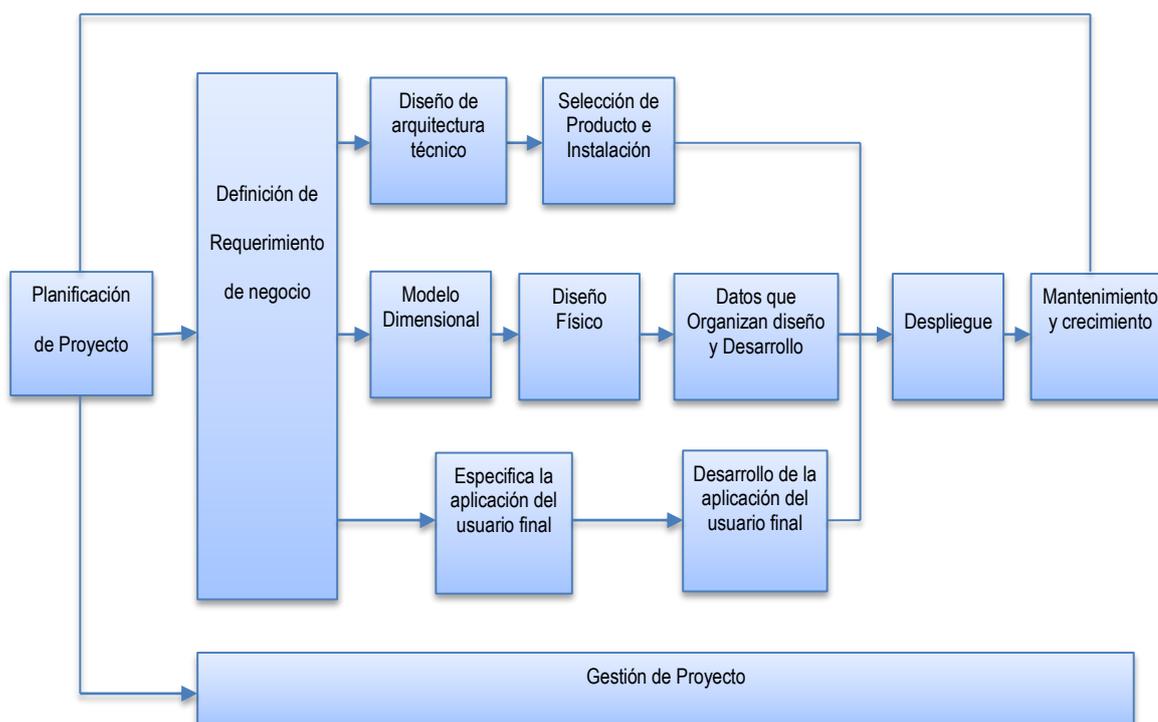
SQL Server Enterprise ofrece funciones avanzadas para el centro de datos y para los más exigentes requisitos de bases de datos y de inteligencia empresarial. Sus principales características son:

- ✓ **Alta disponibilidad mejorada.** Permite obtener un mayor tiempo de actividad, una conmutación por error más rápida, una capacidad de administración mejorada y un uso mejor de los recursos de hardware gracias a AlwaysOn para la alta disponibilidad. Según Microsoft, AlwaysOn entrega hasta cuatro secundarios activos y legibles, operaciones en línea más extensas y un panel integrado para administrar la alta disponibilidad. (ITSitio, 2017)
- ✓ **Rendimiento en memoria.** Ofrece un salto cuántico en el rendimiento del almacenaje de datos, el análisis y la inteligencia empresarial. ¿Cómo? Gracias a las funciones en memoria integradas directamente en los motores de base de datos SQL Server y de análisis, mejora la flexibilidad y se facilita el uso. (ITSitio, 2017)
- ✓ **Escalabilidad de categoría empresarial.** Logre una gran escala, tanto en las instalaciones locales tradicionales como en las implementaciones altamente virtualizadas, con Windows Server 2012 R2. “La escala del procesamiento físico crece hasta 640 procesadores lógicos y hasta 64 procesadores lógicos para las máquinas virtuales con 1 TB de RAM”, dicen desde Microsoft. (ITSitio, 2017)
- ✓ **Inteligencia empresarial de autoservicio.** Descubra conocimientos nuevos mediante Power View, una herramienta de exploración y visualización de datos altamente interactiva para los analistas comerciales y usuarios finales. Power View permite visualizar modelos analíticos tabulares y multidimensionales. (ITSitio, 2017)
- ✓ **Calidad de datos robusta.** Mejore la calidad de los datos al usar conocimientos organizacionales y proveedores de datos de referencia independientes para perfilar, limpiar y hacer coincidir los datos mediante Data Quality Services. (ITSitio, 2017)
- ✓ **Seguridad mejorada.** Mejore la facilidad de administración con respecto a las auditorías con resiliencia a la auditoría, auditorías definidas por el usuario y la capacidad de auditar acciones en las diferentes ediciones. (ITSitio, 2017)

## 2.3. METODOLOGÍA PARA EL DESARROLLO DEL PROYECTO

Existen diversas tecnologías y metodologías para el desarrollo de un BI. Nosotros optamos por la metodología de Ralph Kimball:

### 2.3.1. Metodología de Ralph Kimball



**Figura N° 5:** Metodología de implementación

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

#### 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: (Kimball, 2013)

**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. (Kimball, 2013)

**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. (Kimball, 2013)

**e. Diseño Físico**

El diseño físico de la 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. (Kimball, 2013)

**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. (Kimball, 2013)

**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. (Kimball, 2013)

**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.) (Kimball, 2013)

**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. (Kimball, 2013)

**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. (Kimball, 2013)

**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 está 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. (Kimball, 2013)

### **3. MATERIALES Y MÉTODOS**

#### **3.1. MATERIAL**

##### **3.1.1. Población**

Todos los reportes que sirven en la toma de decisiones en el proceso de Ventas de la empresa Librería Crisol S.A.C. sucursal Trujillo.

##### **3.1.2. Muestra**

Se analizarán 10 reportes principales dentro del proceso de Ventas de la empresa Librería Crisol S.A.C.

##### **3.1.3. Unidad de análisis**

Datos de las ventas realizadas en la sucursal de Trujillo de la empresa Librería Crisol S.A.C.

#### **3.2. MÉTODO**

##### **3.2.1. Tipo de investigación**

Aplicada.

##### **3.2.2. Diseño de Investigación**

Se realizará un diseño cuasi-experimental en el cual se evaluará la eficiencia de la nueva solución (BI) para el soporte en la toma de decisiones del área de ventas.

Se realizarán observaciones antes y después del tratamiento.

Diagrama de investigación cuasi-experimental	G -> O1 -> X -> O2
G (Grupo a investigar)	Tomadores de decisión en la Empresa (02 personas)
X (Tratamiento)	Aplicación del BI
O (Observación)	O1: Observación pre-test
	O2: Observación post-test

**Tabla 1. Diagrama de investigación**

### 3.2.3. Variables de estudio y Operacionalización

- ✓ Independiente (VI): Solución de Inteligencia de negocios
- ✓ Dependiente (VD): Mejorar la Integración de Datos y el análisis de información para un mejor soporte en la toma de decisiones en el proceso de Ventas de la empresa Librería Crisol S.A.C. sucursal Trujillo.

Tabla N° 2: Operacionalización de las variables

Variable	Dimensión	Indicador	Unidad de medida	Instrumento de Investigación
VD	Oportunidad	Tiempo para analizar información	Minutos	Hoja de captura de datos
	Conformidad	Grado de satisfacción del proceso	% Satisfacción	Hoja resumen de grado de satisfacción
	Integración	N° Datos Integrados	Registros	Datos integrados

### **3.2.4. Técnicas e instrumentos de recolección de datos**

#### **3.2.4.1. Técnicas**

Con el fin de obtener los mejores resultados posibles se ha seleccionado el método inductivo - deductivo para efectos investigativos.

Este método permite mediante un conjunto de preguntas adecuadas inducir las respuestas sobre la ratificación de la existencia del problema, una vez que esto ha sucedido se deduce las posibles soluciones al problema.

Entre las técnicas para recopilar información dentro de este proceso investigativo podemos mencionar:

- ✓ La encuesta: que consiste en investigar la sintomatología del problema en base a una serie de preguntas relacionadas con las dificultades que atraviesa la empresa.
- ✓ La entrevista es una indagación directa con una persona directamente relacionada con el problema.

El análisis documental que permitirá obtener información de para la documentación del proyecto y desarrollo de la solución.

#### **3.2.4.2. Instrumentos**

Se utilizarán los siguientes instrumentos en las técnicas mencionadas anteriormente:

- ✓ Cuestionarios sobre las Variable Independiente y Dependiente.
- ✓ Ficha de Observación Sobre la variable Independiente.
- ✓ Entrevistas. Sobre la variable dependiente.
- ✓ Análisis. El análisis que utilizaremos es la Estadística Descriptiva: Media aritmética, desviación estándar, gráfica, cuadros.

### **3.2.5. Técnicas de procesamiento y análisis de datos**

#### **3.2.5.1. Procesamiento de datos**

Una vez la puesta en marcha el cuestionario los datos serán ingresados en el Ms Excel para la realización de análisis descriptivo. Las respuestas que se obtengan se presentarán en tablas y gráficos dinámicos, dando a conocer como es los resultados numéricos y porcentuales de las variables en estudio según los objetivos planteados anteriormente.

#### **3.2.5.2. Análisis de datos**

El análisis de los datos se llevará a cabo por medio de cuadros estadísticos descriptivos (Pruebas hipótesis nula y alternativa y las Pruebas t de Student siendo la mejor opción por el tamaño de la muestra = 10 reportes), ya que la información obtenida será analizada y mostrada por medio de cuadros y gráficos.

El tamaño de la muestra (10) se obtuvo del total de requerimientos de los usuarios finales (Tomadores de decisión) obtenidos de las entrevistas.

## 4. RESULTADOS: APLICACIÓN DE LA METODOLOGÍA

### 4.1. ETAPA 1: PLANEACIÓN Y ADMINISTRACIÓN DEL PROYECTO

#### 4.1.1. EL NEGOCIO

##### a. Descripción de la organización

**Nombre de la empresa o Razón social:** Librería Crisol S.A.C.

Sucursal Trujillo.

**Ruc N°:** 20501457869

**Rubro económico:** Librerías.

La Librería Crisol es una empresa de retail especializada, dedicada a la venta minorista de libros y artículos de entretenimiento cultural. Sus actividades se concentran en la venta de libros diversos en un ambiente agradable, diferenciado por su calidad de servicio y el asesoramiento de sus librereros. En la actualidad, Librerías Crisol cuenta con 29 tiendas a nivel nacional y está presente también en la ciudad de Trujillo, así como en otras ciudades.

##### **Competidores:**

- Librerías el Inca.
- Librerías Book Center
- Librerías San Pablo
- Librerías Communitas
- Librerías el Virrey
- Librerías SBS Internacional

##### **b. Misión:**

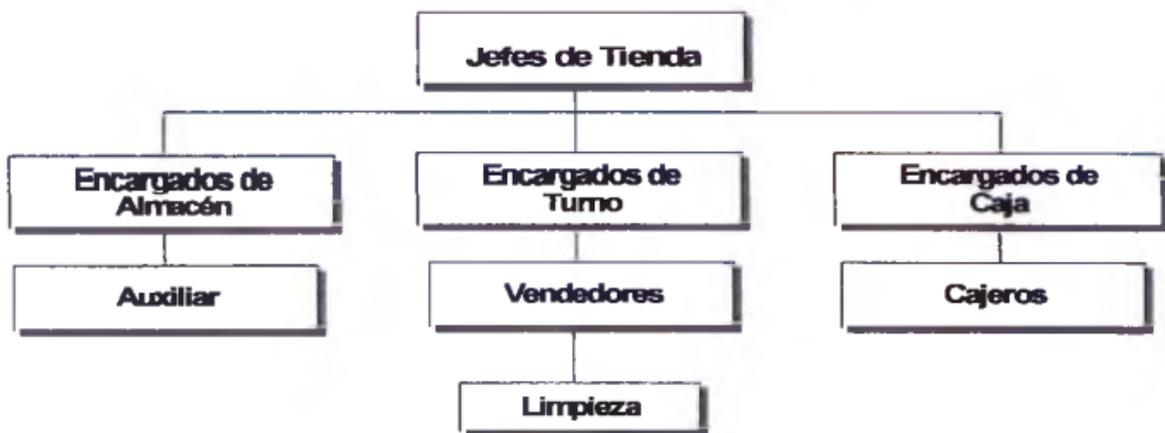
*“Somos la cadena de librerías con la oferta más amplia y variada de entretenimiento cultural, brindando el mejor nivel de servicio en el mercado peruano. Nos caracterizamos por los altos estándares de respeto, calidez y amabilidad de nuestros colaboradores”.*

*“Nuestro objetivo es ofrecer una experiencia de disfrute y crecimiento personal en un ambiente estimulante, contribuyendo a difundir la cultura en nuestro país”.*

**c. Visión:**

*“Convertirnos en la principal cadena de tiendas de entretenimiento cultural ampliando nuestra red de librerías hacia las principales ciudades del país, y de esta manera, convertirnos progresivamente en referente cultural de América Latina”.*

**d. Organigrama empresarial por tienda**



Fuente: Gerencia Comercial

*Figura N° 4: Organigrama por tienda*

#### **4.1.2. Selección de la Estrategia de Implementación**

Los pasos a seguir para realizar un correcto desarrollo, y así mismo, una correcta implementación del proyecto es:

- a) Conocer a la organización e involucrarse con la empresa de tal forma que se pueda conocer cuáles son sus necesidades de información.
- b) Revisar y analizar la documentación para identificar los procesos de la empresa.
- c) Estudiar y familiarizarse con la metodología a trabajar en el proyecto.
- d) Ver el funcionamiento de la base de datos de la empresa y archivos necesarios que guardan información de ella.
- e) Elaborar un cronograma de trabajo.
- f) Recopilar requerimientos empresariales.
- g) Realizar un análisis de cada uno de los requerimientos.
- h) Realizar el diseño para la implementación de un BI.
- i) Construir un nuevo Modelo dimensional para el análisis de la información de las Ventas.
- j) Implementar el BI

#### **4.1.3. Selección de la Metodología de Desarrollo**

La implementación de un Modelo dimensional es un proceso complejo, es por ello que se debe usar la mejor práctica existente. El método propuesto por Ralph Kimball es el ideal para utilizar, adaptándola al caso específico de este proyecto.

#### **4.1.4. Selección del Ámbito de Implementación**

El ámbito de la implementación del BI fue determinado según la cantidad de datos que se maneja en el área de ventas en la sucursal de Trujillo y los requerimientos analizados de la empresa, con el apoyo de los conocimientos adquiridos para su desarrollo.

Para ello se formularon las siguientes preguntas:

- a) ¿Por qué implementar un BI para el área de ventas?
- b) ¿Qué consultas tienen prioridad?
- c) ¿Cuál es el área más recomendable estratégicamente para aplicar el BI?
- d) ¿Cuál es la priorización de desarrollo del BI en los procesos de las principales áreas de la empresa?
- e) ¿Cuál es el rango de consultas empresariales a los que se debe responder inicialmente la solución?

Después del análisis, las respuestas obtenidas fueron:

- a) El área de Ventas en la librería es el núcleo del negocio y es donde más cantidad de datos e información crucial se maneja para la empresa.
- b) Para las revisiones lo que más se genera son consultas de reportes (ventas según periodos, pedidos, etc.), datos, pedidos, entregas.
- c) El área de ventas es el Core del negocio, y es en donde se debe aplicar la solución.
- d) El rango de consultas empresariales que debe responder inicialmente la solución de BI va ser de 10 requerimientos o consultas.

#### **4.1.5. Selección del enfoque arquitectónico**

El enfoque está compuesto por dos niveles

➤ Back Room

- Servidor HP ProLiant ML30 G9 Torre 4U, Intel Xeon E3-1220v6 Quad Core (3.00GHz 8MB) - 8GB DDR4 - HDD 1TB SATA - DVD-RW.

➤ Front Room

- PC's de tienda y vendedores, siendo estas computadoras las que deben de acceder a la información que dará la solución de BI.

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

##### 4.1.6.1. Presupuesto

###### a. Personal.

N°	Personal	Descripción	Cantidad	Costo S/.
01	Investigador	Bach. Denis Anderson Bernabe Barrios	01	2000.00
02	Investigador	Bach. Leonard Leyva Hurtado	01	2000.00
03	Asesor	Ing. Agustín Eduardo Ullón Ramírez	01	500.00
			<b>TOTAL</b>	<b>4500.00</b>

###### b. Bienes.

###### Hardware

N°	Descripción	Precio Unitario S/.	Unidad De Medida	Cantidad	Costo S/.
01	HP LAPTOP Procesador Intel Core i3 1TB Disco - RAM 4GB (Backroom)	2500.00	Unidad	01	2500.00
02	Servidor HP ProLiant ML30 G9 Torre 4U, Intel Xeon E3-1220v6 Quad Core (3.00GHz 8MB) - 8GB DDR4 - HDD 1TB SATA - DVD-RW	3800.00	Unidad	01	3800.00
03	Impresora Multifuncional EPSON L220.	500.00	Unidad	01	500.00
04	Memoria USB HP 16GB	30.00	Unidad	01	30.00
				<b>TOTAL</b>	<b>6830.00</b>

Software

N°	Descripción	Precio Unitario S/.	Unidad De Medida	Cantidad	Costo S/.
01	Microsoft Office 2013.	300.00	Unidad	1	300.00
02	Windows 8.	400.00	Unidad	1	400.00
03	MS SQL SERVER 2014	400.00	Unidad	1	400.00
04	Qlik Sense	400.00	Unidad	1	400.00
05	Microsoft Azure	660.00	Mensual	1	660.00
				<b>TOTAL</b>	<b>2160.00</b>

Insumos

N°	Descripción	Precio Unitario S/.	Unidad De Medida	Cantidad	Costo S/.
01	Papel Bond	20.00	Millar	01	30.00
02	Cartucho de tinta Negro	80.00	Unidad	01	80.00
03	Cartucho de tinta Colores	90.00	Unidad	01	90.00
04	CD-ROM	1.00	Unidad	5	5.00
				<b>TOTAL</b>	<b>205.00</b>

Servicios

N°	Descripción	Precio Unitario S/.	Unidad De Medida	Cantidad	Costo S/.
01	Fotocopias	0.10	Unidad	600	60.00
02	Pasajes	1.00	Unidad	100	100.00
03	Refrigerios	5.00	Unidad	50	250.00
				<b>TOTAL</b>	<b>410.00</b>

**RESUMEN:**

Rubro	Costo S/.
<b>Personal</b>	<b>4500.00</b>
<b>Recursos disponibles</b>	<b>8990.00</b>
<b>Materiales y equipos disponibles (Hardware)</b>	<b>6830.00</b>
<b>Materiales y equipos disponibles (Software)</b>	<b>2160.00</b>
<b>Recursos NO disponibles</b>	<b>615.00</b>
<b>Insumos</b>	<b>205.00</b>
<b>Servicios</b>	<b>410.00</b>
<b>TOTAL PRESUPUESTO</b>	<b>14105.00</b>

**4.1.6.2. Cronograma de actividades para el DataMart**

Actividades	Abril				Mayo				Junio				Julio				Agosto				N° HRS/SEM
	S 1	S 2	S 3	S 4																	
1. Definición de los requerimientos, problema, metas y objetivos.																					3 semana
2. Recolección de información del proceso actual (costo, recursos, tiempo, etc.)																					3 semanas
3. Análisis de Requerimientos.																					3 semanas
4. Preparación de los datos																					3 semanas

Actividades	Abril				Mayo				Junio				Julio				Agosto				N° HRS/SEM
	S 1	S 2	S 3	S 4																	
5. Modelado de datos																					4 semanas
7. ETL																					3 semanas
8. Evaluación de requerimientos																					3 semanas
9. Redacción del informe																					3 semanas
10. Presentación del Informe																					4 semanas

*Figura N° 3: Cronograma de actividades*

#### 4.1.7. Desarrollo del escenario del uso empresarial

El escenario de uso empresarial es el que se muestra a continuación:

##### A. Descripción de los Stakeholder

- Personal Involucrado en el Proyecto

Nombre	Representa	Rol
<b>JEFE DE TIENDA</b>	Responsable del buen funcionamiento de la tienda en la sucursal de Trujillo	<ul style="list-style-type: none"> <li>• Establecer y planificar el trabajo para que se logren los objetivos esperados dentro de la sucursal que dirige.</li> </ul>
<b>ENCARGADO DE TURNO</b>	Persona encargada de llevar el orden y buen funcionamiento de la tienda.	<ul style="list-style-type: none"> <li>• Recopilación de requerimiento por clientes de compras o pedidos.</li> </ul>

- **Usuarios del Data Mart**

Nombre	Representa	Stakeholder
<b>JEFE DE TIENDA</b>	Es la persona encargada de establecer las estrategias necesarias para el buen desempeño de la empresa	<b>Con permiso para acceder a todas las dimensiones del BI</b>
<b>ENCARGADO DE TURNO</b>	Encargado de registrar el movimiento de las ventas, así como de ver la manera como se reabastecen los productos.	<b>Encargado de Turno</b>

## **4.2. ETAPA 2: DETERMINACIÓN DE LOS REQUERIMIENTOS**

### **4.2.1. Requerimientos del Propietario**

Responder las preguntas listadas a continuación para que se pueda tener un mejor conocimiento acerca de la importancia de este proyecto.

#### **a. ¿Para qué un BI?**

Para tener estrategias de cómo obtener un mayor provecho en la información que proporcionan los datos históricos, desarrollando una estrategia empresarial que persigue incrementar el rendimiento de la empresa a través de la organización inteligente de sus datos.

#### **b. ¿Cuál será el impacto sobre la organización?**

Permitirá tomar mejores decisiones favorables a la empresa y actuar en consecuencia basándose en la información analizada y así poder desarrollar e implementar nuevas estrategias, lo cual permitirá tomar mayores ventajas sobre las demás empresas.

**c. ¿Cómo afecta en los costos actuales en cómputo?**

La empresa cuenta con computadoras o terminales de casi todos los usuarios finales, donde sólo se necesitará hacer una inversión en la instalación del BI que es de S/. **14105.00**

**d. ¿Cuáles son los riesgos?**

La implementación de un BI podría presentar riesgos como las inconsistencias en la data, también una mala normalización de las tablas, logrando que la información sea inestable.

**4.2.2. Requerimientos del Usuario Final**

La obtención de los requerimientos fue a través de entrevistas y reuniones técnicas formales con los usuarios finales que utilizará el BI, de lo cual se obtuvo:

**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 categoría en un mes determinado?
- R3. ¿Cuál es el monto en ventas por sub categorías en un mes determinado?
- R4. ¿Cuál es el monto de ventas por editorial y por mes?
- R5. ¿Cuál es el monto que generan las ventas de un libro por mes y año?
- R6. ¿Cuál es la cantidad de libros vendidos por categoría, sub categoría por mes y semestre?
- R7. ¿Cuál es la cantidad de Productos vendidos por un vendedor al mes y al año?
- R8. ¿Cuál es el monto que generan las ventas de libros por tipo de presentación en un mes, semestre o año?
- R9. Listar a los Vendedores y el monto total de cada una de las ventas que han generado
- R10. ¿Cuál cantidad de ventas de libros por autor, por mes y año?

<b>Nº</b>	<b>CONSULTAS</b>	<b>UNIDAD</b>
1	¿Cuál es el monto de un determinado cliente en compras en un mes?	Monto en soles
2	¿Cuál es el Monto de ventas por categoría en un mes determinado?	Monto en soles
3	¿Cuál es el monto en ventas por sub categorías en un mes determinado?	Monto en soles,
4	¿Cuál es el monto de ventas por editorial y por mes ?	Monto en soles,
5	¿Cuál es el monto que generan las ventas de un libro por mes y año?	Monto en soles
6	¿Cuál es la cantidad de libros vendidos por categoría, sub categoría por mes y semestre?	Cantidad de Libros
7	¿Cuál es la cantidad de Productos vendidos por un vendedor al mes y al año?	Cantidad de Libros
8	¿Cuál es el monto que generan las ventas de libros por tipo de presentación en un mes, semestre o año?	Monto en soles
9	Listar a los Vendedores y el monto total de cada una de las ventas que han generado	Monto en soles,
10	¿Cuál cantidad de ventas de libros por autor, por mes y año?	Cantidad de Libros

*Tabla 2: Consultas y medidas*

#### 4.2.3. Requerimientos no Funcionales

- Establecer la seguridad adecuada para la administración del BI.
- Proporcionar información a los usuarios en el menor tiempo posible.

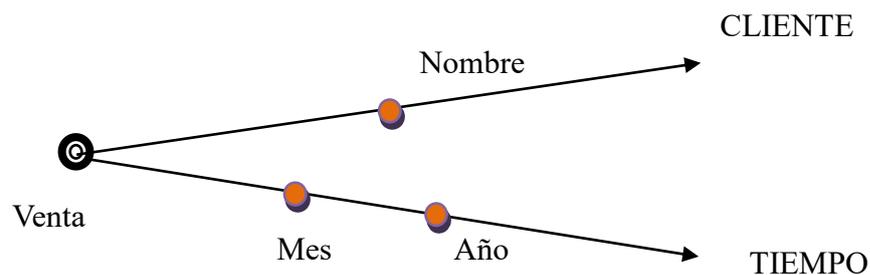
- El BI deben funcionar, con el manejador de base de datos Microsoft SQL Server 2014 y utilizando el Microsoft SQL Server Análisis Services como herramienta de diseño y construcción de los cubos.
- Utilizar Qlik Sense para mostrar los reportes desde los cubos.
- La solución debe poseer un tiempo de respuesta breve.
- El rendimiento del BI debe ser superior a las herramientas utilizadas para la consulta en los sistemas transaccionales.

#### 4.2.4. Análisis de los requerimientos

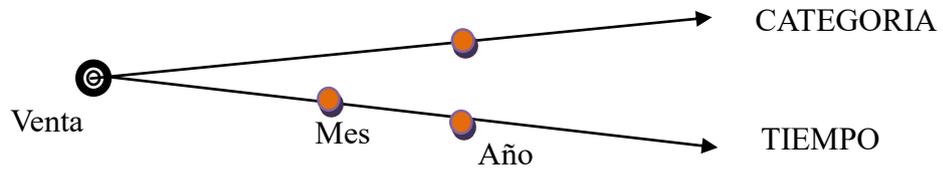
Se procederá a realizar un análisis de cada requerimiento anteriormente mencionado, definiendo algunas tablas de la Base de Datos Operacional de la empresa, que actuarían como dimensión dentro de la solución del BI. También se definirá las posibles medidas dentro del BI.

##### A. Área de Ventas

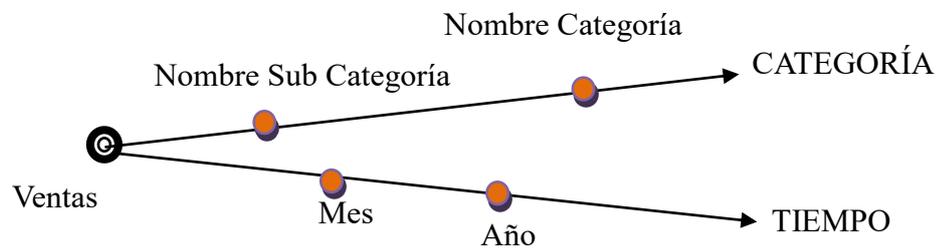
R1 ¿Cuál es el monto de un determinado cliente en compras en un mes?



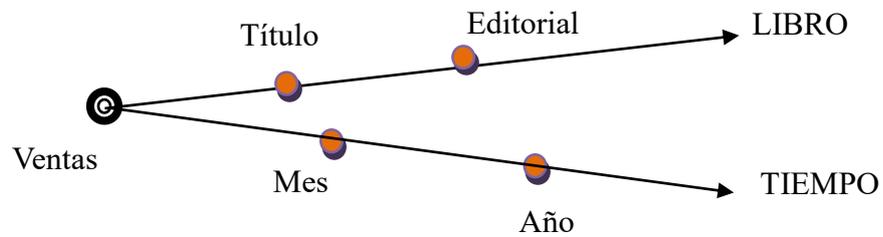
R2 ¿Cuál es el Monto de ventas por categoría en un mes determinado?



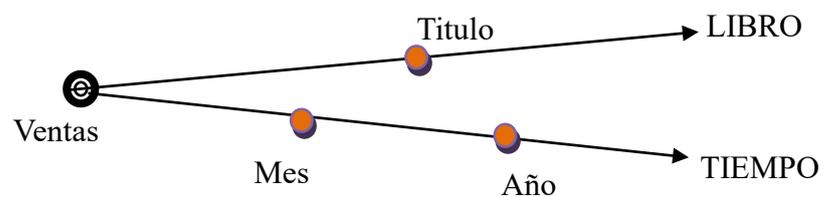
R3 ¿Cuál es el monto en ventas por sub categorías en un mes determinado?



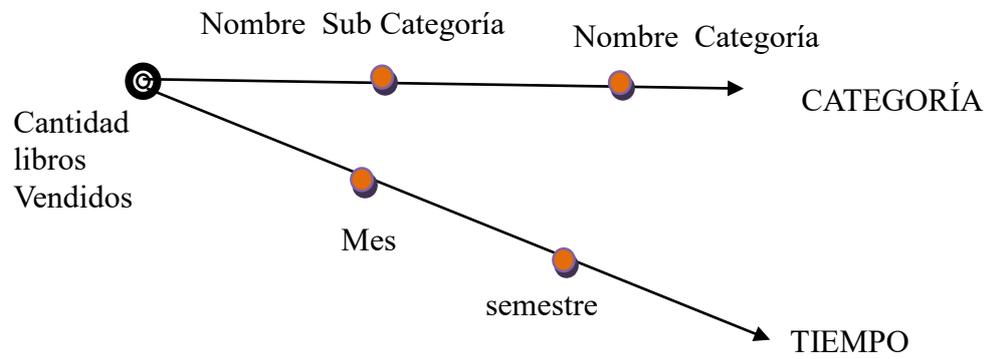
R4 ¿Cuál es el monto de ventas por editorial y por mes?



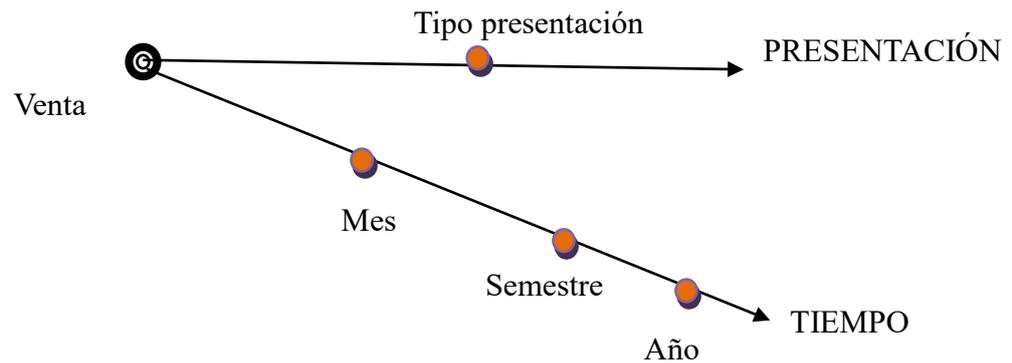
R5 ¿Cuál es el monto que generan las ventas de un libro por mes y año?



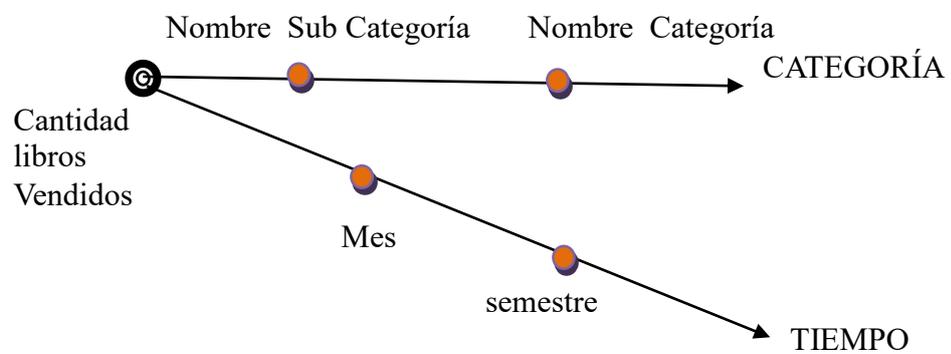
R6 ¿Cuál es la cantidad de libros vendidos por categoría, sub categoría por mes y semestre?



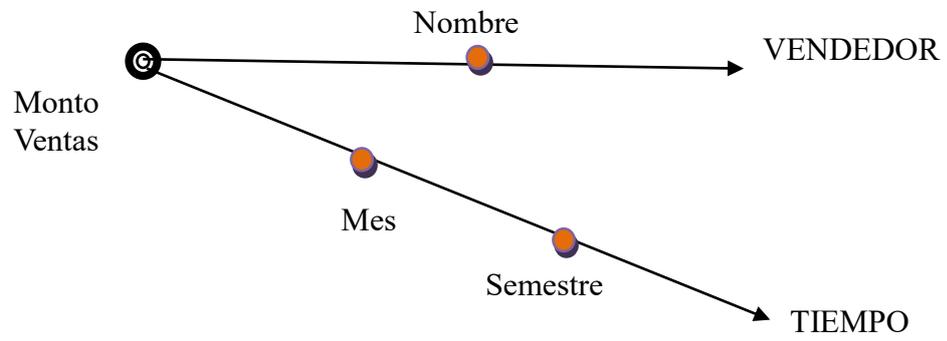
R7 ¿Cuál es el monto que generan las ventas de libros por tipo de presentación en un mes, semestre o año?



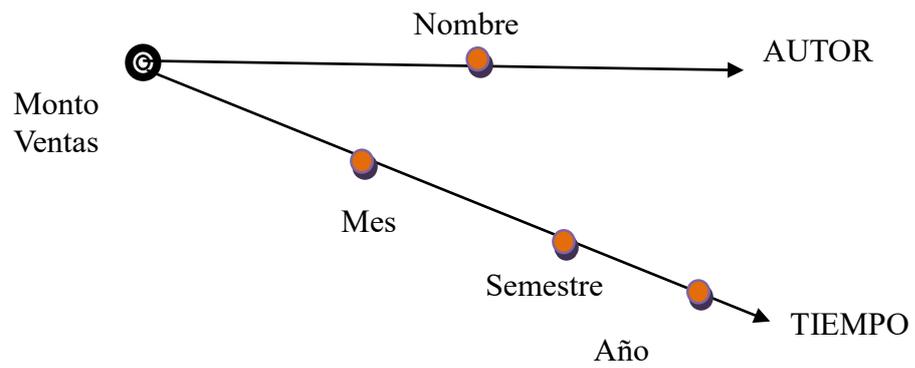
R8 ¿Cuál es la cantidad de libros vendidos por categoría, sub categoría por mes y semestre?



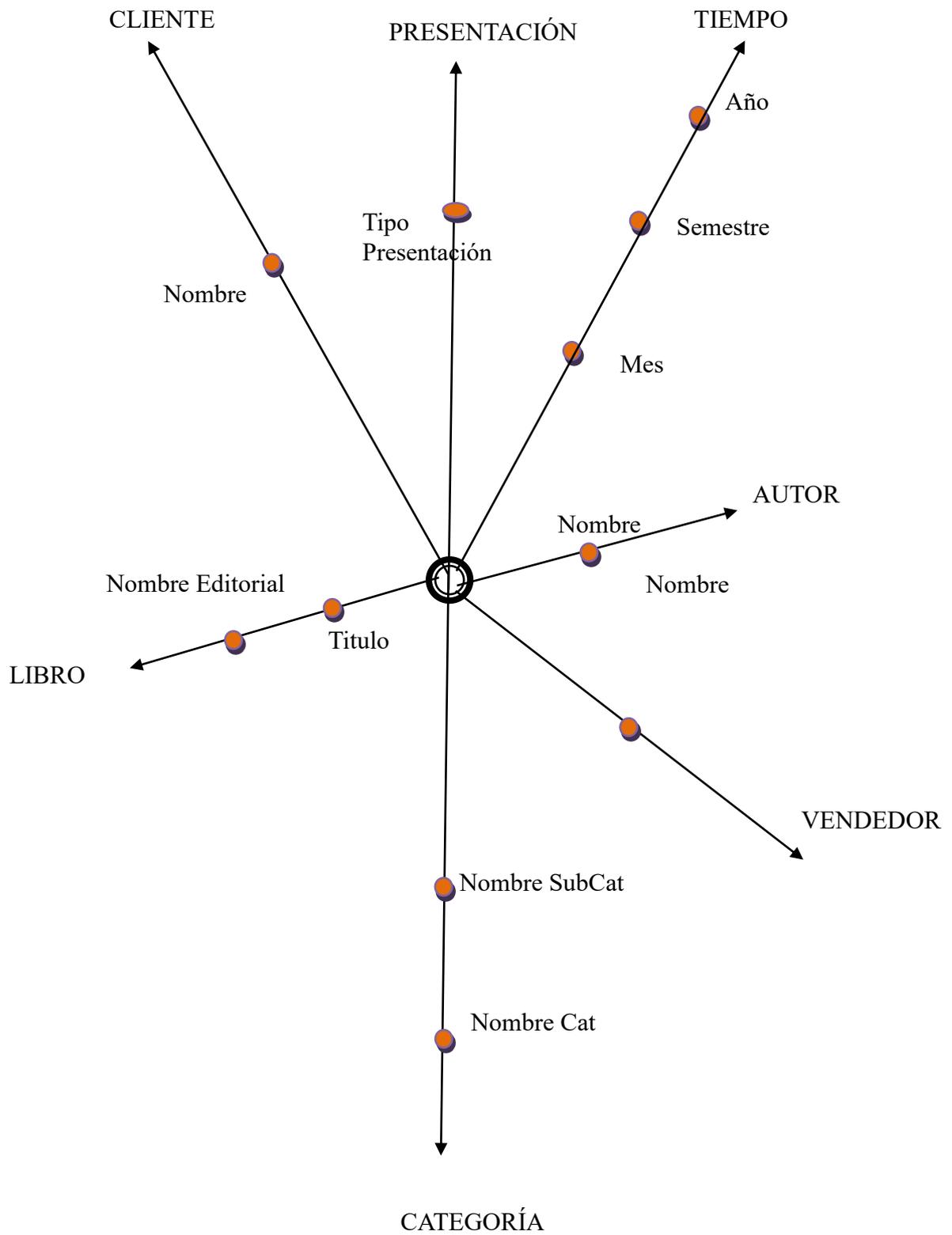
R9 Listar a los Vendedores y el monto total de cada una de las ventas que han generado?



R10 ¿Cuál cantidad de ventas de libros por autor, por mes y año?



### Modelo Startnet: Área de ventas



### 4.3. ETAPA 3: DISEÑO TÉCNICO DE LA ARQUITECTURA

#### 4.3.1. Nivel de Datos (Base de Datos Transaccional)

La sucursal de la empresa Crisol en Trujillo trabaja con las fuentes de datos relacional, así como hojas de cálculo. El tipo de base de datos y la estructura que conforman las tablas se muestra en la siguiente imagen:

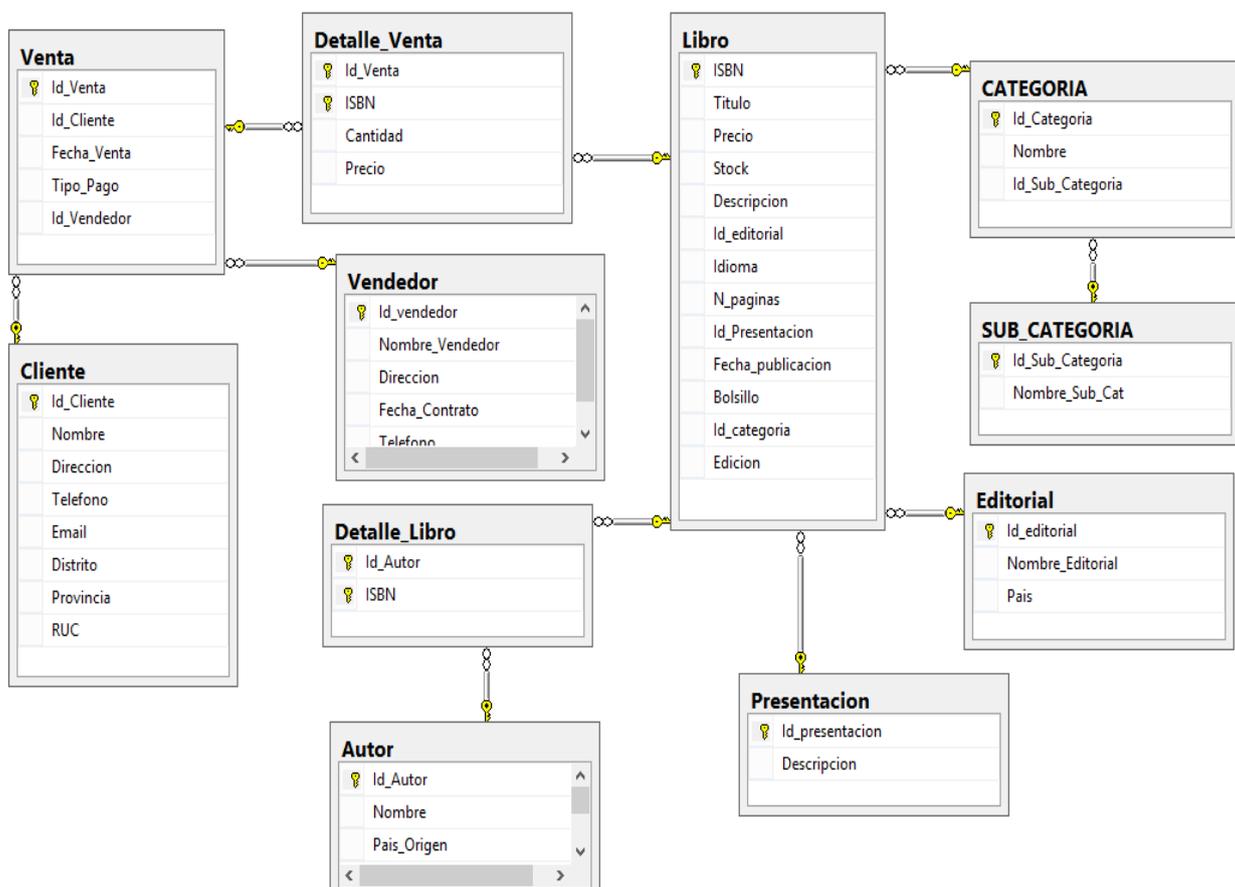
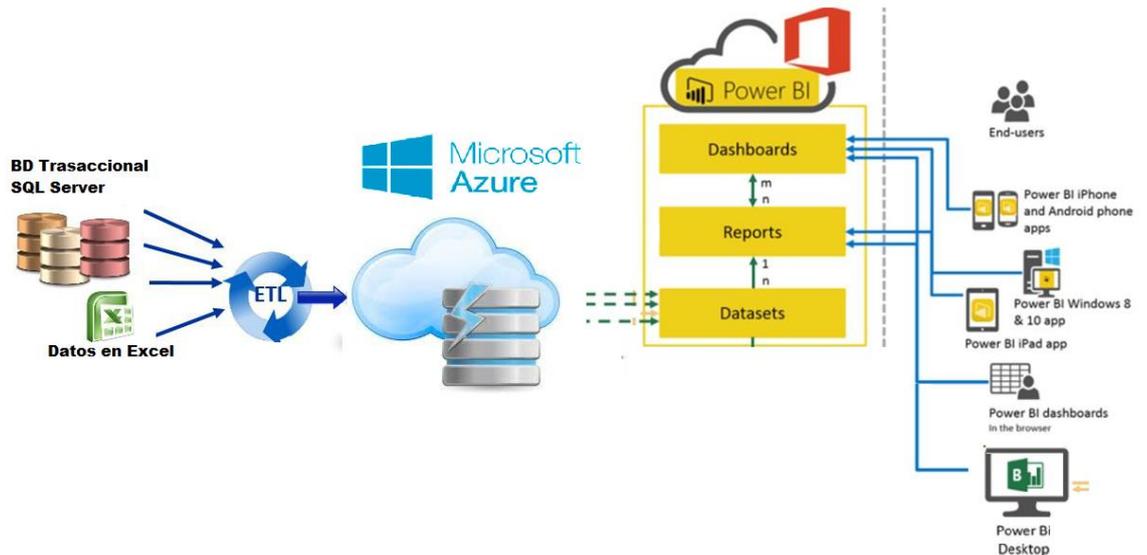


Figura 6: Diagrama E-R de la base de datos

### 4.3.2. Nivel Técnico



*Figura 7: Nivel Técnico*

La solución de BI está compuesta por dos niveles:

➤ **Back Room:**

- Servidor HP ProLiant ML30 G9 Torre 4U, Intel Xeon E3-1220v6 Quad Core (3.00GHz 8MB) - 8GB DDR4 - HDD 1TB SATA - DVD-RW
- Almacenamiento del modelo dimensional en Microsoft Azure, que servirá como repositorio del BI para la empresa.

➤ **Front Room**

- PC's clientes, dispositivos móviles con acceso a la información que brindaran la solución, a través de Qlik sense.

## 4.4. ETAPA 4: MODELADO DIMENSIONAL

### 4.4.1. Identificación de los Componentes del Modelo

De acuerdo a la metodología de Ralph Kimball, para el siguiente paso se procederá a aplicar el “método de los cuatro pasos” que nos ayudará a modelar el esquema estrella para luego ser poblado a través de procesos ETL.

### **Método de los cuatro pasos**

- ✓ Elección del Modelo dimensional
- ✓ Elección del objetivo de la tabla de hechos
- ✓ Elección de las dimensiones
- ✓ Elección de los hechos

#### **a) Elección del Data Mart**

En este paso se procede a identificar el modelo dimensional ideal para la problemática de la empresa Crisol sucursal Trujillo. Para ello utilizaremos el método de la matriz.

**Método de la Matriz:** Aquí se listan las entidades que apoyan a los hechos. De acuerdo a la situación actual del área de la empresa, las fuentes más urgentes para el esquema en el área de ventas, está basada en las filas de la matriz es el Área de Ventas.

**Listado de las dimensiones:** De lo elaborado anteriormente en el análisis de los requerimientos, a continuación, se listarán las posibles dimensiones del modelo.

- ✓ Autor
- ✓ Tiempo
- ✓ Libro
- ✓ Categoría
- ✓ Cliente
- ✓ Presentación
- ✓ Vendedor

#### **b) Elección del objetivo de la tabla de hechos**

En este paso se declara cómo es el registro de las medidas en las tablas de hechos, donde se debe definir de manera precisa que los registros de cada tabla de hechos aparezcan en el diseño del modelado dimensional.

Tabla de Hechos	Objetivo
VENTAS	“Gestionar mejor de las ventas en base a los resultados de las ventas realizadas día a día”

Tabla 3: Objetivos de las tablas de Hecho

**c) Elección de las dimensiones**

Se procede a definir las dimensiones de las tablas de hechos en base a sus objetivos planteados anteriormente.

Tabla de Hechos	Objetivo	Dimensiones
Ventas	“Gestionar mejor de las ventas en base a los resultados de las ventas realizadas día a día”	<ul style="list-style-type: none"> <li>✓ Autor</li> <li>✓ Tiempo</li> <li>✓ Libro</li> <li>✓ Categoría</li> <li>✓ Cliente</li> <li>✓ Presentación</li> <li>✓ Vendedor</li> </ul>

Tabla 4: Dimensiones de la tabla de Hechos Ventas

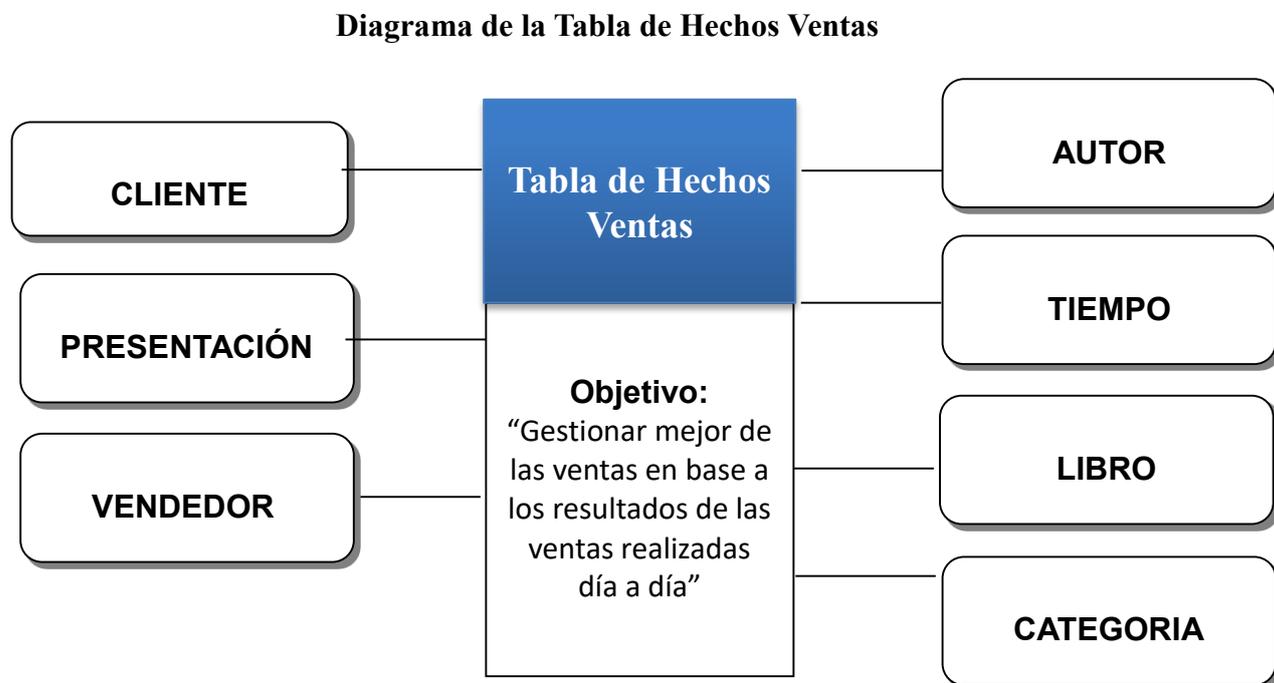
**d) Elección de los hechos**

Cada tabla de hechos tiene un objetivo de permite tener una idea sobre cuál es su alcance de dichos hechos.

Tabla de Hechos	Medida
Ventas	<ul style="list-style-type: none"> <li>✓ Cantidad de Libros</li> <li>✓ Monto de las Ventas</li> </ul>

Tabla 5: Elección de los hechos

#### 4.4.2. Diagrama de la Tabla de Hechos



*Figura N° 8: Diagrama de Tabla de Hechos*

#### Detalle de tabla de hechos y dimensiones:

Nombre de la Tabla	Nombre de la Columna	Descripción de la Columna
Dimensión Tiempo	Tiempo_Key	Llave primaria única para la Dimensión Tiempo
Dimensión Categoría	Categoría_Key	Llave primaria única para la Dimensión Categoría
Dimensión Cliente	Cliente_Key	Llave primaria única para la Dimensión Cliente
Dimensión Presentación	Presentacion_Key	Llave primaria única para la Dimensión Presentación

Dimensión Autor	Autor_Key	Llave primaria única para la Dimensión Autor
Dimensión Libro	Libro_Key	Llave primaria única para la Dimensión Libro
Dimensión Vendedor	Vendedor_Key	Llave primaria única para la Dimensión Vendedor

*Tabla 6: Tabla de dimensiones*

- **Detalle de las Medidas de las Tablas de Hecho**

Nombre de la Tabla	Nombre de la Columna	Descripción de la Columna
Hecho Ventas	Cantidad_Libros	Cantidad Total de Libros vendidos
	Monto_Ventas	Monto Total de ventas

*Tabla 7: Tabla de Hechos ventas*

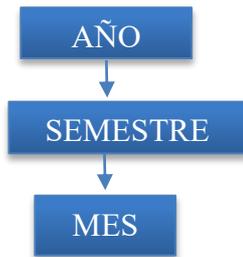
TABLA DE HECHOS VENTAS
Tiempo_Key
Categoria_Key
Vendedor_Key
Cliente_Key
Presentacion_Key
Autor_Key
Libro_Key
Cantidad_Libros
Monto_Ventas

*Figura N° 9: Tabla de Hechos*

- **DETALLE DE LAS TABLAS DIMENSIÓN**

Cada Dimensiones elegidas anteriormente se le realiza una descripción de la construcción de los detalles y sus jerarquías de cada una de ellas.

**DIMENSIÓN TIEMPO**

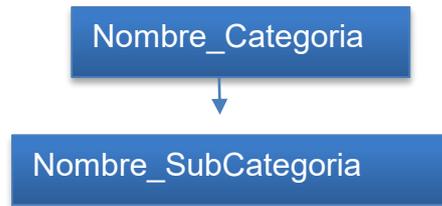


*Figura 10: Jerarquía de los Atributos de la Dimensión Tiempo*

**Tabla 8: Detalle de Dimensión Tiempo Ventas**

Nombre del Atributo	Descripción del Atributo	Cambiando Política	Valores de Muestra
Año	Representa el año.	No Actualizar	2016,2017,...
Semestre	Representa a un Semestre	No actualizar	1, 2,
Mes	Representa los meses que posee un año	No Actualizar	1, 2, ...12

## DIMENSIÓN CATEGORÍA



*Figura 11: Jerarquía de los Atributos de la Dimensión Categoría*

**Tabla 9: Detalle de Dimensión Categoría**

Nombre del Atributo	Descripción del Atributo	Cambiando Política	Valores de Muestra
Nombre_categoria	Representa el nombre de la categoría del libro	No Actualizar	Arte, Ciencias y Humanidades, etc.
Nombre_Sub Categoría	Representa el nombre de la SubCategoría	No Actualizar	Cine, teatro Música, Arte gráfico, etc

## DIMENSIÓN CLIENTE



*Figura 12: Jerarquía de los Atributos de la Dimensión Cliente*

**Tabla 102: Detalle de Dimensión Cliente**

Nombre del Atributo	Descripción del Atributo	Cambiando Política	Valores de Muestra
Nombre_Cliente	Representa el nombre del cliente	No Actualizar	José Castillo, Carlos Luna Victoria, etc

## DIMENSIÓN PRESENTACIÓN

Tipo\_presentacion

Figura 13: Jerarquía de los Atributos de la Dimensión Presentación

Tabla 13: Detalle de Dimensión Presentacion

Nombre del Atributo	Descripción del Atributo	Cambiando Política	Valores de Muestra
Tipo_presentacion	Representa el tipo de presentación que tiene un libro	No Actualizar	Tapa dura, tapa blanda, etc

## DIMENSIÓN AUTOR

Nombre\_Autor

Figura 14: Jerarquía de los Atributos de la Dimensión Autor

Tabla 114: Detalle de Dimensión Autor

Nombre del Atributo	Descripción del Atributo	Cambiando Política	Valores de Muestra
Nombre_Autor	Representa el nombre del autor de los libros	No Actualizar	

## DIMENSIÓN VENDEDOR

Nombre\_Vendedor

Figura 15: Jerarquía de los Atributos de la Dimensión Comprobante

## DIMENSIÓN LIBRO



Figura 16: Jerarquía de los Atributos de la Dimensión Editorial

Tabla 125: Detalle de Dimensión Libro

Nombre del Atributo	Descripción del Atributo	Cambiando Política	Valores de Muestra
Nombre_editorial	Representa las editoriales de los libros	No Actualizar	Lunweg Editors, Edimat Libros, etc
Título	Representa el Título del libros	No Actualizar	Perú. Arquitectura Y Espacios Sagrados, etc

### 4.4.3. ESQUEMA ESTRELLA

Conociendo la cantidad de Tablas de Hechos y las dimensiones asociadas a ella, orientamos nuestro Modelo dimensional a un Esquema Estrella para luego llevarlos al Diseño Físico.

Para un mejor entendimiento, listamos los componentes que contiene el esquema estrella en el Diseño Lógico del BI.

- ✓ Componente: TABLA DE HECHOS VENTAS
- ✓ Componente: TABLA DE DIMENSIÓN TIEMPO

- ✓ Componente: TABLA DE DIMENSIÓN CATEGORIA
- ✓ Componente: TABLA DE DIMENSIÓN LIBRO
- ✓ Componente: TABLA DE DIMENSIÓN VENDEDOR
- ✓ Componente: TABLA DE DIMENSIÓN PRESENTACION
- ✓ Componente: TABLA DE DIMENSIÓN AUTOR
- ✓ Componente: TABLA DE DIMENSIÓN CLIENTE

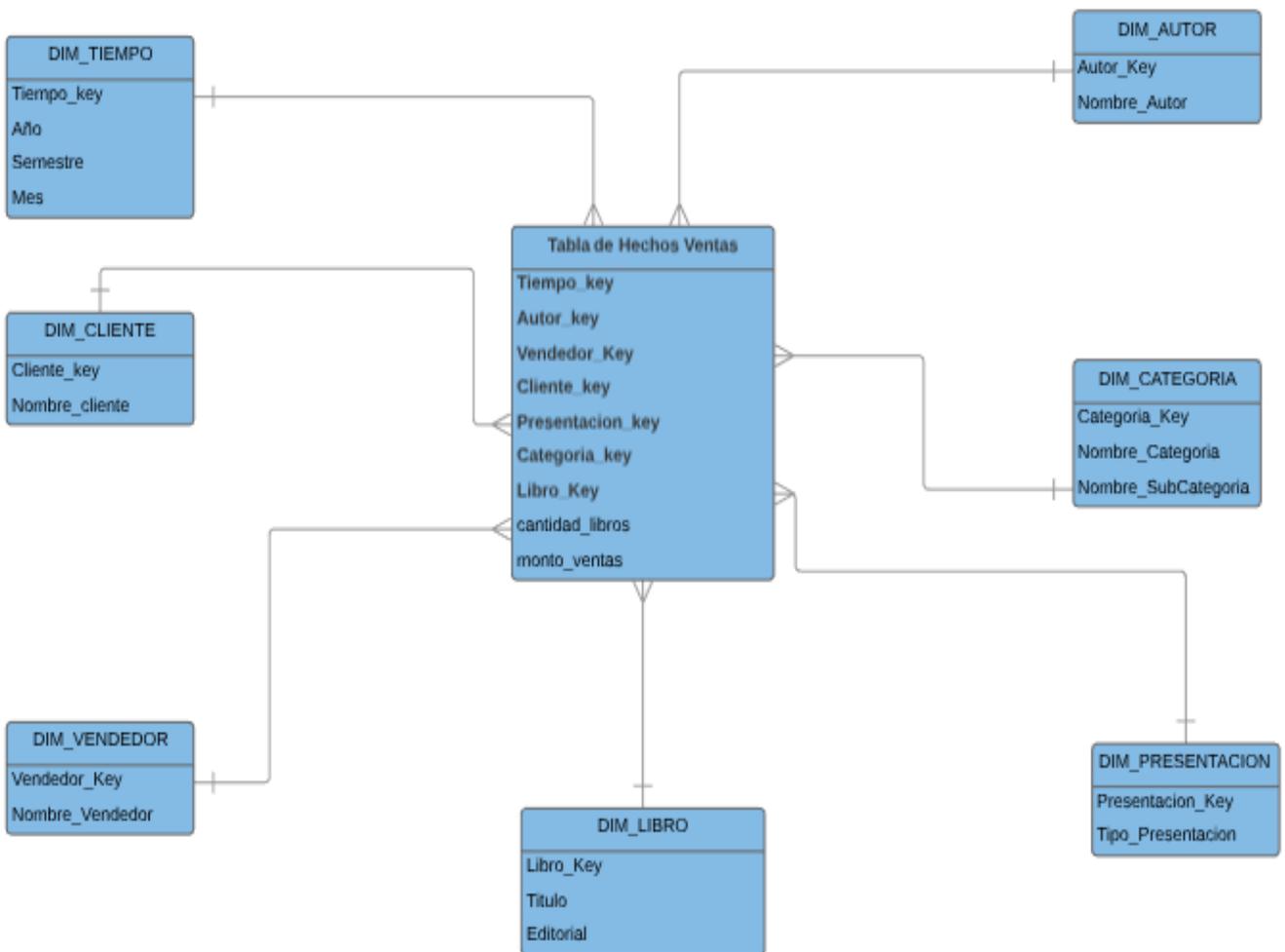


Figura N° 17: Esquema estrella del BI

#### 4.5. ETAPA 5: DISEÑO FÍSICO

Tener en cuenta si es necesario Modificar nombres a nombres estándar.

Para nuestro BI se realizaron los siguientes cambios:

*Tabla 13: Nombres estándares para la Tabla de Hecho y Dimensiones*

<b>DISEÑO LÓGICO</b>	<b>DISEÑO FÍSICO</b>
Hecho_Ventas	Hecho_Ventas
Dim_tiempo	Dim_tiempo
Dim_cliente	Dim_cliente
Dim_categoria	Dim_categoria
Dim_presentacion	Dim_Presentacion
Dim_Autor	Dim_Autor
Dim_Libro	Dim_Libro
Dim_Vendedor	Dim_Vendedor

*Tabla 17: Nombres estándares para los atributos de las dimensiones*

<b>NOMBRE DE LA TABLA</b>	<b>DISEÑO LÓGICO</b>	<b>DISEÑO FÍSICO</b>
Dim_tiempo	Año	Año
Dim_tiempo	Semestre	Semestre
Dim_tiempo	Mes	Mes
Dim_Cliente	Nombre_Cliente	Nombre_Cliente
Dim_Categoria	Nombre_Categoria	Nombre_Categoria
Dim_Presentacion	Tipo_presentacion	Tipo_presentacion

Dim_Libro	Titulo	Titulo
Dim_Autor	Nombre_Autor	Nombre_Autor
Dim_Vendedor	Nombre_Vendedor	Nombre_Vendedor

*Tabla 18: Nombres estándares para los atributos de las Tablas de Hechos*

<b>NOMBRE DE LA TABLA</b>	<b>DISEÑO LÓGICO</b>	<b>DISEÑO FÍSICO</b>
Hecho_ventas	Cantidad_libros	Cantidad_libros
Hecho_ventas	Monto_Ventas	Monto_Ventas

## **ESTRUCTURAS DE LAS TABLAS DIMENSIONES Y TABLA DE HECHO**

También se determina la estructura de las tablas, así como el Tipo de Dato de cada atributo:

### **TABLA: DIM\_TIEMPO**

*Tabla 19: Tipo de Dato para dim\_tiempo*

<b>NOMBRE DEL CAMPO</b>	<b>TIPO DE DATO</b>	<b>LONGITUD</b>
Año	INT	4
Semestre	INT	2
Mes	INT	2

### **TABLA: DIM\_CLIENTE**

*Tabla 20: Tipo de Dato para dim\_cliente*

<b>NOMBRE DEL CAMPO</b>	<b>TIPO DE DATO</b>	<b>LONGITUD</b>
Nombre_Cliente	NVARCHAR	140

**TABLA: DIM\_VENDEDOR***Tabla 21: Tipo de Dato para Dim\_Vendedor*

<b>NOMBRE DEL CAMPO</b>	<b>TIPO DE DATO</b>	<b>LONGITUD</b>
Nombre_Vendedor	NVARCHAR	140

**TABLA: DIM\_CATEGORIA***Tabla 14: Tipo de Dato para Dim\_Categoria*

<b>NOMBRE DEL CAMPO</b>	<b>TIPO DE DATO</b>	<b>LONGITUD</b>
Nombre_Categoria	NVARCHAR	100
Nombre_SubCategoria	NVARCHAR	100

**TABLA: DIM\_AUTOR***Tabla23: Tipo de Dato para Dim\_Autor*

<b>NOMBRE DEL CAMPO</b>	<b>TIPO DE DATO</b>	<b>LONGITUD</b>
Nombre_Autor	NVARCHAR	140

**TABLA: DIM\_LIBRO***Tabla 24: Tipo de Dato para Dim\_Libro*

<b>NOMBRE DEL CAMPO</b>	<b>TIPO DE DATO</b>	<b>LONGITUD</b>
Titulo	NVARCHAR	160
Editorial	NVARCHAR	100

## TABLA: DIM\_PRESENTACION

Tabla 25: Tipo de Dato para Dim\_Presentacion

NOMBRE DEL CAMPO	TIPO DE DATO	LONGITUD
Tipo_Presentacion	NVARCHAR	50

Para las claves primarias de cada dimensión se determinó el tipo de Datos y que se generaran en forma automática (autoincremental) y ordenada:

Tiempo_Key	}	Tipo de Datos Entero
Cliente_Key		
Vendedor_key		
Autor_key		
Presentacion_Key		
Libro_Key		
Categoria_Key		

Especificar las claves foráneas para la Tabla de Hechos:

TABLA DE HECHOS VENTAS
Tiempo_Key Cliente_Key Vendedor_key Autor_key Presentacion_Key Libro_Key Categoria_Key

Figura 1: Claves Foráneas de las Tablas de Hechos

#### 4.5.1. DETERMINACIÓN DE LAS AGREGACIONES

Las Agregaciones se determinaron en su mayoría en operaciones de sumas, como veremos a continuación:

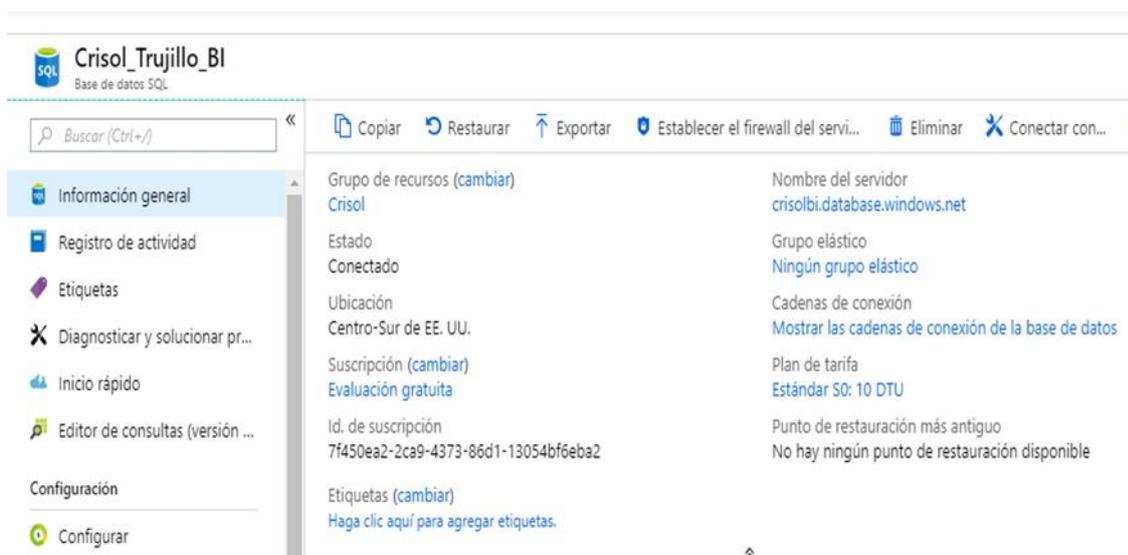
Tabla 15: Determinación de las Agregaciones

Tabla de Hechos	Hecho	Regla de Agregación	Fórmula
Hecho_Ventas	Cantidad_libros	Sum	Select Sum(Cantidad) From detalle_venta
Hecho_Ventas	Monto_Ventas	Sum	Select Sum(cantidad*precio) from detalle_venta

#### 4.5.2. CONSTRUCCIÓN DE LAS TABLAS Y LA BASE DE DATOS

##### ✓ CREACIÓN DE LA BASE DE DATOS EN MICROSOFT AZURE

Después de iniciar sesión en Microsoft Azure se debe de proceder a la creación de una base de dimensional como se detalla en la figura a continuación:



## ✓ CREACIÓN DE LAS DIMENSIONES Y TABLA DE HECHOS EN MICROSOFT AZURE

Luego de haber realizado el diseño físico de las tablas se procede a realizar la construcción de dichas tablas en la base de datos de SQL de Microsoft Azure.

Nota: Se agrega a cada dimensión el id de la tabla transaccional para futuras actualizaciones de las dimensiones.

### TABLA: DIM\_TIEMPO

Column Name	Data Type	Allow Nulls
Tiempo_key	int	<input type="checkbox"/>
Fecha	date	<input type="checkbox"/>
Año	int	<input type="checkbox"/>
Semestre	int	<input type="checkbox"/>
Mes	int	<input type="checkbox"/>

### TABLA: DIM\_CLIENTE

Column Name	Data Type	Allow Nulls
Cliente_key	int	<input type="checkbox"/>
Nombre_Cliente	nvarchar(140)	<input type="checkbox"/>
Id_Cliente	int	<input type="checkbox"/>

### TABLA: DIM\_VENDEDOR

Column Name	Data Type	Allow Nulls
Vendedor_Key	int	<input type="checkbox"/>
Nombre_Vendedor	nvarchar(140)	<input type="checkbox"/>
Id_Vendedor	int	<input type="checkbox"/>

**TABLA: DIM\_CATEGORIA**

CRISOL_TRU...bo.DIM_CATEGORIA			
	Column Name	Data Type	Allow Nulls
▶	Categoria_key	int	<input type="checkbox"/>
	Nombre_Categoria	nvarchar(100)	<input type="checkbox"/>
	Id_Categoria	int	<input type="checkbox"/>
	Sub_categoria	nvarchar(100)	<input type="checkbox"/>

**TABLA: DIM\_AUTOR**

CRISOL_TRUJ...I - dbo.DIM_AUTOR			
	Column Name	Data Type	Allow Nulls
▶	Autor_key	int	<input type="checkbox"/>
	Nombre_Autor	nvarchar(140)	<input type="checkbox"/>
	Id_Autor	int	<input type="checkbox"/>

**TABLA: DIM\_LIBRO**

CRISOL_TRUJ...BI - dbo.DIM_LIBRO			
	Column Name	Data Type	Allow Nulls
▶	Libro_Key	int	<input type="checkbox"/>
	Titulo	nvarchar(160)	<input type="checkbox"/>
	ISBN	int	<input type="checkbox"/>
	Editorial	nvarchar(50)	<input type="checkbox"/>

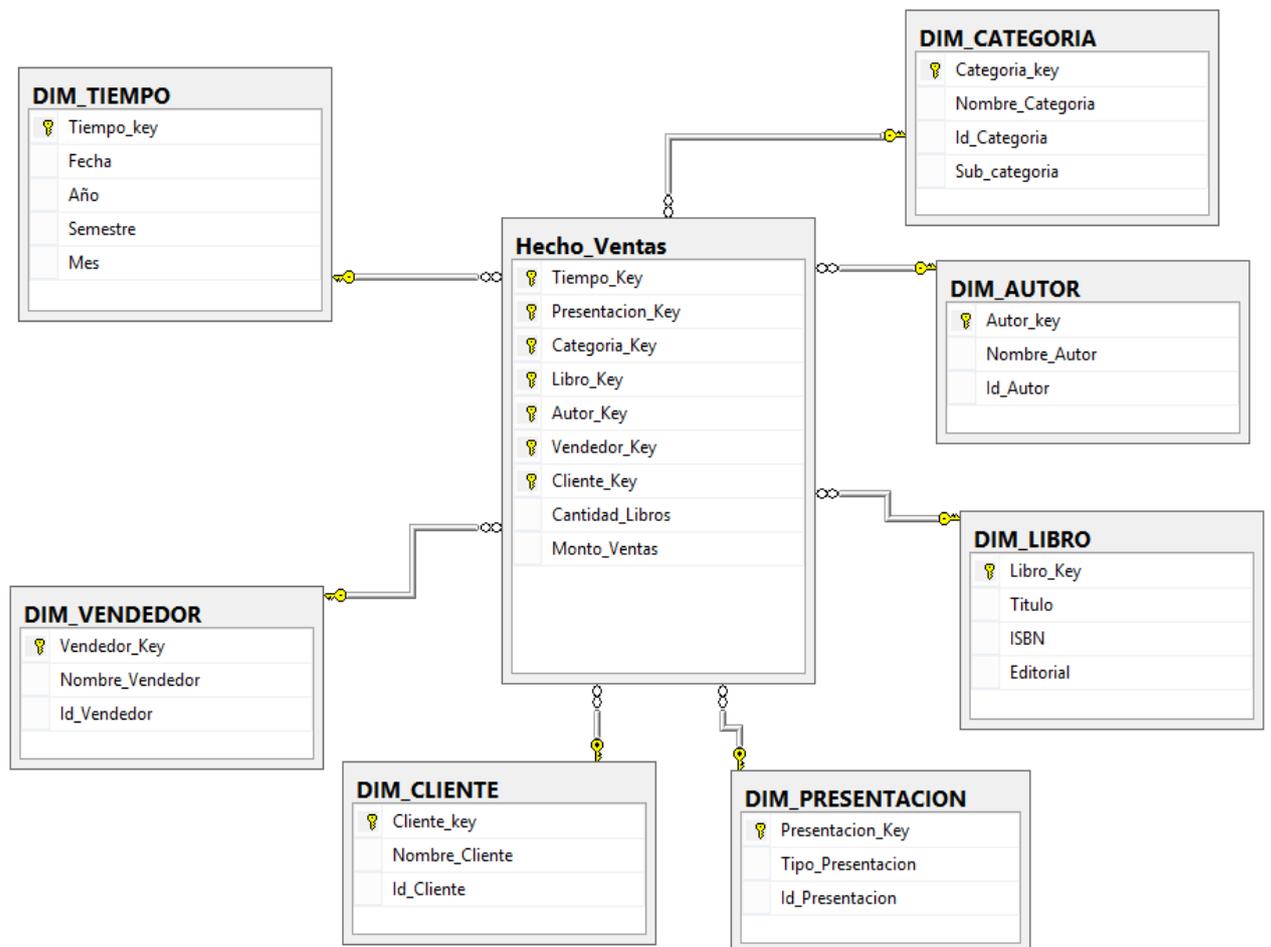
**TABLA: DIM\_PRESENTACION**

CRISOL_TR...DIM_PRESENTACION			
	Column Name	Data Type	Allow Nulls
▶	Presentacion_Key	int	<input type="checkbox"/>
	Tipo_Presentacion	nvarchar(50)	<input type="checkbox"/>
	Id_Presentacion	int	<input type="checkbox"/>

**TABLA: HECHO\_VENTAS**

Column Name	Data Type	Allow Nulls
⚡ Tiempo_Key	int	<input type="checkbox"/>
⚡ Presentacion_Key	int	<input type="checkbox"/>
⚡ Categoria_Key	int	<input type="checkbox"/>
⚡ Libro_Key	int	<input type="checkbox"/>
⚡ Autor_Key	int	<input type="checkbox"/>
⚡ Vendedor_Key	int	<input type="checkbox"/>
⚡ Cliente_Key	int	<input type="checkbox"/>
Cantidad_Libros	int	<input type="checkbox"/>
Monto_Ventas	money	<input type="checkbox"/>

**DIAGRAMA DEL MODELO DIMENSIONAL CONSTRUIDO**



*Figura 18: Diseño Físico de la Base de Datos del BI*

#### 4.6. ETAPA 6: PROCESO DE ETL

Después de haber construido el modelo estrella en la Nube de Microsoft procedemos a poblar cada Tabla de la Base de Datos. Este proceso será realizado usando Integration Service de Microsoft y consta en 3 fases:

- Determinación de los Pasos de Transformación.
- Definición de los Flujos de trabajo.
- Definición y Creación de las transformaciones de datos

##### A. Determinación de los Pasos de Transformación.

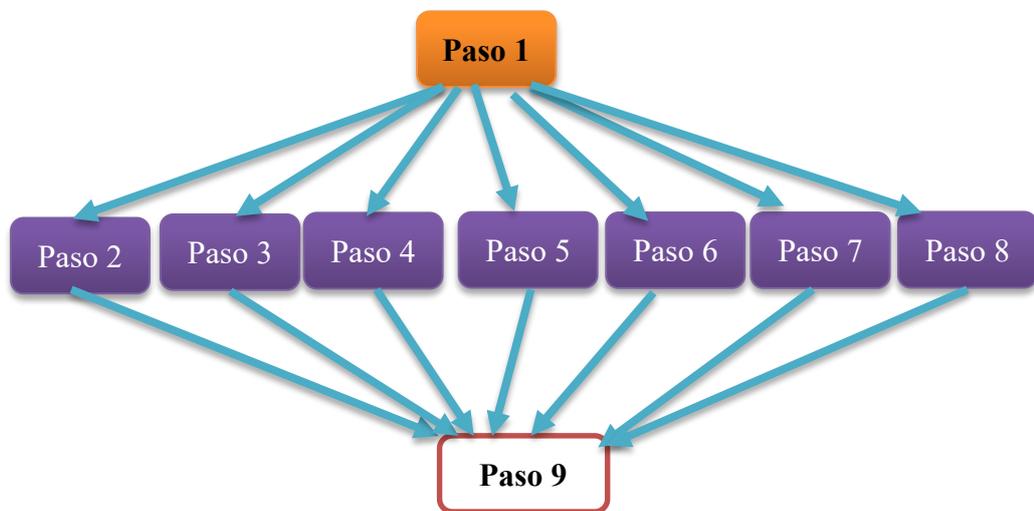
Para un mejor poblamiento de los datos en el modelo dimensional implementado debemos definir los *Pasos de Transformación* en base a una unidad de trabajo.

Para poblar el BI se tiene que realizar los siguientes pasos:

- ✓ **Paso 1:** Limpiar todas las Dimensiones, eliminando los registros de todas las Dimensiones y Tablas de Hechos, para que de esta manera nos podamos asegurar que no existan datos repetidos que afecten la consistencia de la base de datos.
- ✓ **Paso 2:** Poblamiento de la Dimensión Tiempo, consiste en ejecutar una transformación que permita a través de funciones en SQL de extraer cada parte de una fecha de venta y llevar esos datos a la dimensión Tiempo
- ✓ **Paso 3:** Poblamiento de la Dimensión Presentación, consiste en ejecutar una transformación, para transferir datos a la Dimensión Presentación desde la base de datos transaccional.
- ✓ **Paso 4:** Poblamiento de la Dimensión Categoría, consiste en ejecutar una transformación, para transferir datos a la Dimensión Categoría del BI.
- ✓ **Paso 5:** Poblamiento de la Dimensión Cliente, consiste en ejecutar una transformación, para transferir datos a la Dimensión Cliente del BI.
- ✓ **Paso 6:** Poblamiento de la Dimensión Libro, consiste en ejecutar una transformación, para transferir datos a la Dimensión Libro del BI.

- ✓ **Paso 7:** Poblamiento de la Dimensión Vendedor, consiste en ejecutar una transformación, para transferir datos a la Dimensión Vendedor del BI.
- ✓ **Paso 8:** Poblamiento de la Dimensión Autor, consiste en ejecutar una transformación, para transferir datos a la Dimensión Autor del BI.
- ✓ **Paso 9:** Poblamiento de las Tabla de Hechos Ventas.

## B. Definición de los Flujos de Trabajo.



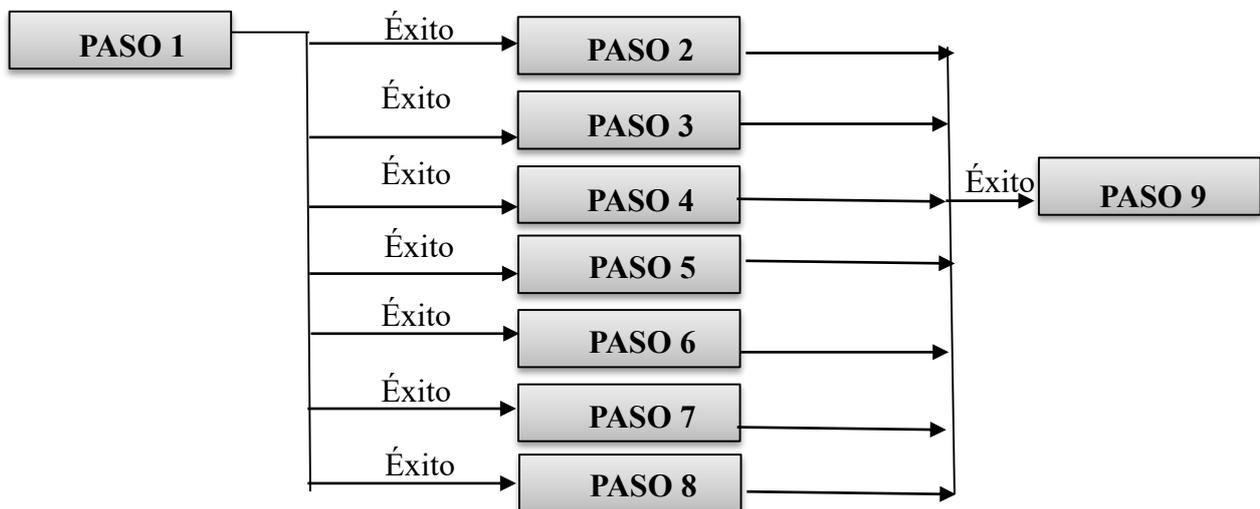
*Figura 29: Workflow de los Pasos de Transformación*

Restricciones de Precedencia:

- a) La Limpieza de las Dimensiones (Paso 1) debe realizarse al inicio del proceso.
- b) El Poblamiento de la Dimensión Tiempo (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 Presentación (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 Categoría (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 Cliente (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 Libro (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 Vendedor (Paso 7) debe realizarse sólo cuando se tenga la seguridad de que el paso 1 se ha ejecutado con éxito.
- h) El Poblamiento de la Dimensión Autor (Paso 8) debe realizarse sólo cuando se tenga la seguridad de que el paso 1 se ha ejecutado con éxito
- i) El Poblamiento de las Tablas de Hechos Ventas (Paso 9) debe realizarse sólo cuando se tenga la seguridad de que el paso 2 al 8 se hallan ejecutado con éxito.

Después de haber descrito las relaciones de precedencias diagramaremos el flujo de trabajo que se necesitará realizar para construir el paquete ETL de poblamiento del BI, para el proceso de Ventas de la empresa.



*Figura 20: Diagrama Workflow con Restricciones de Precedencia.*

### C. Creación de los paquetes de Servicio de Transformación de Datos (DTS).

Luego de haber definido las precedencias o pasos se procede a construir el paquete DTS o la estructura que tendrá la solución.

Para su implementación se usó el Integration Services basándonos en Workflow con restricciones de la figura anterior, creamos un Control Flow cual llamaremos “Poblamiento” para realizar el poblamiento del BI de ventas.

#### DIAGRAMA DEL PROCESO ETL

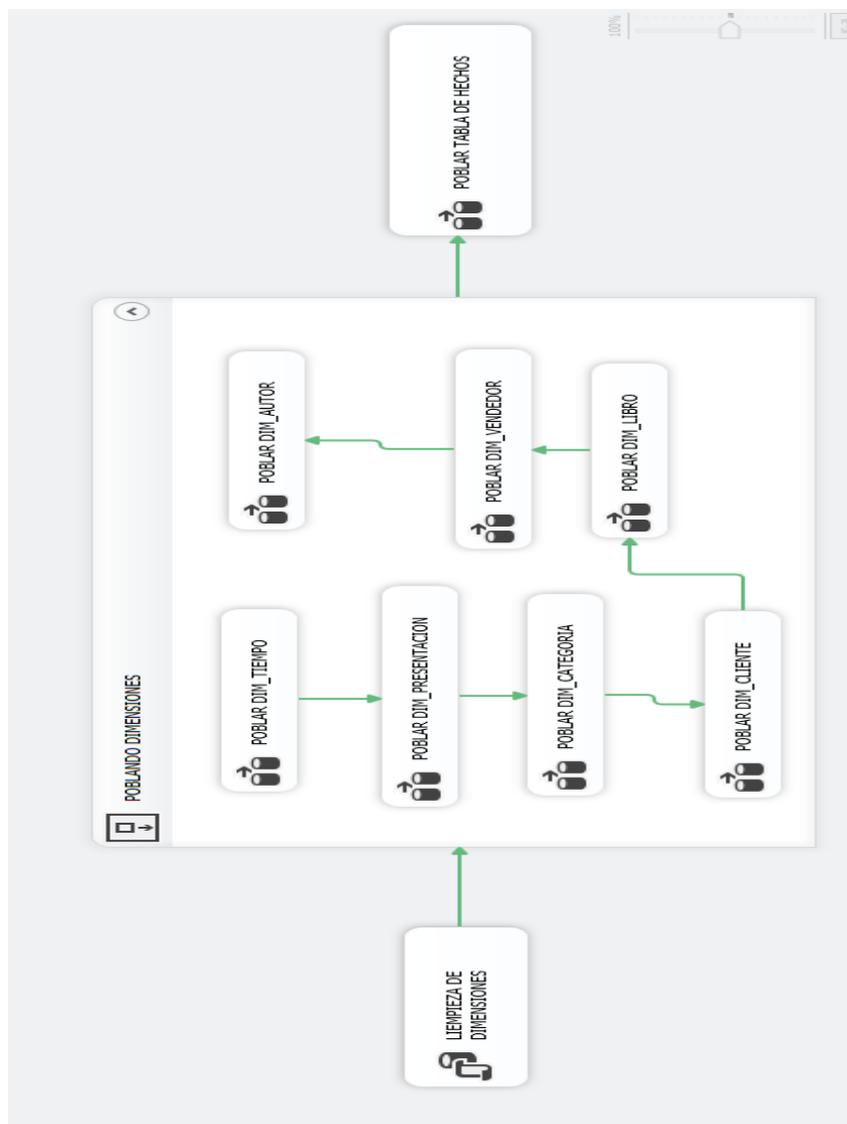
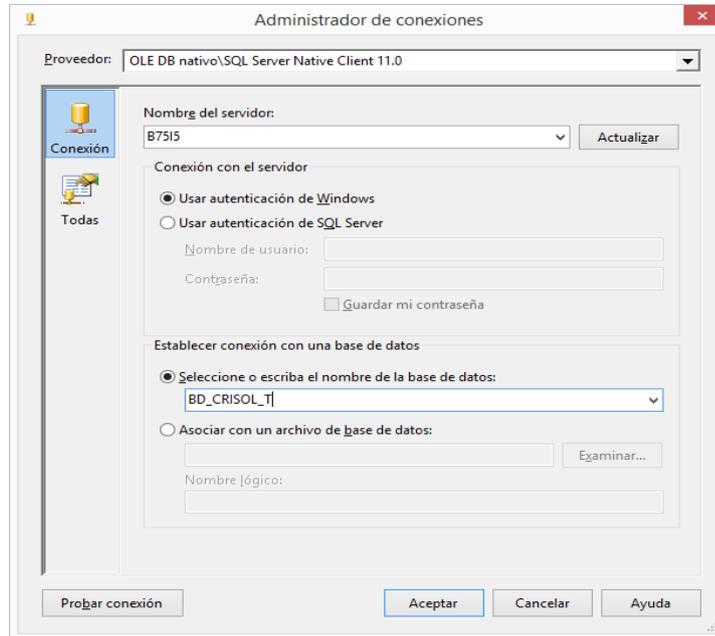


Figura 21: ETL de la solución de BI

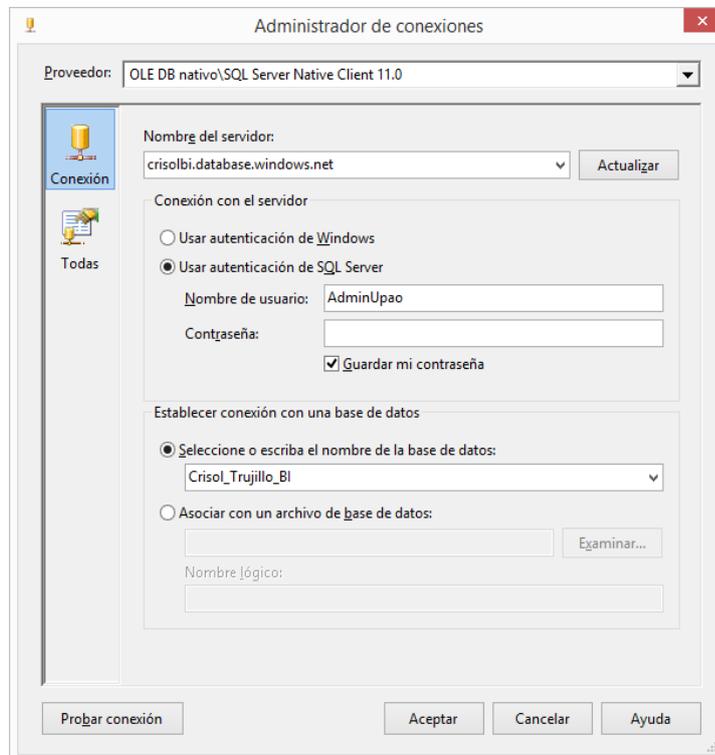
## DESCRIPCION DE CADA PASO DEL ETL:

Se trabajará en base a dos conexiones:

- **BD TRANSACCIONAL.**

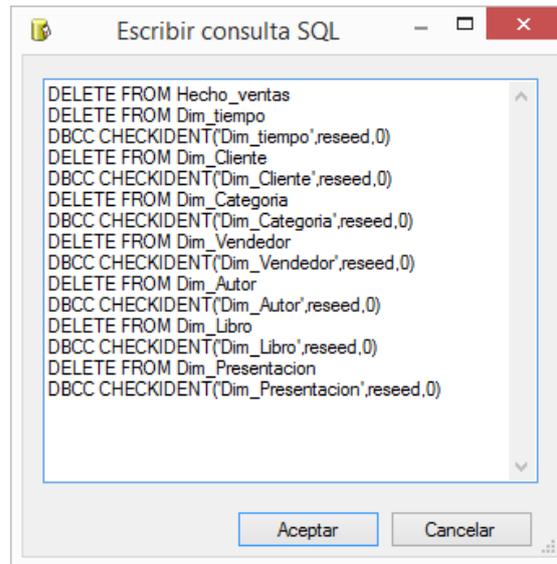


- **BI EN LA NUBE DE MICROSOFT AZURE.**



✓ **PASO 1: LIMPIAR TODAS LAS DIMENSIONES**

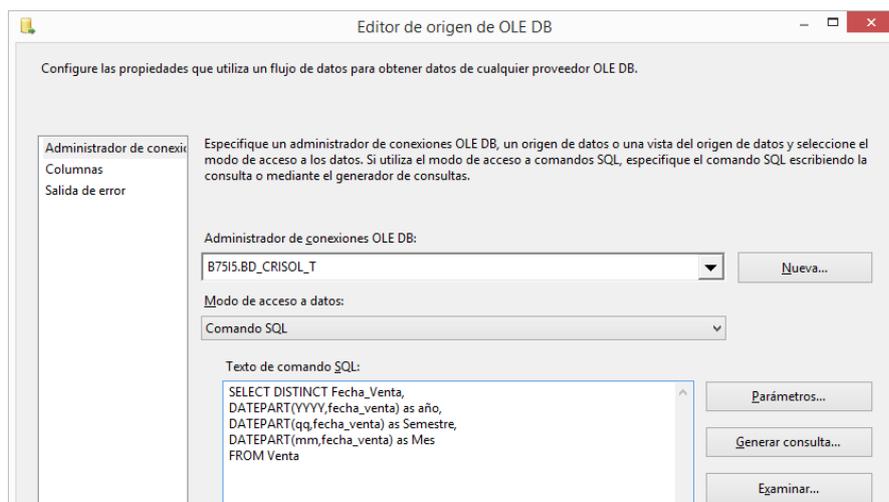
Para este paso se agrega una tarea al paquete que tendrá por nombre “Limpieza de dimensiones”. Esta tarea nos permite limpiar los datos de todas las tablas de nuestra solución de BI. Esta limpieza asegura que no se dupliquen los datos. Por lo que la ejecutamos con la siguiente sentencia:



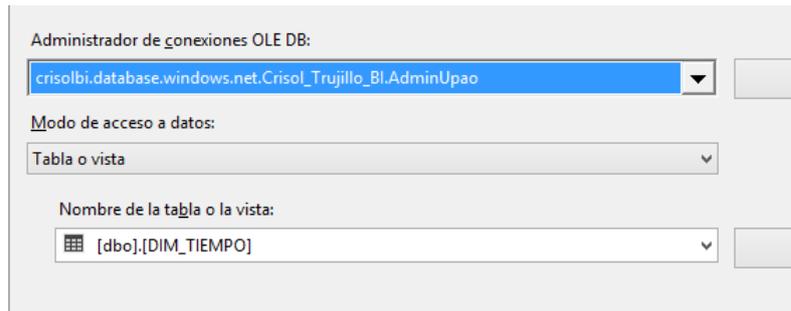
✓ **PASO 2: POBLAMIENTO DE LA DIMENSIÓN TIEMPO**

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

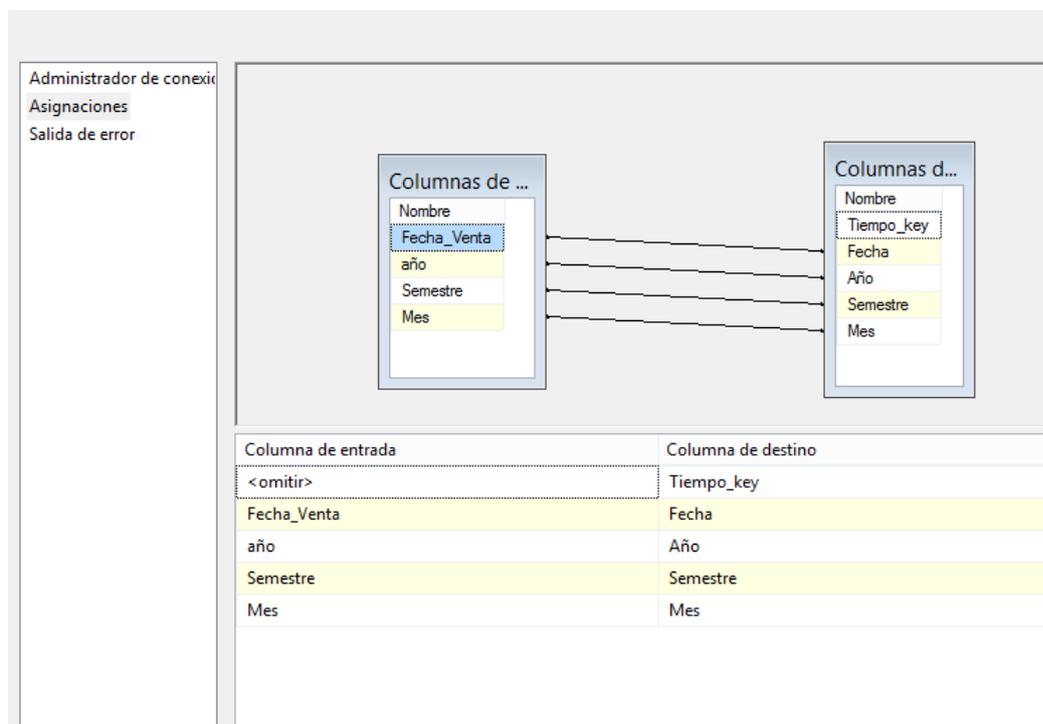
- a. Se define una sentencia SQL como origen de datos de la Base de Datos Operacional.



- b. Determinamos la Dimensión Tiempo como destino de Datos.



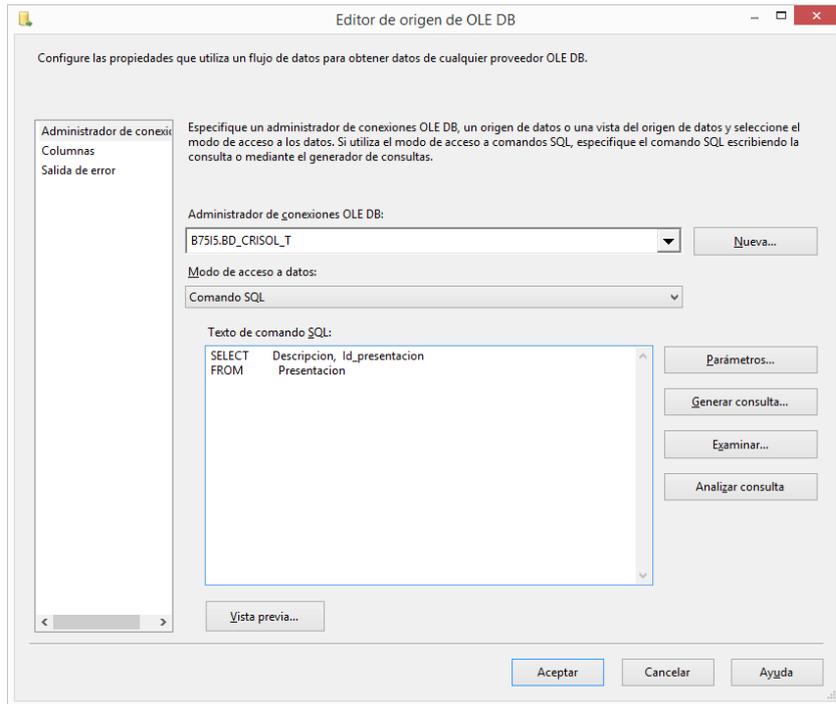
- c. En las transformaciones se determina que columnas del origen corresponderán con las del destino.



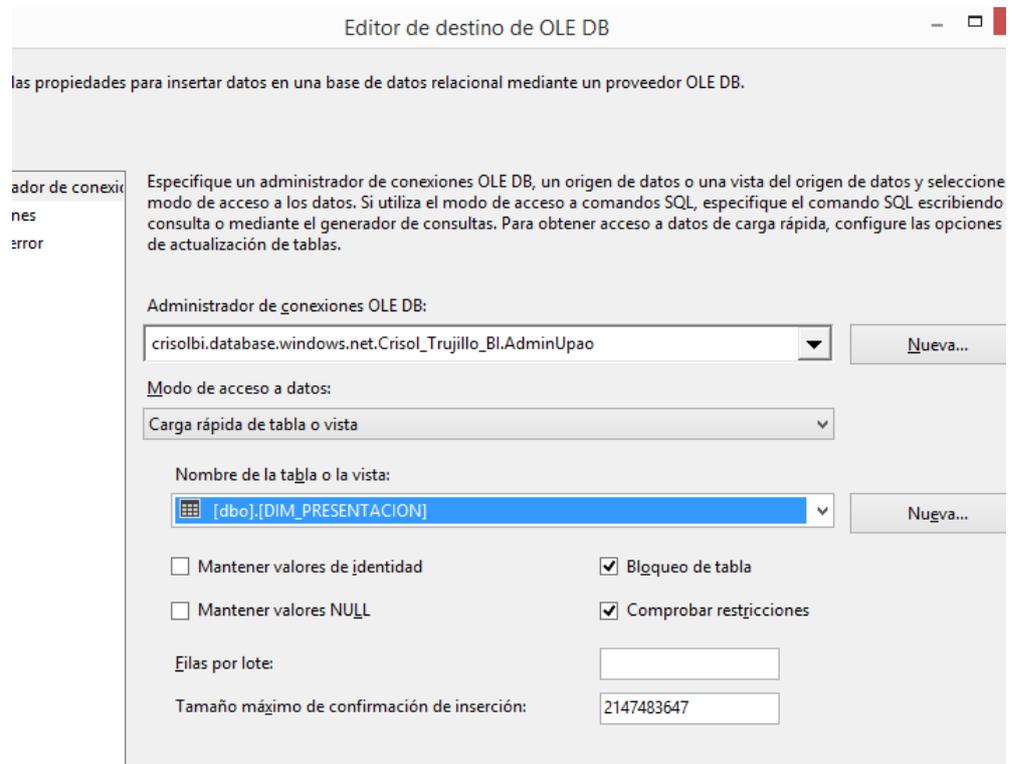
✓ **PASO 3: POBLAMIENTO DE LA DIMENSIÓN PRESENTACIÓN**

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

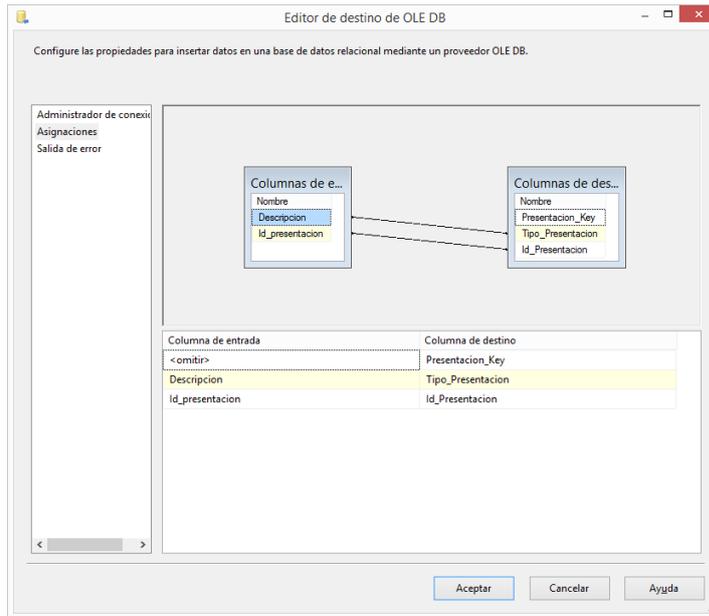
- a. Se define una sentencia SQL como origen de datos de la Base de Datos Operacional.



b. Determinamos la Dimensión presentación como destino de Datos.



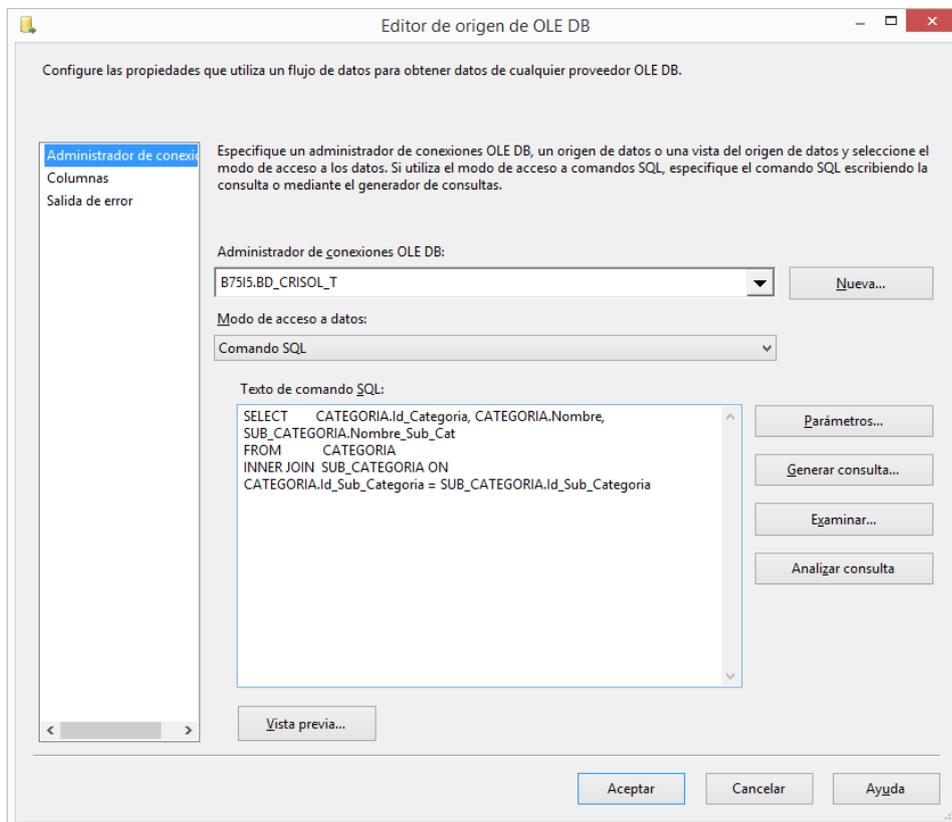
c. En las transformaciones se determina que columnas del origen corresponderán con las del destino.



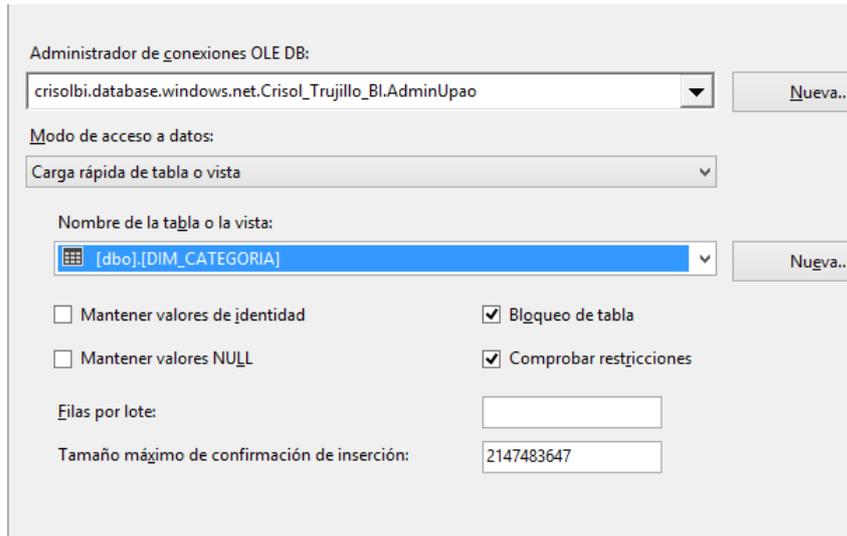
✓ **PASO 4: POBLAMIENTO DE LA DIMENSIÓN CATEGORÍA.**

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

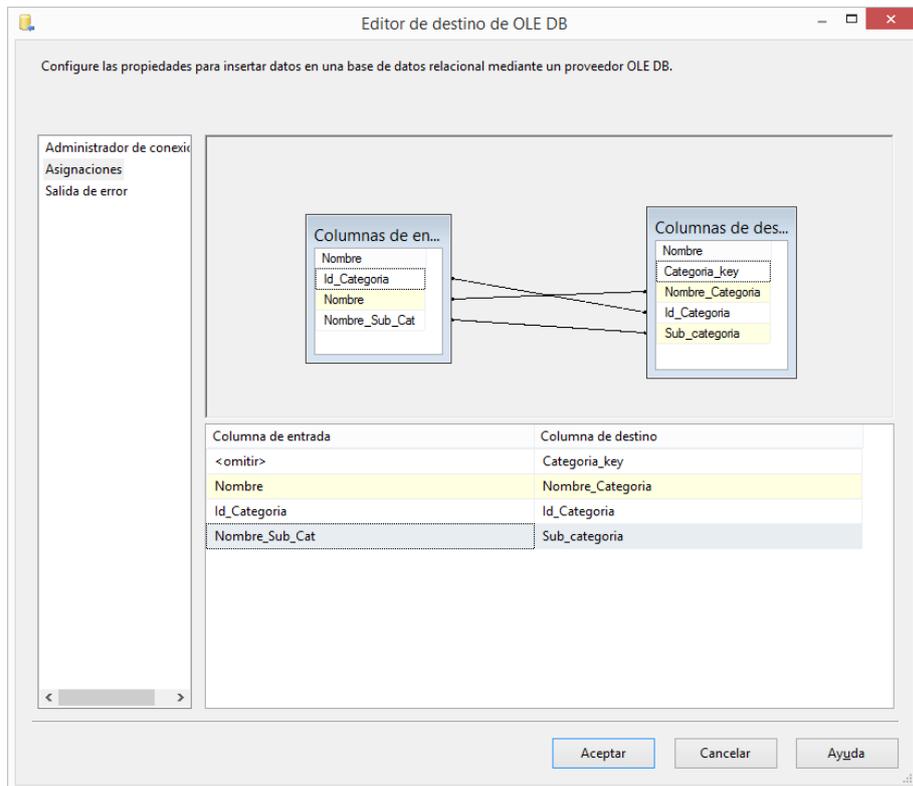
- a. Se define una sentencia SQL como origen de datos de la Base de Datos Operacional.



- b. Determinamos la Dimensión Categoría como destino de Datos.



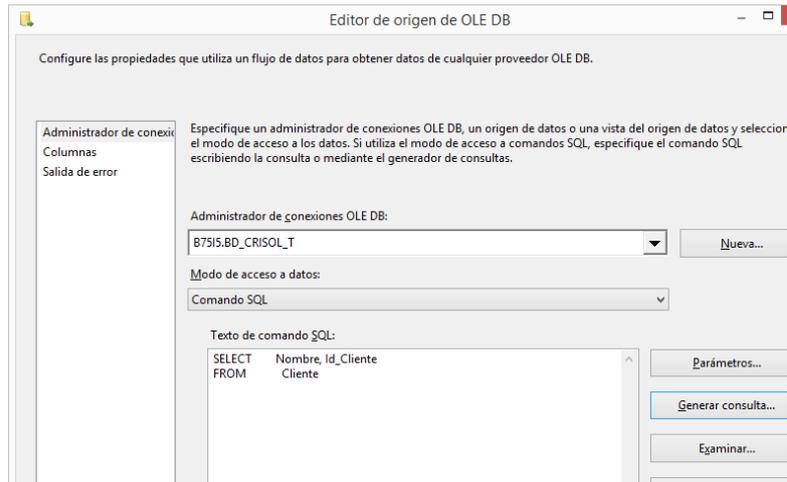
- c. En las transformaciones se determina que columnas del origen corresponderán con las del destino.



✓ **PASO 5: POBLAMIENTO DE LA DIMENSIÓN CLIENTE.**

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

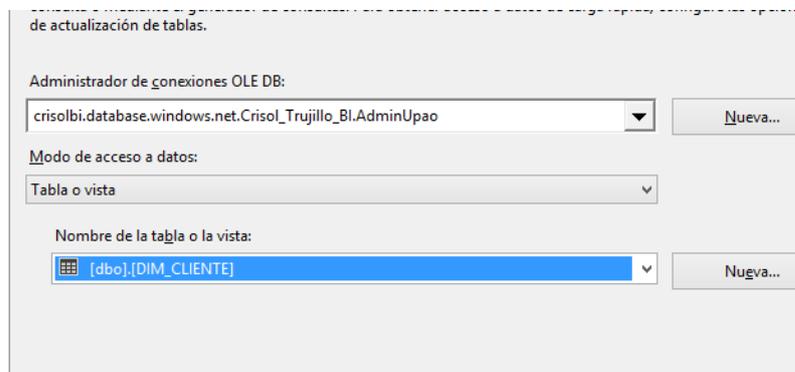
- a. Se define una sentencia SQL como origen de datos de la Base de Datos Operacional.



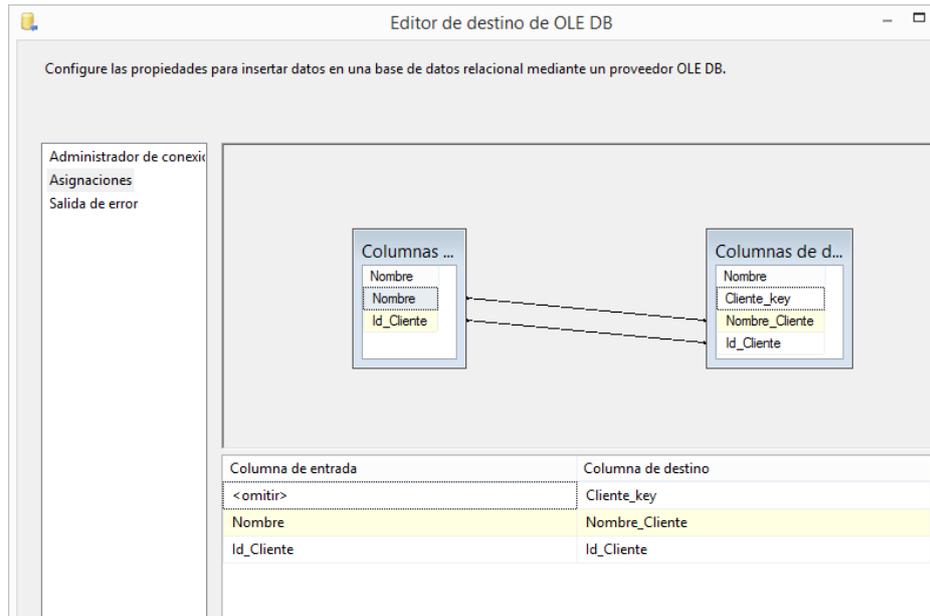
## b. INTEGRACION DE DATOS DE CLIENTES



## c. Determinamos la Dimensión Cliente como destino de Datos.



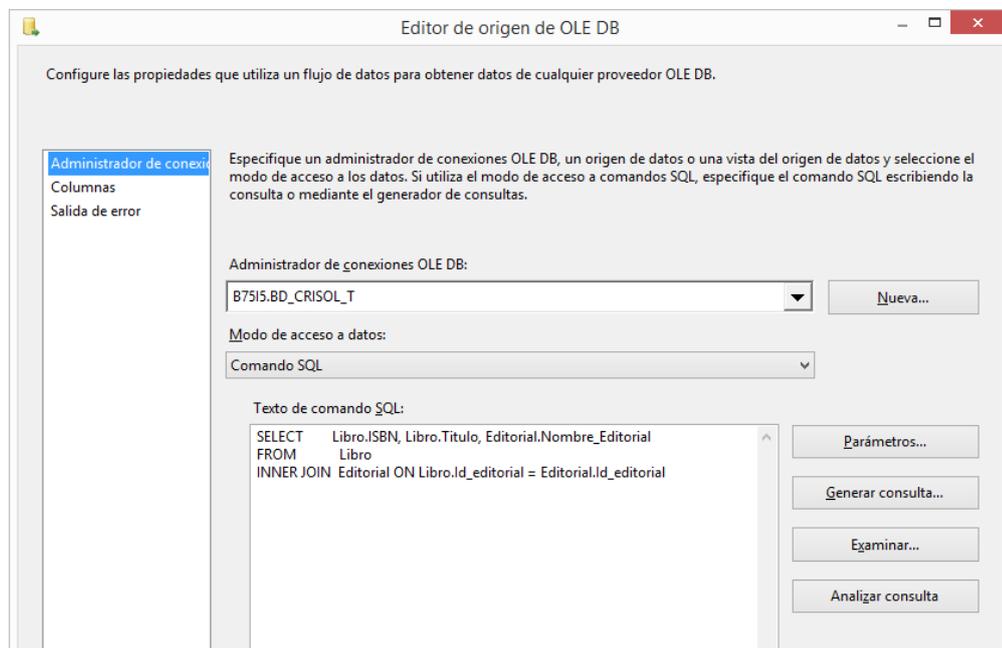
## d. En las transformaciones se determina que columnas del origen corresponderán con las del destino.



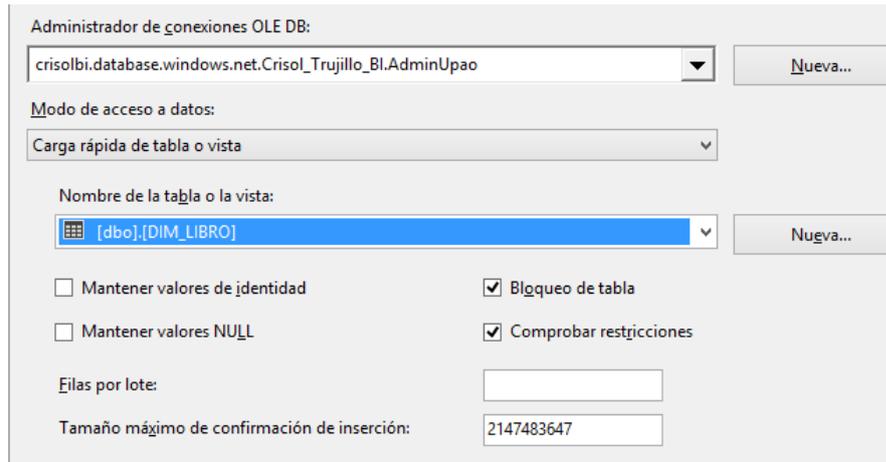
✓ **PASO 6: POBLAMIENTO DE LA DIMENSIÓN LIBRO.**

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

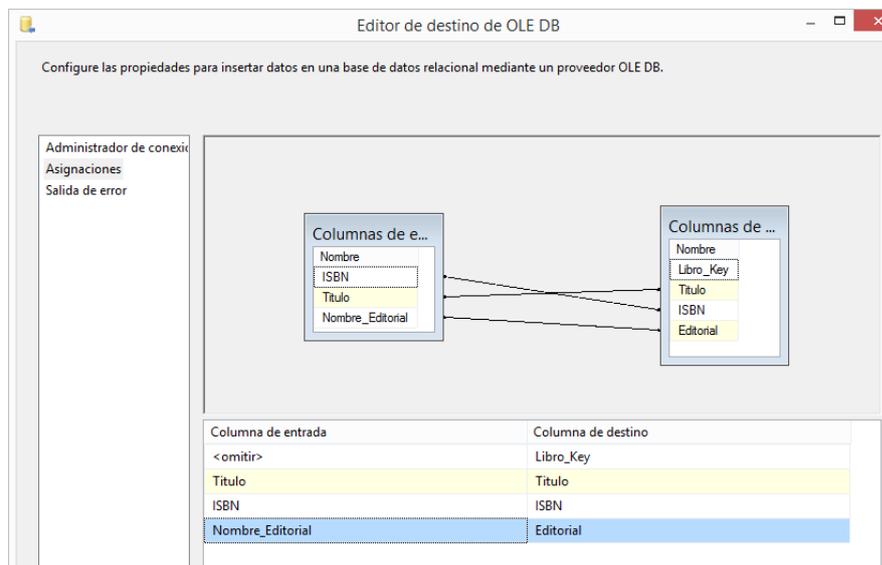
- a. Se define una sentencia SQL como origen de datos de la Base de Datos Operacional.



- b. Determinamos la Dimensión Libro como destino de Datos.



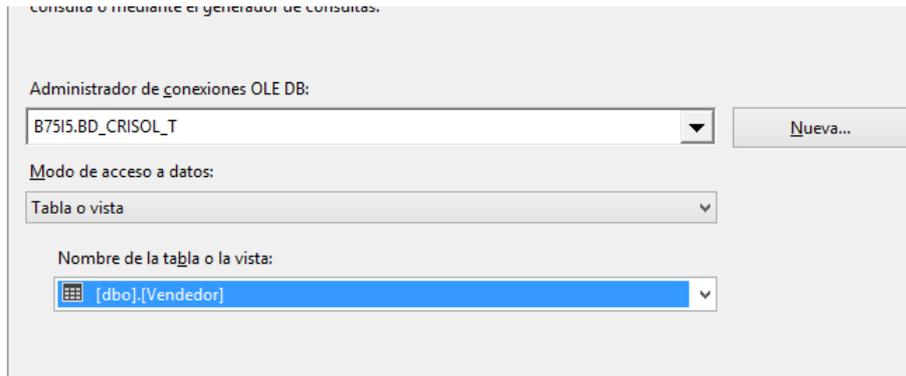
- c. En las transformaciones se determina que columnas del origen corresponderán con las del destino.



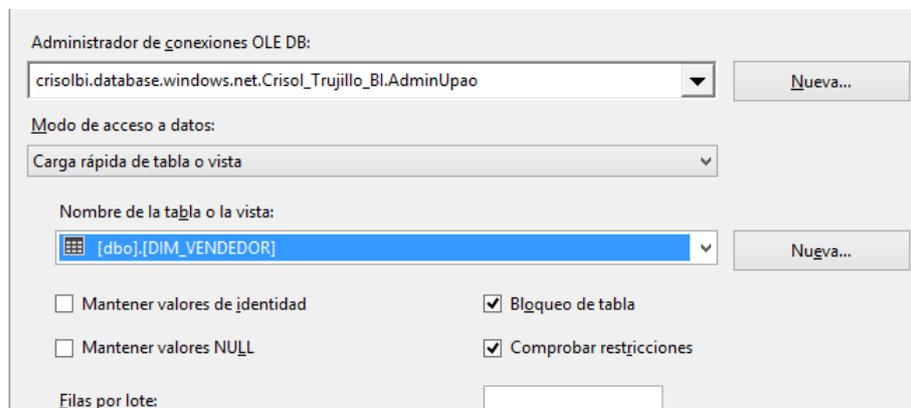
✓ **PASO 7: POBLAMIENTO DE LA DIMENSIÓN VENDEDOR.**

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

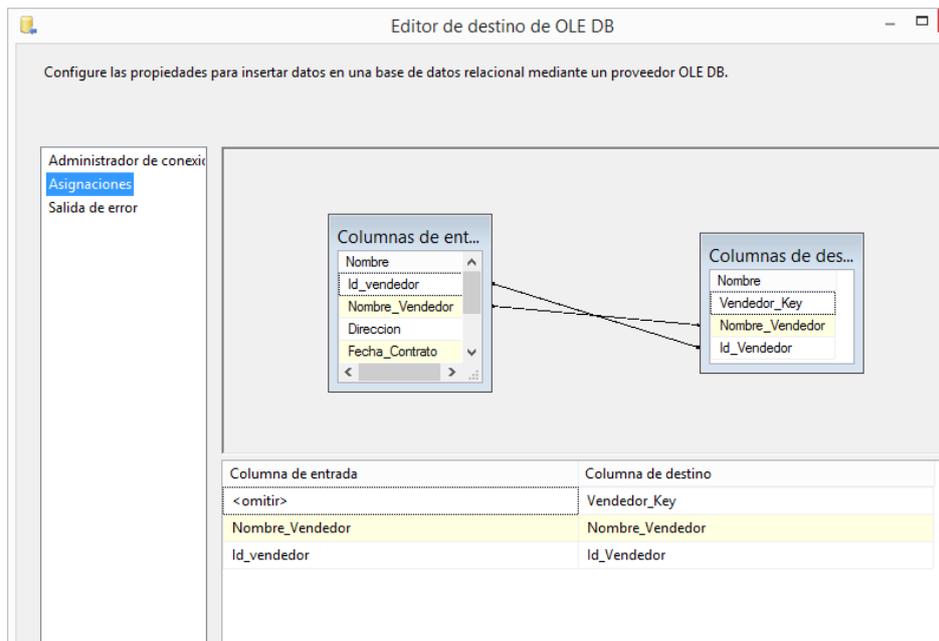
- a. Se define una sentencia SQL como origen de datos de la Base de Datos Operacional.



b. Determinamos la Dimensión Vendedor como destino de Datos.



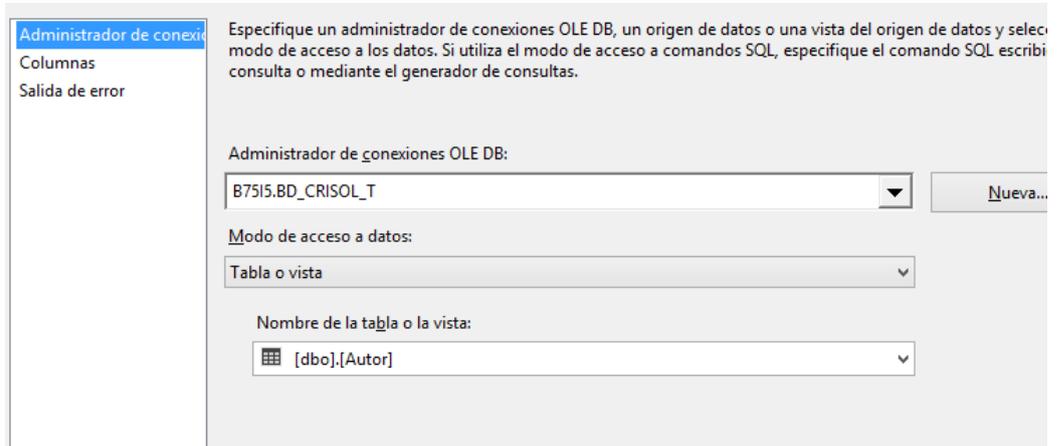
c. En las transformaciones se determina que columnas del origen corresponderán con las del destino.



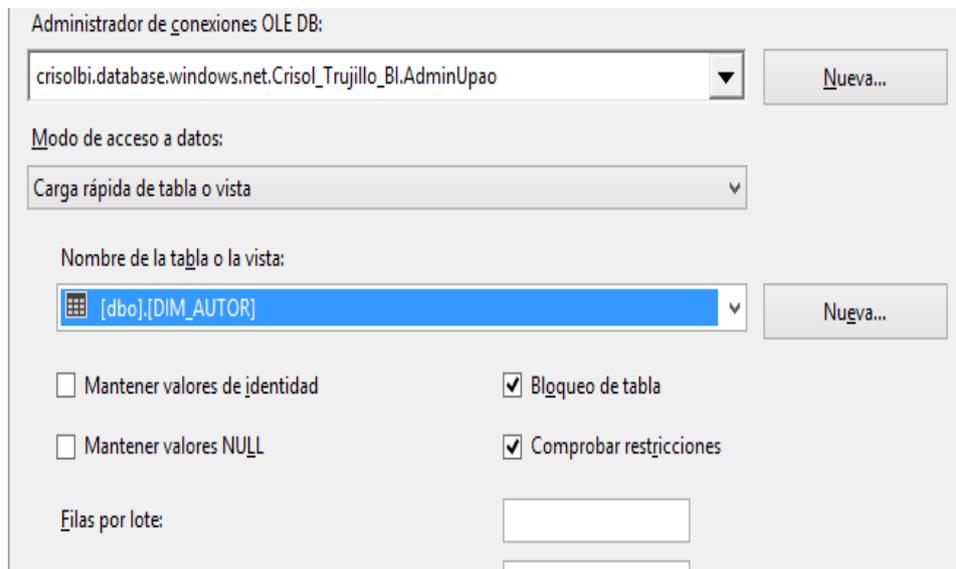
✓ **Paso 8: Poblamiento de la Dimensión Autor.**

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

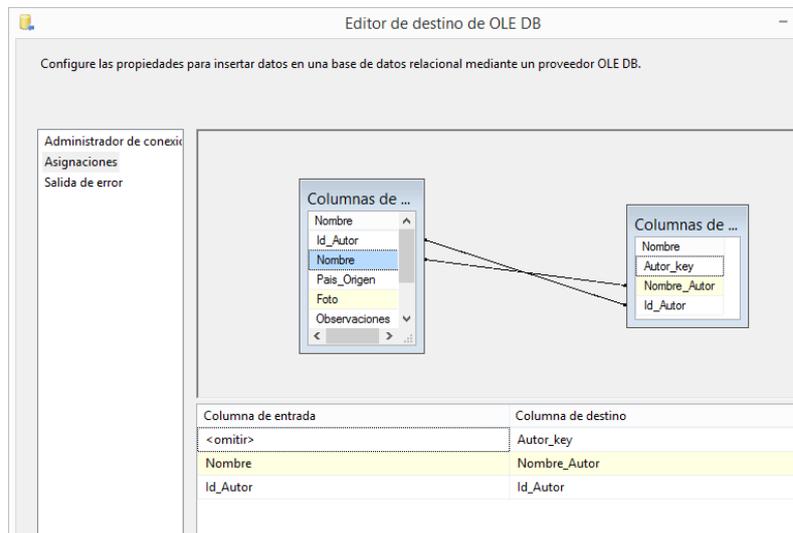
- a. Se define una sentencia SQL como origen de datos de la Base de Datos Operacional.



- b. Determinamos la Dimensión Autor como destino de Datos.



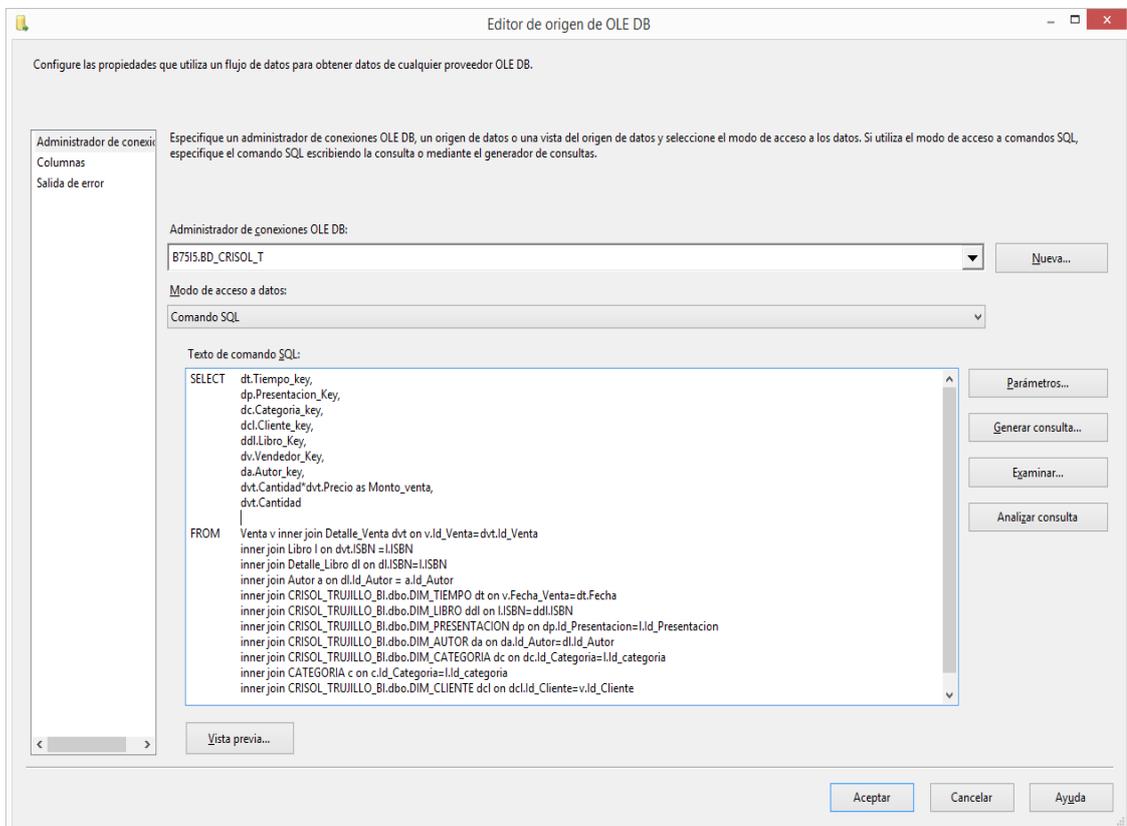
- c. En las transformaciones se determina que columnas del origen corresponderán con las del destino.



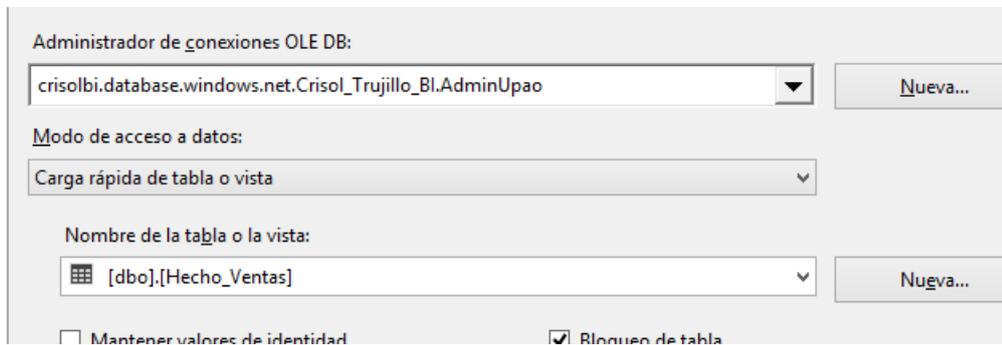
✓ **Paso 9: Poblamiento de las Tabla de Hechos Ventas.**

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

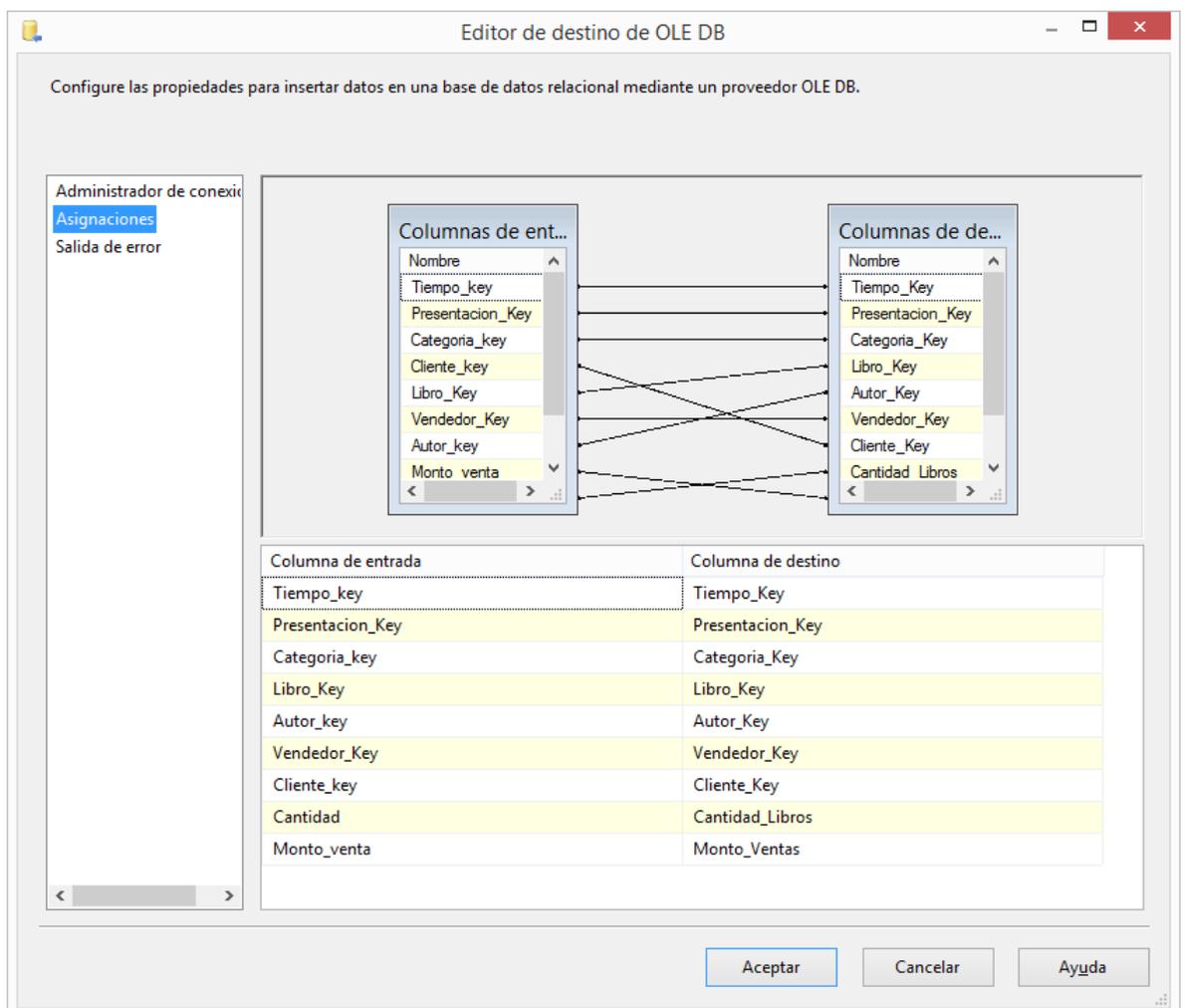
- a. Se define una sentencia SQL como origen de datos de la Base de Datos Operacional.



b. Determinamos la tabla de HechoVentas como destino de Datos.



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



## 4.7. ETAPA 7: SELECCIÓN DE PRODUCTOS

### 4.7.1. Hardware

Solo se mantendrá un Hardware donde se aloja la base de datos Transaccional:

- ✓ Servidor HP ProLiant ML30 G9 Torre 4U, Intel Xeon E3-1220v6 Quad Core (3.00GHz 8MB) - 8GB DDR4 - HDD 1TB SATA - DVD-RW

### 4.7.2. Software

Para el desarrollo de todo este trabajo se consideró los siguientes productos:

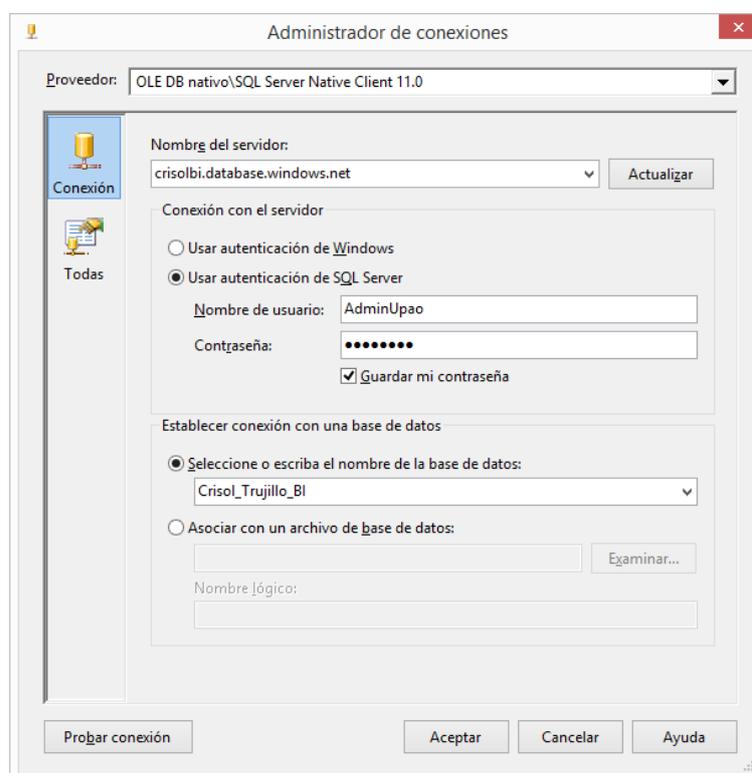
- ✓ MS SQL Server 2014
- ✓ Microsoft Azure
- ✓ Qlik Sense

## 4.8. ETAPA 8: ESPECIFICACIÓN DE LA APLICACIÓN DEL USUARIO FINAL

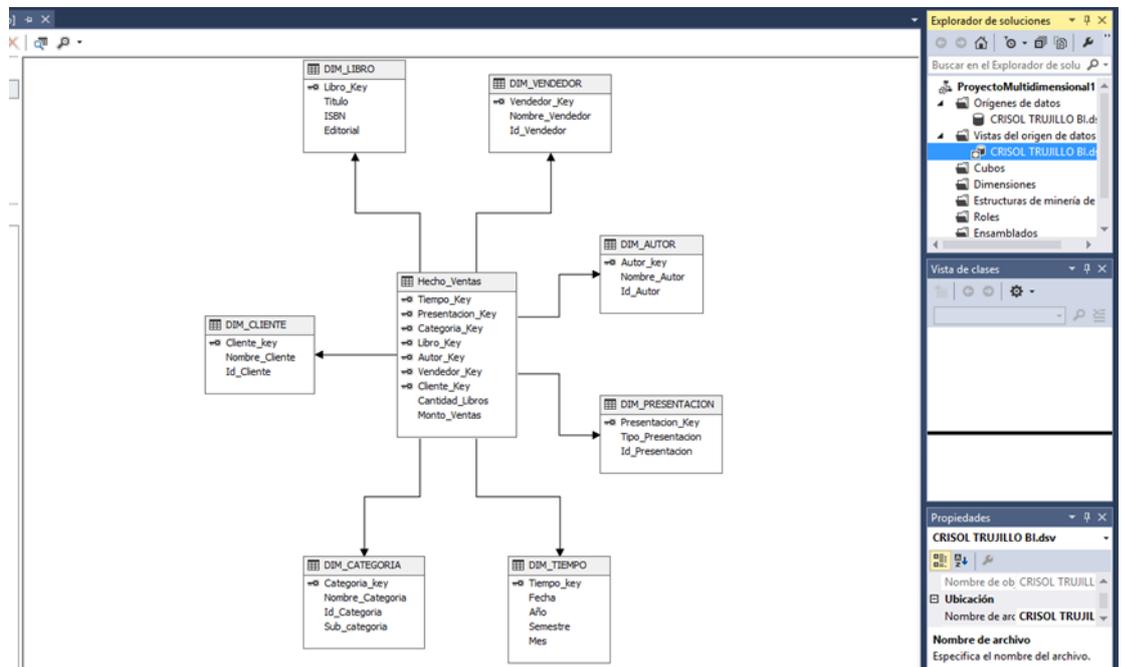
### 4.8.1. Estructura de Cubo

A continuación, se detalla paso a paso el desarrollo del cubo para el análisis.

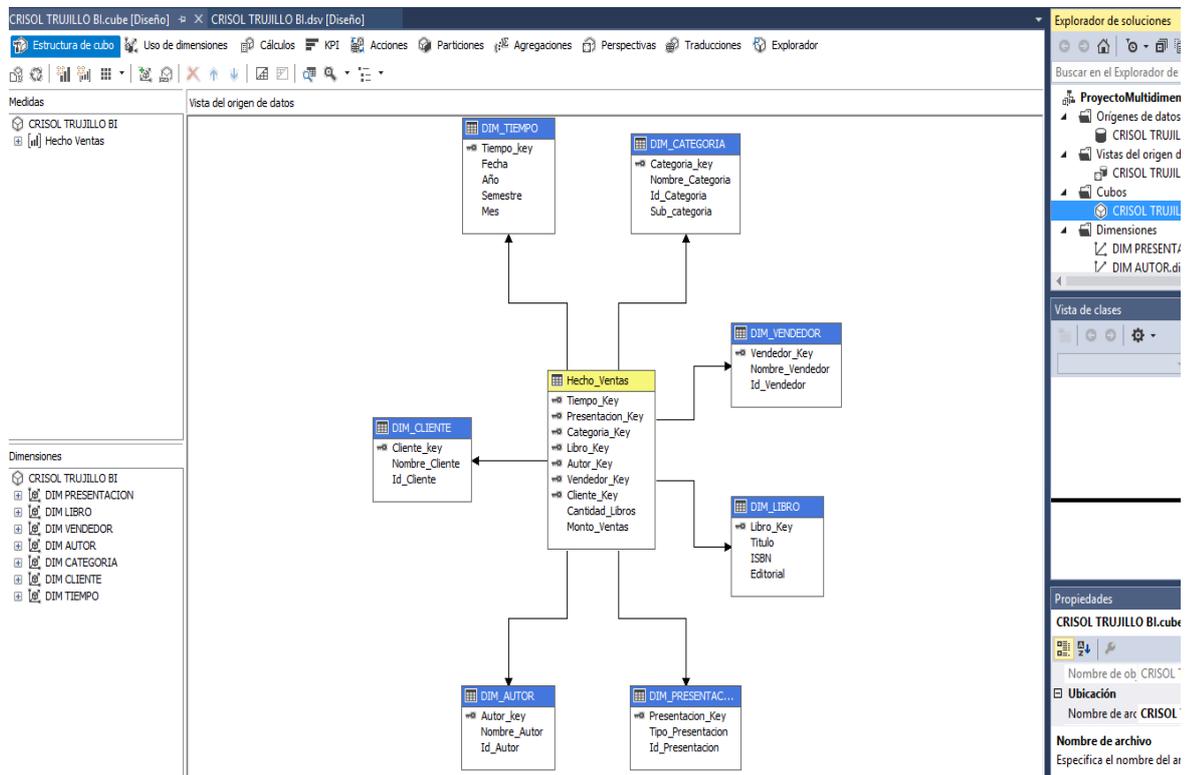
#### Paso 1: Creando la conexión a la base de datos en la Nube



## Paso 2: Creando la vista de datos



## Paso 3: Creando un nuevo cubo y añadiendo la tabla de hechos



## Paso 4: Creando las jerarquías

### JERARQUIA DIMENSION TIEMPO:

The screenshot shows the SQL Server Enterprise Manager interface for creating a hierarchy for the DIM TIEMPO dimension. The interface is divided into three main panes: Atributos, Jerarquías, and Vista del origen de datos.

- Atributos:** Lists the attributes for DIM TIEMPO: Año, Fecha, Mes, Semestre, and Tiempo Key.
- Jerarquías:** Shows a hierarchy named 'Jerarquía' with levels: Año, Semestre, Mes, and Fecha. A note states: "Para crear un nuevo atributo, arrastre un atributo hasta aquí."
- Vista del origen de datos:** Shows the data source view for DIM\_TIEMPO, listing the attributes: Tiempo\_key, Fecha, Año, Semestre, and Mes.

### JERARQUIA DIMENSION CATEGORIA:

The screenshot shows the SQL Server Enterprise Manager interface for creating a hierarchy for the DIM CATEGORIA dimension. The interface is divided into three main panes: Atributos, Jerarquías, and Vista del origen de datos.

- Atributos:** Lists the attributes for DIM CATEGORIA: Categoría Key, Nombre Categoría, and Sub Categoría.
- Jerarquías:** Shows a hierarchy named 'CAT-SUBCAT' with levels: Nombre Categoría, Sub Categoría, and a placeholder '<nuevo nivel>'. A note states: "Para crear un nuevo atributo, arrastre un atributo, arrastre un atributo, arrastre un atributo hasta aquí."
- Vista del origen de datos:** Shows the data source view for DIM\_CATEGORIA, listing the attributes: Categoría\_key, Nombre\_Categoría, Id\_Categoría, and Sub\_categoria.

### JERARQUIA DIMENSION LIBRO:

The screenshot shows the SQL Server Enterprise Manager interface for creating a hierarchy for the DIM LIBRO dimension. The interface is divided into three main panes: Atributos, Jerarquías, and Vista del origen de datos.

- Atributos:** Lists the attributes for DIM LIBRO: Editorial, Libro Key, and Título.
- Jerarquías:** Shows a hierarchy named 'EDITORIAL - LIBRO' with levels: Editorial, Título, and a placeholder '<nuevo nivel>'. A note states: "Para crear un nuevo atributo, arrastre un atributo hasta aquí."
- Vista del origen de datos:** Shows the data source view for DIM\_LIBRO, listing the attributes: Libro\_Key, Título, ISBN, and Editorial.

## 4.9. ETAPA 9: DESARROLLO DE LA APLICACIÓN DEL USUARIO FINAL

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

La herramienta a usar para el procesamiento analítico será:

- ✓ Qlik Sense



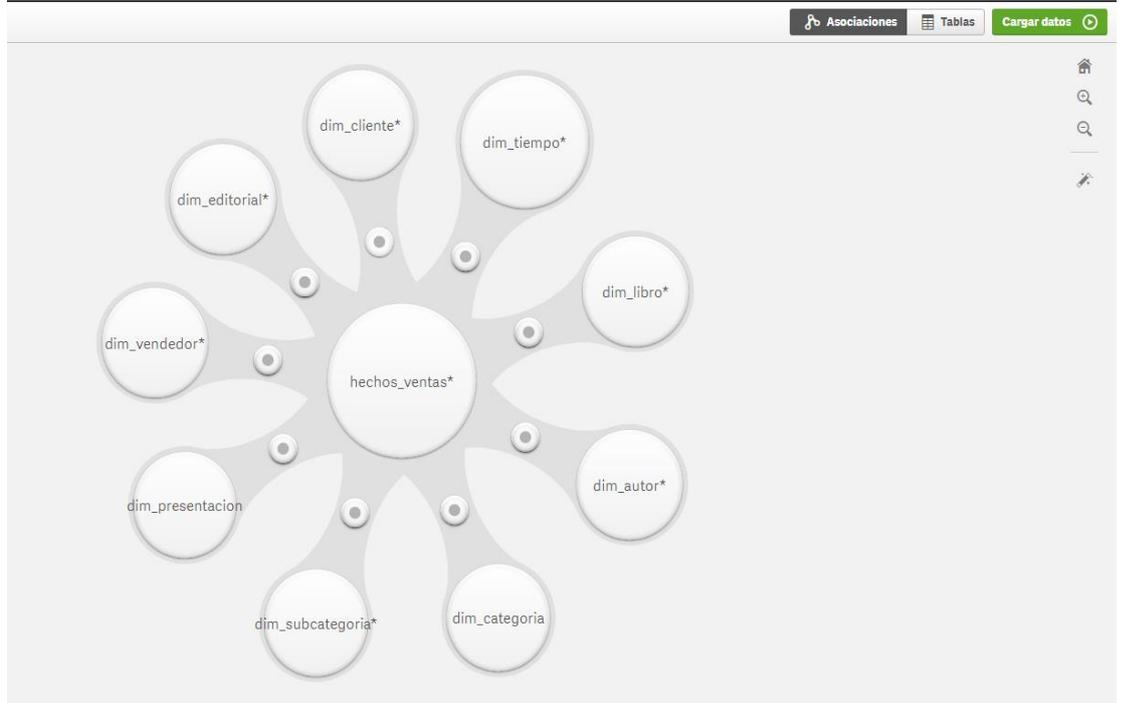
### 4.9.2. Lista los resúmenes de información requeridos por los usuarios.

Los requerimientos elegidos por el gerente son:

- ✓ ¿Cuál es el monto de un determinado cliente en compras en un mes?
- ✓ ¿Cuál es el Monto de ventas por categoría en un mes determinado?
- ✓ ¿Cuál es el monto en ventas por sub categorías en un mes determinado?
- ✓ ¿Cuál es el monto de ventas por editorial y por mes?
- ✓ ¿Cuál es el monto que generan las ventas de un libro por mes y año?
- ✓ ¿Cuál es la cantidad de libros vendidos por categoría, sub categoría por mes y semestre?
- ✓ ¿Cuál es la cantidad de Productos vendidos por un vendedor al mes y al año?
- ✓ ¿Cuál es el monto que generan las ventas de libros por tipo de presentación en un mes, semestre o año?
- ✓ Listar a los Vendedores y el monto total de cada una de las ventas que han generado
- ✓ ¿Cuál cantidad de ventas de libros por autor, por mes y año?

### 4.9.3. Implementación de los reportes en Qlik Sense.

- **Conexión a Qlik Sense**



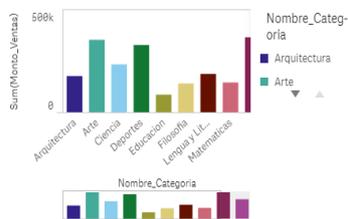
#### R1. ¿Cuál es el monto de un determinado cliente en compras en un mes?

##### Librería Ventas

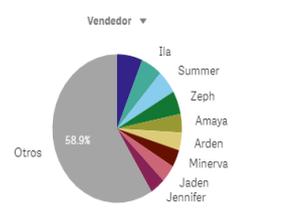
Q Cliente	Q Vendedor	Q Editorial	Q Libro	Q Autor	Q Categoría	Q Subcategoría	Q Presentación
Adele	Alexander	A Enim Ltd	Aliquet Phasellus Foundati...	Acton	Arquitectura	Aerobics	Tapa Blanda
Almee	Amaya	A Facilis Non Company	Amet Corporation	Ahmed	Arte	Algebra	Tapa Delgada
Alexis	Anne	Ac Fermentum Industries	Amet LLP	Aimee	Ciencia	Animacion	Tapa Deluxe
Alika	Arden	Ac Mattis Consulting	Arcu Vel Inc.	Akeem	Deportes	Asiático	Tapa Dura
Alisa	Benjamin	Adipiscing Ligula Aenean PC	Auctor Nunc Nulla LLP	Ali	Educacion	Baloncesto	Tapa Gruesa
Allen	Beverly	Aenean Associates	Congue In Corporation	Beau	Filosofía	Béisbol	
Alma	Brenda	Aenean Gravida Nunc LLC	Convallis Foundation	Benedict	Lengua y Literatura	Biología	

Q Año	Q Mes	Q Semestre	Monto de Venta	Cantidad de Libros
2000		1	1	Sum(Monto_Ventas)
2001		2	2.26M	Sum(Cantidad_Libros)
				46.2k

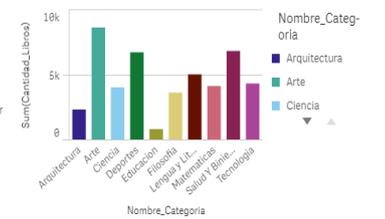
Categoría por el Monto de venta



Ventas del Vendedor

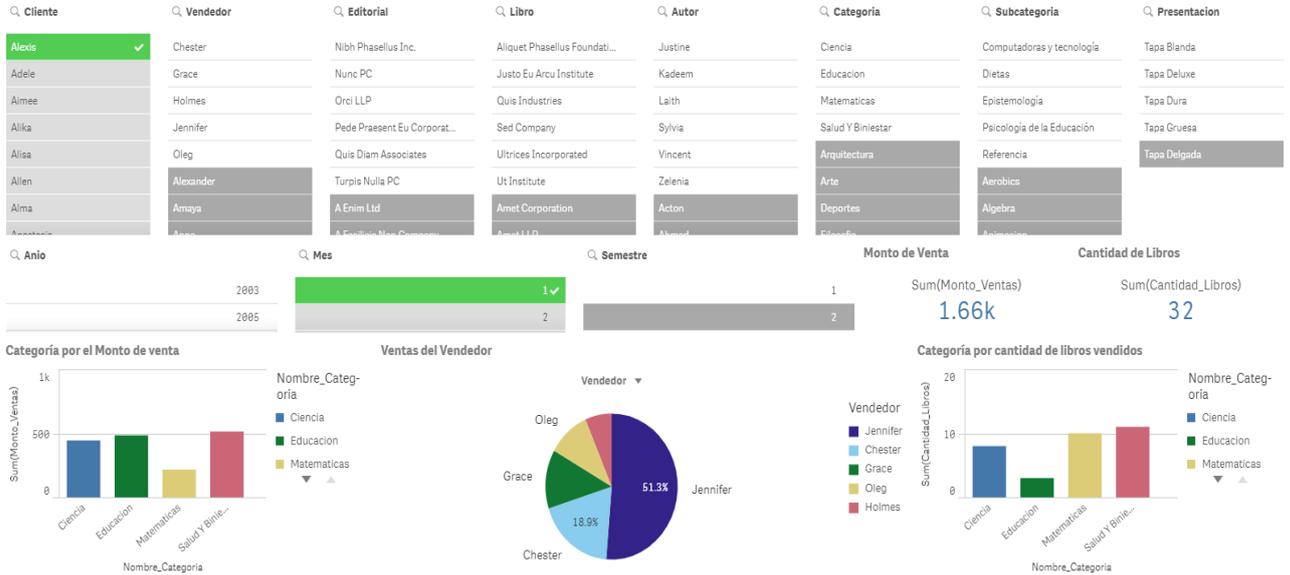


Categoría por cantidad de libros vendidos



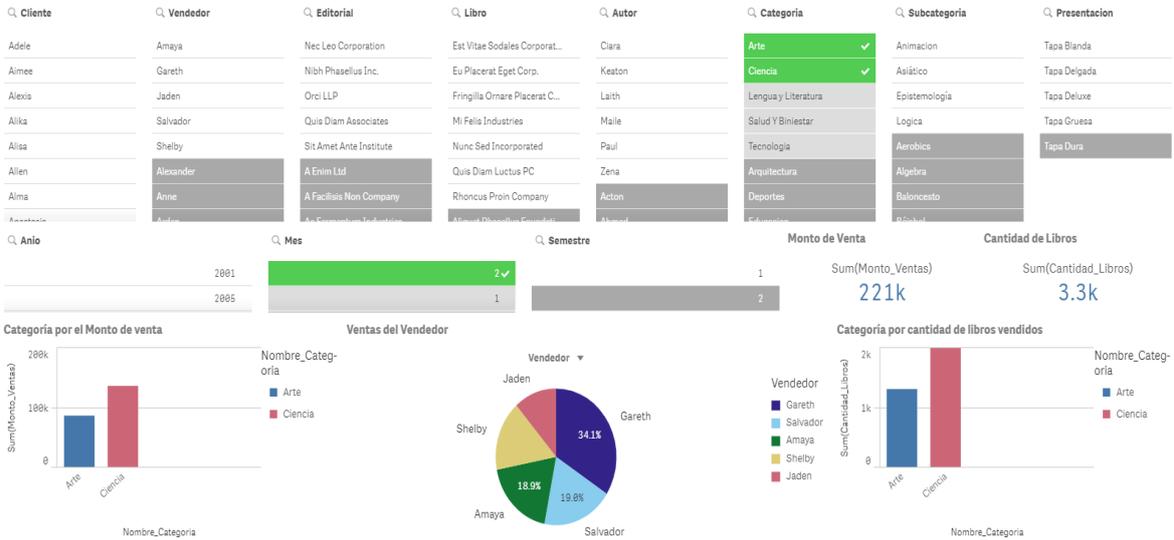
## R2. ¿Cuál es el Monto de ventas por categoría en un mes determinado?

### Librería Ventas



## R3. ¿Cuál es el monto en ventas por sub categorías en un mes determinado?

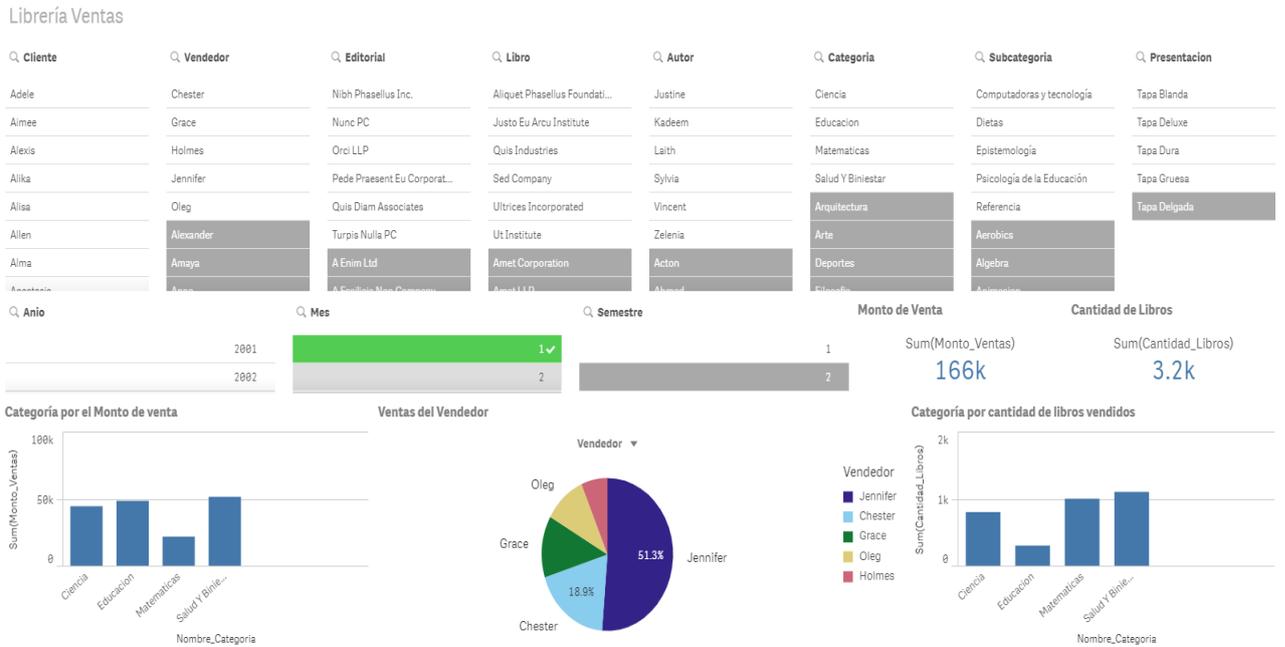
### Librería Ventas



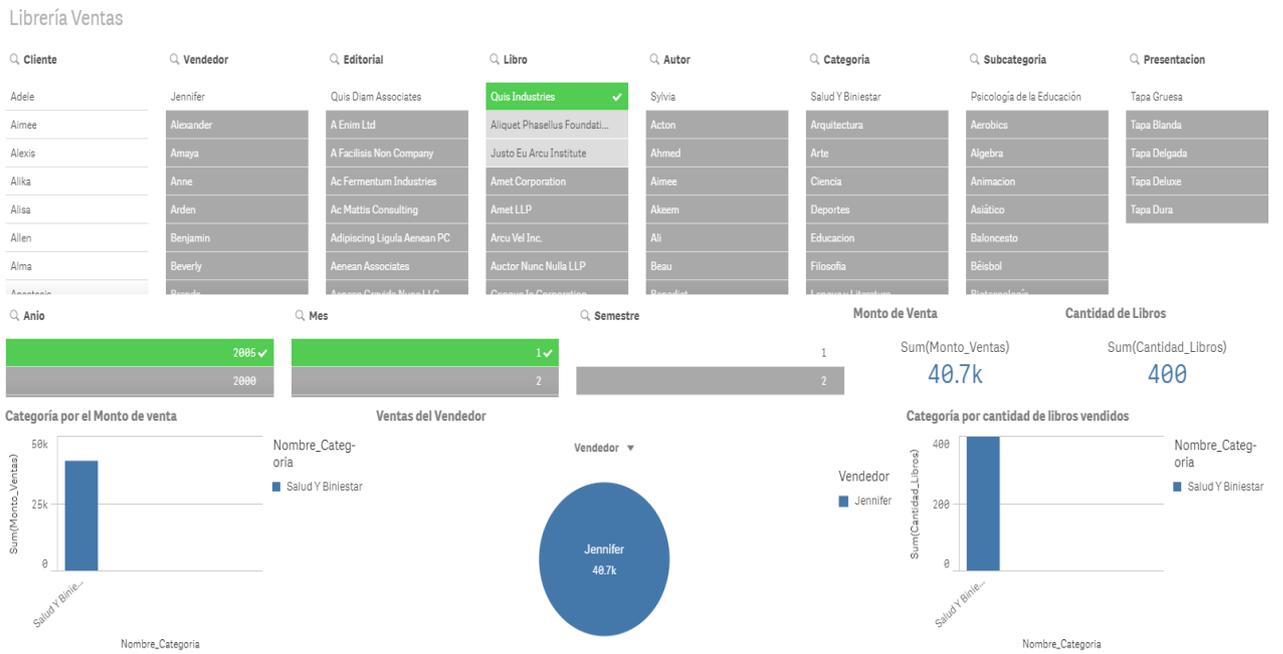
## R4. ¿Cuál es el monto de ventas por editorial y por mes?



## R5. ¿Cuál es el monto que generan las ventas de un libro por mes y año?



## R6. ¿Cuál es la cantidad de libros vendidos por categoría, sub categoría por mes y semestre?

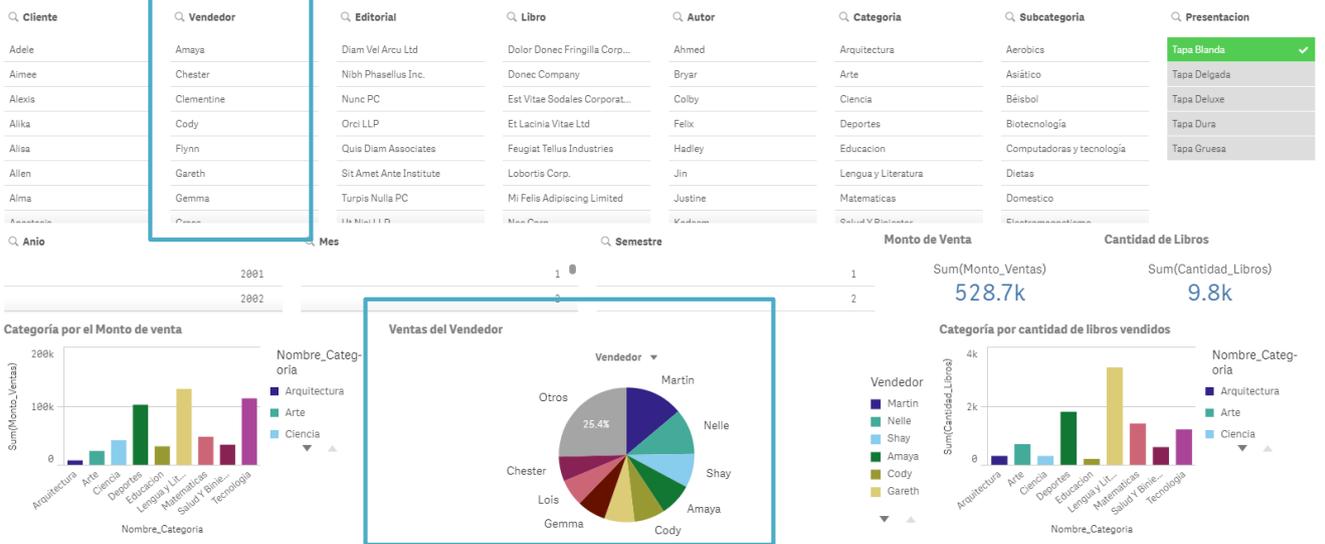


## R7. ¿Cuál es la cantidad de Productos vendidos por un vendedor al mes y al año?



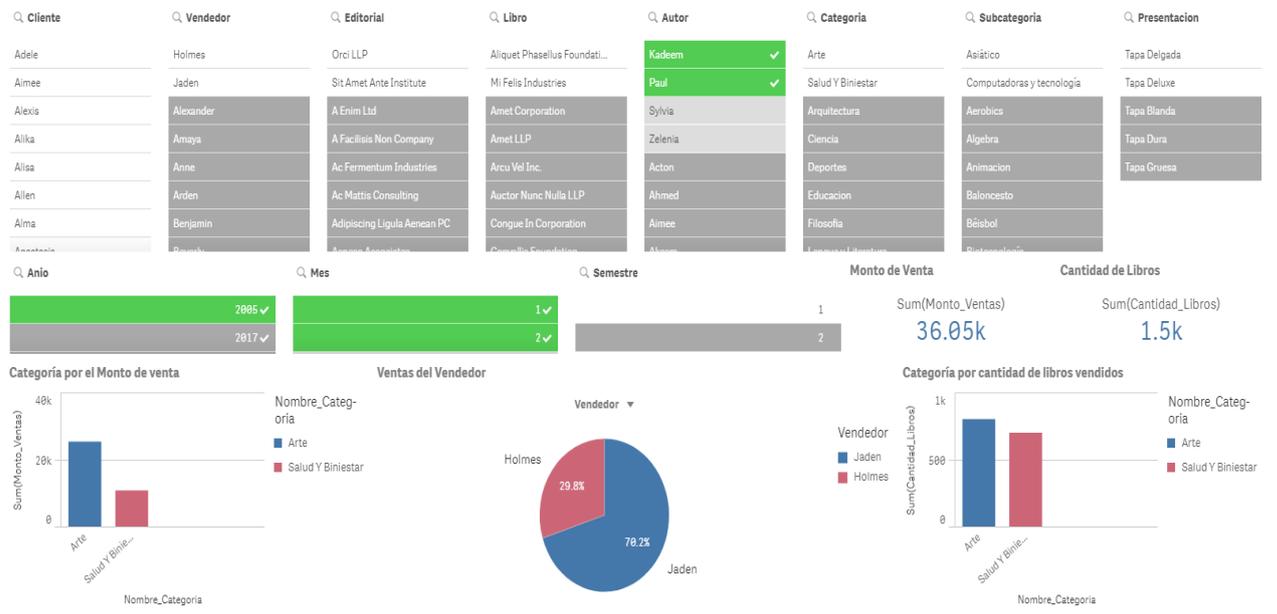


## Librería Ventas



## R10. ¿Cuál cantidad de ventas de libros por autor, por mes y año?

## Librería Ventas



## **5. DISCUSION DE RESULTADOS**

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

### **5.1. Formulación del Problema**

¿Cómo se podría mejorar la integración de datos y el soporte en la toma de decisiones para el proceso de ventas de la empresa Crisol S.A.C. sucursal Trujillo, utilizando tecnologías de información?

### **5.2. Hipótesis**

“El Desarrollo de una Solución de Inteligencia de negocios usando la arquitectura tecnológica de Microsoft y la metodología de Kimball mejora la Integración de Datos y el análisis de información para un mejor soporte en la toma de decisiones en el proceso de Ventas de la empresa Librería Crisol S.A.C. sucursal Trujillo.”

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

- ✓ Independiente (VI): Solución de Inteligencia de Negocios (BI).
- ✓ Dependiente (VD): mejora la Integración de Datos y el análisis de información para un mejor soporte en la toma de decisiones en el proceso de Ventas de la empresa Librería Crisol S.A.C. sucursal Trujillo.

### **5.3. Población y muestra.**

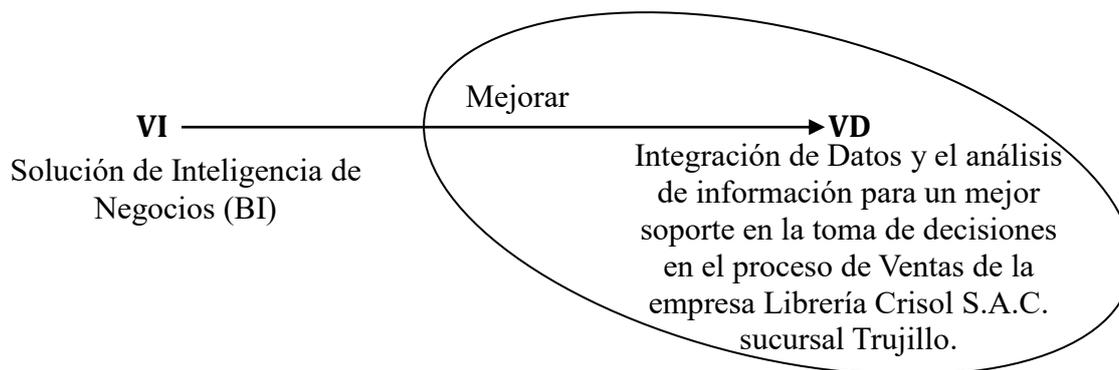
#### **5.3.1. Población**

Todos los reportes que sirven en la toma de decisiones en el proceso de Ventas de la empresa Librería Crisol S.A.C. sucursal Trujillo.

#### **5.3.2. Muestra**

Se analizarán 10 reportes principales para el proceso de Ventas de la empresa Librería Crisol S.A.C.

#### 5.4. MANERA PRESENCIAL



#### 5.5. 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 (BI). Este diseño se presenta de la siguiente manera:

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

Donde:

X: Tratamiento, estímulo (BI)

O: Medición a sujetos (Cuestionario)

G: Grupo de sujetos (Empleados)

El espacio de la muestra que se tomó para la medición de los indicadores de la hipótesis, correspondió al total de personas que operarán el BI, siendo estos 2; a estas personas se le aplicó un cuestionario, antes de interactuar con el BI (**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.

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

Para el cálculo de los indicadores de la hipótesis en el Desarrollo de la solución de Inteligencia de Negocios Propuesto (BIP) propuesto y el Sistema Actual (SA), se realizó un cuestionario donde se evaluó a los usuarios luego de haber interactuado con el BI.

Los valores que los usuarios dieron a las respuestas del cuestionario fueron aplicados según el siguiente Rango de valoración:

<b>RANGO</b>	<b>GRADO DE VALORACIÓN</b>
1	Malo
2	Regular
3	Bueno
4	Muy Bueno
5	Excelente

Tabla 10: Rango de grado de valoración

<b>ESCALA DE VALORACION TOTAL</b>	
<b>Inadecuado</b>	<b>15-45</b>
<b>Adecuado</b>	<b>46- 75</b>

### 5.5.2. APLICACIÓN DEL RANGO DE VALORACIÓN A LOS INDICADORES DE LA HIPÓTESIS

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

**Evaluación de los indicadores de la hipótesis:**

N°	INDICADORES	VALORACION					$\bar{X}$
		1	2	3	4	5	
1	La carga de datos al BI es la suficiente o requerida				2		4.0
2	Los reportes obtenidos son los requeridos u óptimo				1	1	4.5
3	Con el BI es posible obtener reportes aplicando un número de cruces de requerimientos.					2	5.0
4	El grado de acceso al sistema es lo requerido					2	5
5	Se puede acceder al sistema desde varios tipos de dispositivos					2	5
6	La interface de reportes es intuitiva, amigable y sencilla				1	1	4.5
7	La implementación del BI es una decisión acertada.					2	5
8	La información presentada por el BI cumple con el criterio de exactitud					2	5
9	El personal se encuentra satisfecha con el cambio del modo anterior de proceso de toma de decisiones al actual con el BI				1	1	4.5
10	La información presentada apoya al proceso de toma de decisiones del área de ventas.					2	5
$\sum \bar{X}$							<b>47.5</b>

Dónde:  $X = (\text{Valor Valoración} * \text{Número de empleados respondieron en nivel valoración}) / 2$

Tabla N° 11: Evaluación de los indicadores de la hipótesis.

**Interpretación:** La evaluación nos determinó la Solución de inteligencia de negocios es la más **adecuada**, por ser sumatoria de los promedios 47.5 y superior a 45.

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

#### Paso 1: Planteamiento de hipótesis.

$$H_0: O_1 \geq O_2$$

$$H_1: O_2 \geq O_1$$

Dónde:

**H<sub>0</sub> es la hipótesis Nula:** “El Desarrollo de una Solución de Inteligencia de negocios usando la arquitectura tecnológica de Microsoft y la metodología de Kimball no mejora la Integración de Datos y el análisis de información para un mejor soporte en la toma de decisiones en el proceso de Ventas de la empresa Librería Crisol S.A.C. sucursal Trujillo”

**H<sub>1</sub> es la hipótesis Alternativa:** “El Desarrollo de una Solución de Inteligencia de negocios usando la arquitectura tecnológica de Microsoft y la metodología de Kimball mejora la Integración de Datos y el análisis de información para un mejor soporte en la toma de decisiones en el proceso de Ventas de la empresa Librería Crisol S.A.C. sucursal Trujillo”

#### Paso 2: Nivel de significancia.

Para todo valor de probabilidad igual o menor que 0.05, se acepta H<sub>1</sub> y se rechaza H<sub>0</sub>.  $\alpha = 0,05$ .

#### Paso 3: Prueba estadística.

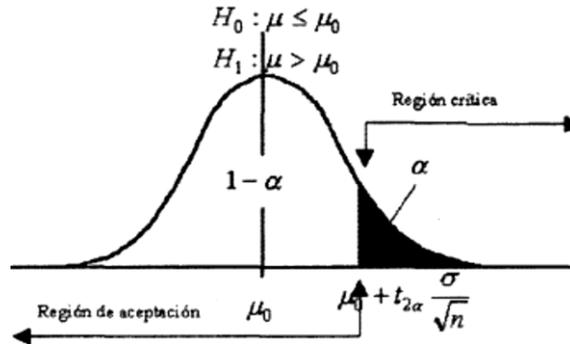
De acuerdo a que la muestra es de  $n = 10$ , y siendo menor a 30, se escogió la prueba estadística t-student, donde hay dos momentos uno antes y otro después de la solución. En el primer período, las observaciones servirán de control o testigo, para conocer los cambios que se susciten después de aplicar una variable experimental.

#### Paso 4: Zona de rechazo.

Si el valor de probabilidad es mayor que 0.05, se acepta  $H_0$  y se rechaza  $H_1$ .

Si la  $t_c > t_t$  se rechaza  $H_0$  y se acepta  $H_1$ .

Dónde:  $t_c$  es la t calculada y  $t_t$  es la t de tabla



#### Paso 5: Calculo de $t_t$ y $t_c$

##### Calculo de la t de tabla $t_t$

$t_t (95\%, 2) = 2,92 \quad \rightarrow \text{Ver Anexo B.}$

##### Calculo de la t calculado $t_c$

$$\bar{D} = \frac{\sum D}{n}, \delta = \sqrt{\frac{\sum (D_i - \bar{D})^2}{n - 1}}, t_c = \frac{\bar{D}}{\frac{\delta}{\sqrt{n}}}$$

Donde:

- $t_c$  : T calculado.
- $\delta$  : Desviación estándar
- $n$  : Tamaño de la muestra
- $\bar{D}$  : Valor promedio o media aritmética de las diferencias entre los momentos antes y después.

## Para el cálculo del valor de t calculado

Para el cálculo del valor T calculado se realizó un cuestionario (Ver Anexo A) donde se evaluó el grado de satisfacción a los usuarios luego de haber interactuado con el BI.

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
2.6 – 5.0	Medianamente Satisfecho
5.1 – 7.5	Satisfecho
7.6 – 10.0	Muy Satisfecho

Nº	INDICADORES	Media Pre U <sub>1</sub>	Media Post U <sub>2</sub>	D= (U <sub>2</sub> - U <sub>1</sub> )	(D <sub>i</sub> - $\bar{D}$ )	(D <sub>i</sub> - $\bar{D}$ ) <sup>2</sup>
1	¿Cuál es el monto de un determinado cliente en compras en un mes?	5.0	9.0	4	-1.4	1.96
2	¿Cuál es el Monto de ventas por categoría en un mes determinado?	3.0	8.0	5	-0.4	0.16
3	¿Cuál es el monto en ventas por sub categorías en un mes determinado?	2.0	9.0	7	1.6	2.56
4	¿Cuál es el monto de ventas por editorial y por mes?	4.0	9.0	5	-0.4	0.16
5	¿Cuál es el monto que generan las ventas de un libro por mes y año?	6.0	9.0	3	-2.4	5.76
6	¿Cuál es la cantidad de libros vendidos por categoría, sub categoría por mes y semestre?	3.0	9.0	6	0.6	0.36

7	¿Cuál es la cantidad de Productos vendidos por un vendedor al mes y al año?	4.0	9.0	5	-0.4	0.16
8	¿Cuál es el monto que generan las ventas de libros por tipo de presentación en un mes, semestre o año?	3.0	9.0	6	0.6	0.36
9	Listar a los Vendedores y el monto total de cada una de las ventas que han generado	3.0	9.0	6	0.6	0.36
10	¿Cuál cantidad de ventas de libros por autor, por mes y año?	2.0	9.0	7	1.6	2.56

$$N = 10 ; \sum D = 54 ; \bar{D} = 5.4 ; \sum (D_i - \bar{D})^2 = 14.4 ; \delta = 1.26 ; \sqrt{N} = 3.16$$

$$t_c = \frac{\bar{D}}{\frac{\delta}{\sqrt{n}}}$$

$$t_c = 36.12$$

**Interpretación:** Al tener como resultado que  $t_c > t_t$ , entonces se acepta la hipótesis alternativa, entendiéndose que una Solución de Inteligencia de negocios usando la arquitectura tecnológica de Microsoft y la metodología de Kimball mejora la Integración de Datos y el análisis de información para un mejor soporte en la toma de decisiones en el proceso de Ventas de la empresa Librería Crisol S.A.C. sucursal Trujillo.

## 5.6. CUADRO DE LA COMPARACIÓN DE TIEMPO DE DEMORA EN LA EJECUCIÓN DE LAS CONSULTAS.

NRO	CONSULTAS	CON EL SISTEMA OLTP	CON LA SOLUCION DE BI
1	¿Cuál es el monto de un determinado cliente en compras en un mes?	Se debe de procesar 1/2 hora = 1 800 seg	02 seg.
2	¿Cuál es el Monto de ventas por	Se debe de procesar 1 horas = 3 600 seg	02 seg.

	categoría en un mes determinado?		
3	¿Cuál es el monto en ventas por sub categorías en un mes determinado?	Se debe de procesar 01 horas = 3 600 seg	02 seg.
4	¿Cuál es el monto de ventas por editorial y por mes?	Se debe de procesar 1/2 hora = 1 800 seg	03 seg.
5	¿Cuál es el monto que generan las ventas de un libro por mes y año?	Se debe de procesar 01 horas = 3 600 seg	03 seg
6	¿Cuál es la cantidad de libros vendidos por categoría, sub categoría por mes y semestre?	Se debe de procesar 01 horas = 3 600 seg	03 seg
7	¿Cuál es la cantidad de Productos vendidos por un vendedor al mes y al año?	Se debe de procesar 1 hora = 3 600 seg	02 seg.
8	¿Cuál es el monto que generan las ventas de libros por tipo de presentación en un mes, semestre o año?	Se debe de procesar 1 horas = 3 600 seg	03 seg.
9	Listar a los Vendedores y el monto total de cada una de las ventas que han generado	Se debe de procesar 01 hora = 3 600 seg	02 seg.
10	¿Cuál cantidad de ventas de libros por autor, por mes y año?	Se debe de procesar 01 hora = 3 600 seg	02 seg.
		Fuente: La empresa	Fuente: Solución BI

## 6. CONCLUSIONES

- ✓ Se logró determinar los Stakeholders (02) involucrados en el proceso de ventas y obtener sus requerimientos a través de entrevistas siendo el resultado de 10 requerimientos que son necesarios para la toma de decisiones en el área de ventas de la empresa (ver pág. 52)
- ✓ Se logró analizar cada requerimiento para la elaboración del Modelo de Datos Dimensional que de soporte a la toma de decisiones en la organización obteniendo un modelo estrella de 09 dimensiones y 01 tabla de hechos.
- ✓ Se logró implementar el Modelo estrella en la Nube de Microsoft Azure creando un servidor en la nube que aloje a la base de datos estratégica de la solución. (ver pág. 74)
- ✓ Se desarrolló el proceso ETL del proyecto utilizando la herramienta Integration Service el cual permitió mejorar la extracción de los datos de las fuentes transaccionales para poblar la base de datos dimensional en la Nube de Azure, poblando las 09 dimensiones y la tabla de hechos. (ver pág. 81)
- ✓ Se diseñó la estructura del cubo OLAP de la aplicación BI utilizando Analysis Service de Microsoft, creando las dimensiones (09), sus jerarquías y las métricas de la tabla de hecho para obtener un nivel de conocimiento amplio y así mejorar el análisis de los datos, permitiendo una mejor navegabilidad de la información.
- ✓ Se logró crear los reportes de cada requerimiento (10) a través de la herramienta de Qlik Sense permitiendo una mayor flexibilidad en la elaboración de reportes con una mejor visualización de la información.

## 7. RECOMENDACIONES

- ✓ Se recomienda conocer a fondo todos los procesos del negocio de la empresa en donde se va a desarrollar un proyecto, con esto se evitará futuros problemas que traigan retrasos en tiempo y además gastos económicos no previstos.
- ✓ Se recomienda dirigir proyectos de TI hacia los lineamientos estratégicos de la institución y realizar una evaluación periódica de los procesos de Tecnologías de Información de la empresa.
- ✓ Se recomienda seguir la metodología de Ralph Kimball para el desarrollo de soluciones de Business Intelligence ya que esto ayudara a un mejor desarrollo y es la que más se ajusta a los requisitos y necesidades de un proyecto en BI.
- ✓ Mantener una relación estrecha con los usuarios involucrados en el área de donde se realizará el proyecto, ya que ellos en cada una de las fases que comprende la metodología aplicada nos brindaran información y ayudan a identificar los puntos necesarios para el desarrollo del proyecto.
- ✓ Se recomienda crear backups de la información en las bases de datos almacenada en la nube, debido a que pueda ocurrir una falla temporal del proveedor del servicio en la Nube.
- ✓ Se recomienda realizar capacitaciones a los tomadores de decisiones asignados para manejar la aplicación de Inteligencia de Negocios.

## 8. REFERENCIAS BIBLIOGRAFICAS

- Alcami, R. I., Carañana, C. D., & Herrando, J. G. (2009). *Introduccion a la gestion de sistemas de informacion en la empresa*. Universitat jaume.
- Arturo L, C. C. (2001). *GUÍA PARA OBTENER EL RETORNO A LA INVERSION EN PROYECTOS DE DATAWAREHOUSE*.
- Benites Robles, J., & López Girón, S. (2015). *Solución de Inteligencia de Negocio para Empresas de Servicio de Importación y exportación de Calzado del departamento La Libertad*. Trujillo: Repositorio UPAO.
- Cano, J. L. (2007). *BUSINESS INTELLIGENCE: competir con informacion*. madrid: fundacion banesto.
- Curto, J. (2007). *Introduccion al business intelligence*. españa: UOC.
- Flores, R. (2004). *Aplicación de minería de datos en ambiente universitario*.
- Gastañaudi Caballero, Y., & Tamayo Avalos, J. (2016). *Inteligencia de negocios y automatización en la gestión de puntos y fuerza de ventas en una empresa de tecnología*. Trujillo: Repositorio UPAO.
- Ian H. witten, E. F. (2000). *Data mining practical Machine learning tools and techniques*. san francisco: academic press.
- IBM. (29 de 04 de 2017). *IBM Knowledge Center*. Obtenido de [https://www.ibm.com/support/knowledgecenter/es/SS9UM9\\_9.1.0/com.ibm.datatools.dimensionai.ui.doc/topics/c\\_dm\\_star\\_schemas.html](https://www.ibm.com/support/knowledgecenter/es/SS9UM9_9.1.0/com.ibm.datatools.dimensionai.ui.doc/topics/c_dm_star_schemas.html)
- ITSitio. (11 de 08 de 2017). *SQL Server*. Obtenido de <http://www.itsitio.com/ar/sql-server-para-que-sirve-y-cual-es-la-version-que-necesito/>
- Kimball, R. (2013). *The Data Warehouse Toolkit*. Indiana: Wiley.
- Medina, D., & Calzado, L. (2008). *Data Mart para el área de ventas para la empresa Mi Mercado S.A.C. utilizando la metodología de Ralph Kimball y SQL Server como herramienta tecnológica*. Trujillo: Repositorio UPAO.

- MSDN Microsoft. (05 de 08 de 2017). *SQL server 2014*. Obtenido de <https://msdn.microsoft.com/es-es/library/bb545450.aspx>
- Oracle. (01 de 12 de 2016). *¿Qué es Inteligencia de Negocios?* . Obtenido de [http://www.oracle.com/ocom/groups/public/@otn/documents/webcontent/317529\\_esa.pdf](http://www.oracle.com/ocom/groups/public/@otn/documents/webcontent/317529_esa.pdf)
- Powerdata. (02 de 08 de 2017). *Integracion de datos*. Obtenido de <http://www.powerdata.es/integracion-de-datos>
- Rio, L. M. (2006). *Mas alla del business intelligence*. barcelona: gestion 2000.
- Rodriguez Torres, E., & Pereda Morales, P. (2013). *Implementación de un Dashboard para la toma de decisiones estratégicas en la unidad de negocio de producción de huevo incubable de la empresa Avícola Santa Fe S.A.C. Usando Tecnologías Oracle Business Intelligence*. Trujillo: Repositorio de Tesis UPAO.
- Roussel, G. (2006). *Decision support systems serving the company : the secrets to a succesfull project*.
- Sánchez Guevara, O. (2014). *Modelo de inteligencia de negocio para la toma de decisiones en la empresa San Roque S.A*. Trujillo: Repositorio UPAO.
- sanchez, L. z. (2008). *Metodología para el Diseño*. valencia.
- Silberschatz , A., Korth , H., & Sudarshan, S. (2006). *Fundamentos de Base de Datos*. Madrid: McGraw-Hill.
- Sinnexus. (20 de 04 de 2014). *sinnexus*. Obtenido de [http://www.sinnexus.com/business\\_intelligence/datawarehouse.aspx](http://www.sinnexus.com/business_intelligence/datawarehouse.aspx)
- Stackowiak, R., Rayman, J., & Rick, R. (2007). *Oracle Data Warehousing and Business Intelligence*. wiley.
- Wang, j. (2009). *Encyclopedia of Data Warehousing and Mining, Second Edition*.

# **ANEXOS**

## ANEXO A

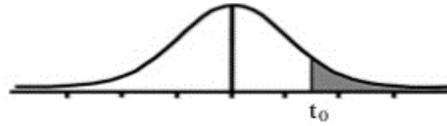
### CUESTIONARIO DIRIGIDO: Jefe de ventas y Administrador

PREGUNTAS	VALORES										
	0	1	2	3	4	5	6	7	8	9	10
¿Cuál es el monto de un determinado cliente en compras en un mes?											
¿Cuál es el Monto de ventas por categoría en un mes determinado?											
¿Cuál es el monto en ventas por sub categorías en un mes determinado?											
¿Cuál es el monto de ventas por editorial y por mes?											
¿Cuál es el monto que generan las ventas de un libro por mes y año?											
¿Cuál es la cantidad de libros vendidos por categoría, sub categoría por mes y semestre?											
¿Cuál es la cantidad de Productos vendidos por un vendedor al mes y al año?											
¿Cuál es el monto que generan las ventas de libros por tipo de presentación en un mes, semestre o año?											
Listar a los Vendedores y el monto total de cada una de las ventas que han generado											
¿Cuál cantidad de ventas de libros por autor, por mes y año?											

*Tabla A1. Jefe de ventas y Administrador*

## ANEXO B

Tabla t-Student



Grados de libertad	0.25	0.1	0.05	0.025	0.01	0.005
1	1.0000	3.0777	6.3137	12.7062	31.8210	63.6559
2	0.8165	1.8856	2.9200	4.3027	6.9645	9.9250
3	0.7649	1.6377	2.3534	3.1824	4.5407	5.8408
4	0.7407	1.5332	2.1318	2.7765	3.7469	4.6041
5	0.7267	1.4759	2.0150	2.5706	3.3649	4.0321
6	0.7176	1.4398	1.9432	2.4469	3.1427	3.7074
7	0.7111	1.4149	1.8946	2.3646	2.9979	3.4995
8	0.7064	1.3968	1.8595	2.3060	2.8965	3.3554
9	0.7027	1.3830	1.8331	2.2622	2.8214	3.2498
10	0.6998	1.3722	1.8125	2.2281	2.7638	3.1693
11	0.6974	1.3634	1.7959	2.2010	2.7181	3.1058
12	0.6955	1.3562	1.7823	2.1788	2.6810	3.0545
13	0.6938	1.3502	1.7709	2.1604	2.6503	3.0123
14	0.6924	1.3450	1.7613	2.1448	2.6245	2.9768
15	0.6912	1.3406	1.7531	2.1315	2.6025	2.9467
16	0.6901	1.3368	1.7459	2.1199	2.5835	2.9208
17	0.6892	1.3334	1.7396	2.1098	2.5669	2.8982
18	0.6884	1.3304	1.7341	2.1009	2.5524	2.8784
19	0.6876	1.3277	1.7291	2.0930	2.5395	2.8609
20	0.6870	1.3253	1.7247	2.0860	2.5280	2.8453
21	0.6864	1.3232	1.7207	2.0796	2.5176	2.8314
22	0.6858	1.3212	1.7171	2.0739	2.5083	2.8188
23	0.6853	1.3195	1.7139	2.0687	2.4999	2.8073
24	0.6848	1.3178	1.7109	2.0639	2.4922	2.7970
25	0.6844	1.3163	1.7081	2.0595	2.4851	2.7874
26	0.6840	1.3150	1.7056	2.0555	2.4786	2.7787
27	0.6837	1.3137	1.7033	2.0518	2.4727	2.7707
28	0.6834	1.3125	1.7011	2.0484	2.4671	2.7633
29	0.6830	1.3114	1.6991	2.0452	2.4620	2.7564
30	0.6828	1.3104	1.6973	2.0423	2.4573	2.7500
31	0.6825	1.3095	1.6955	2.0395	2.4528	2.7440
32	0.6822	1.3086	1.6939	2.0369	2.4487	2.7385
33	0.6820	1.3077	1.6924	2.0345	2.4448	2.7333
34	0.6818	1.3070	1.6909	2.0322	2.4411	2.7284
35	0.6816	1.3062	1.6896	2.0301	2.4377	2.7238
36	0.6814	1.3055	1.6883	2.0281	2.4345	2.7195
37	0.6812	1.3049	1.6871	2.0262	2.4314	2.7154
38	0.6810	1.3042	1.6860	2.0244	2.4286	2.7116
39	0.6808	1.3036	1.6849	2.0227	2.4258	2.7079
40	0.6807	1.3031	1.6839	2.0211	2.4233	2.7045
41	0.6805	1.3025	1.6829	2.0195	2.4208	2.7012
42	0.6804	1.3020	1.6820	2.0181	2.4185	2.6981
43	0.6802	1.3016	1.6811	2.0167	2.4163	2.6951
44	0.6801	1.3011	1.6802	2.0154	2.4141	2.6923
45	0.6800	1.3007	1.6794	2.0141	2.4121	2.6896
46	0.6799	1.3002	1.6787	2.0129	2.4102	2.6870
47	0.6797	1.2998	1.6779	2.0117	2.4083	2.6846
48	0.6796	1.2994	1.6772	2.0106	2.4066	2.6822
49	0.6795	1.2991	1.6766	2.0096	2.4049	2.6800