



UNIVERSIDAD TÉCNICA DEL NORTE
FACULTAD DE INGENIERÍA EN CIENCIAS APLICADAS
CARRERA DE INGENIERÍA EN SISTEMAS COMPUTACIONALES

TRABAJO DE GRADO PREVIO A LA OBTENCIÓN DEL TÍTULO
DE INGENIERO EN SISTEMAS COMPUTACIONALES

TEMA:

“Diseño de un DATAMART orientado al proceso de Ventas usando la herramienta de Inteligencia de Negocios SQL Server 2014”

AUTOR:

Ana Mercedes Montenegro Rivera

DIRECTOR:

Ing. Mauricio Rea

Ibarra – Ecuador

2015

CERTIFICACIÓN DIRECTOR

Por medio de la presente Yo Ing. Mauricio Rea, certifico: que la Sra. Ana Mercedes Montenegro Rivera, portadora de la cedula de identidad número: 0401404184, ha trabajado en el desarrollo del proyecto de tesis **“Diseño de un DATAMART orientado al proceso de Ventas usando la herramienta de Inteligencia de Negocios SQL Server 2014”** previo a la obtención del Título de Ingeniera en Sistemas Computacionales, la cual la ha realizado en su totalidad con responsabilidad.

Es todo cuanto puede certificar en honor a la verdad.



Ing. Mauricio Rea
DIRECTOR DE TESIS

CERTIFICADO DE CESION DE DERECHOS DE AUTOR

Yo, Ana Mercedes Montenegro Rivera, con cedula de identidad numero 0401401484, manifiesto mi voluntad de ceder a la Universidad Tecnica del Norte los derechos patrimoniales consagrados en la ley de Propiedad Intelectual del Ecuador, articulos 4, 5, 6, en calidad de autor del trabajo de grado denominado **“Diseño de un DATAMART orientado al proceso de Ventas usando la herramienta de Inteligencia de Negocios SQL Server 2014”**, que ha sido desarrollado para optar por el título de Ingeniero en Sistemas Computacionales, en la Universidad Técnica del Norte, quedando la Universidad facultada para ejercer plenamente los derechos cedidos anteriormente.

En mi condición de autor me reservo los derechos morales de la obra antes citada.

En concordancia suscribo este documento en el momento que hago la entrega del trabajo final en formato impreso y digital a la Biblioteca de la Universidad Técnica del Norte.



Nombre: Ana Mercedes Montenegro Rivera

Cedula: 0401401484

Ibarra 15 de julio de 2015

AUTORIZACIÓN DE USO Y PUBLICACIÓN

A FAVOR DE LA UNIVERSIDAD TÉCNICA DEL NORTE

1. IDENTIFICACIÓN DE LA OBRA

La UNIVERSIDAD TÉCNICA DEL NORTE dentro del proyecto Repositorio Digital institucional determina la necesidad de disponer los textos completos de forma digital con la finalidad de apoyar los procesos de investigación, docencia y extensión de la universidad. Por medio del presente documento dejo sentada mi voluntad de participar en este proyecto, para lo cual ponemos a disposición la siguiente investigación:

DATOS DE CONTACTO	
CEDULA DE IDENTIDAD	0401401484
APELLIDOS Y NOMBRES	MONTENEGRO RIVERA ANA MERCEDES
DIRECCIÓN	El Olivo Calle Plutarco Larrea 3-45
E-MAIL	montenegroany@gmail.com
TELÉFONO FIJO	062601063
TELÉFONO MÓVIL	0958970278
DATOS DE TRABAJO DE GRADO	
TITULO	“Diseño de un DATAMART orientado al proceso de Ventas usando la herramienta de Inteligencia de Negocios SQL Server 2014”
AUTOR	MONTENEGRO RIVERA ANA MERCEDES
FECHA	15 de julio de 2015
PROGRAMA	PREGRADO
TITULO	INGENIERÍA EN SISTEMAS COMPUTACIONALES
DIRECTOR	ING. MAURICIO REA

2. AUTORIZACIÓN DE USO A FAVOR DE LA UNIVERSIDAD

Yo, ANA MERCEDES MONTENEGRO RIVERA, con cédula de identidad Nro. 0401401484, en calidad de autor y titular de los derechos patrimoniales del trabajo de grado descrito anteriormente, hago entrega del ejemplar respectivo en forma digital y autorizo a la Universidad Técnica del Norte, la publicación de la obra en el Repositorio Digital Institucional y el uso del archivo digital en la biblioteca de la universidad con fines académicos, para ampliar la disponibilidad del material y como apoyo a la educación, investigación y extensión, en concordancia con la Ley de Educación Superior Artículo 143.



Nombre: Ana Mercedes Montenegro Rivera

Cedula: 0401401484

Ibarra 15 de julio de 2015

3. CONSTANCIAS

La autora manifiesta que la obra objeto de la presente autorización es original y se la desarrolló, sin violar derechos de autor de terceros, por lo tanto la obra es original y que es el titular de los derechos patrimoniales, por lo que asume la responsabilidad sobre el contenido de la misma y saldrá en defensa de la Universidad en caso de reclamación por parte de terceros.

Ibarra, a los 15 días del mes de Julio de 2015

EL AUTOR:



Nombre: Ana Mercedes Montenegro Rivera

DEDICATORIA

El presente proyecto está dedicado de manera especial a mi hija Andy, que ha sido desde siempre fuente de amor e inspiración.

A mi esposo por todo el cariño y apoyo.

A mi familia que siempre ha estado apoyándome, especialmente a mis queridos padres que con sus consejos apoyo y comprensión han sabido guiarme durante todas las etapas de mi vida.

Ana Montenegro

AGRADECIMIENTO

A MIS PADRES, por ser las personas que con su amor incondicional me han apoyado y guiado en todos los aspectos de mi vida.

A MI ESPOSO E HIJA, que con su paciencia y amor han sabido apoyarme siempre, les amo gracias por ser parte de mi vida.

A MI QUERIDA FAMILIA, hermanos y sobrinos, por su apoyo incondicional en todo aspecto, soy muy afortunada de contar con todos ustedes.

A MIS QUERIDOS SUEGROS Y CUÑADOS, por su valioso aporte en esta etapa de mi vida, especialmente a Ligiesita por su cariño y comprensión que ha sido como mi segunda madre.

A LA UTN, por brindarme sus conocimientos, al Ing. Mauricio Rea por su gran apoyo que aparte de ser un buen amigo ha sabido guiarme adecuadamente como asesor de Tesis, y de igual manera a los amigos que me han sido de gran apoyo en esta etapa, de manera especial a Jorge y Martin.

A todos les quedo muy agradecida

Ana Montenegro

ÍNDICE DE CONTENIDOS

CERTIFICACIÓN DIRECTOR	II
DEDICATORIA	VI
AGRADECIMIENTO	VIII
ÍNDICE DE CONTENIDOS	IX
ÍNDICE DE ILUSTRACIONES	XII
ÍNDICE DE TABLAS	XVI
RESUMEN	XVII
SUMARY	XVIII
CAPITULO 1	1
1. INTRODUCCIÓN	2
1.1 ANTECEDENTES	2
1.2 DEFINICIÓN DEL PROBLEMA	2
1.3 OBJETIVOS	3
1.3.1 OBJETIVO GENERAL	3
1.3.2 OBJETIVOS ESPECÍFICOS	3
1.4 ALCANCE	4
1.5 JUSTIFICACIÓN	5
CAPITULO 2	7
2. INTELIGENCIA DE NEGOCIOS	8
2.1 INTRODUCCIÓN	8
2.2 DEFINICIÓN Y CARACTERÍSTICAS	8
2.3 PRINCIPALES CARACTERÍSTICAS DE INTELIGENCIA DE NEGOCIOS	11
2.4 ARQUITECTURA	13
2.5 HERRAMIENTAS DE INTELIGENCIA DE NEGOCIOS	14
2.5.1 CARACTERÍSTICAS, VENTAJAS Y DESVENTAJAS	14
2.6 ELEMENTOS DE INTELIGENCIA DE NEGOCIOS	17
2.6.1 FUENTES DE INFORMACIÓN	17
2.6.2 PROCESOS ETL	17

2.6.3	REPOSITORIO DE INFORMACIÓN DATAWAREHOUSE	19
2.6.4	DATAMART.....	21
2.6.4.1	DATAMART DEPENDIENTES	21
2.6.4.2	DATAMART INDEPENDIENTES	22
2.6.5	METODOLOGÍAS DE DATAWAREHOUSE	23
2.6.5.1	METODOLOGÍA KIMBALL	23
2.6.5.2	METODOLOGÍA BILL INNMON	25
2.6.6	HERRAMIENTAS OLAP.....	27
2.6.7	BASE DE DATOS MULTIDIMENSIONALES.....	29
2.6.8	CUBOS.....	31
2.6.9	MODELAMIENTO DE UN DATAWAREHOUSE.....	32
2.6.9.1	ESQUEMA ESTRELLA	32
2.6.9.2	ESQUEMA COPO DE NIEVE	34
2.6.10	HERRAMIENTAS DE VISUALIZACIÓN	35
2.7	SQL SERVER BUSINESS INTELLIGENCE 2014	35
2.7.1	DEFINICIÓN Y CARACTERÍSTICAS.....	35
2.7.2	SQL SERVER DATA TOOLS.....	38
2.7.3	ANALYSIS SERVICES.....	38
2.7.3.1	POWER PIVOT.....	40
2.7.3.2	POWER VIEW.....	41
2.7.4	INTEGRATION SERVICES	42
2.7.5	REPORTING SERVICES.....	43
	CAPITULO 3	47
3.	DISEÑO DEL DATAMART DE VENTAS	48
3.1	PLANIFICACIÓN.....	48
3.1.1	ACTIVIDADES Y ALCANCE	48
3.1.2	DEFINICIÓN DE RECURSOS.....	49
3.1.3	IDENTIFICACIÓN DE REQUERIMIENTOS.....	50
3.2	ANÁLISIS Y DISEÑO	54
3.2.1	ORÍGENES DE DATOS.....	54

3.2.2	MAPEO DE LA INFORMACIÓN	58
3.2.3	DEFINICIÓN DE DIMENSIONES.....	64
3.2.4	IDENTIFICACIÓN DE TABLA DE HECHOS Y SUS MEDIDAS	65
3.2.5	MODELO DE DATOS	66
3.3	CONSTRUCCIÓN Y DESARROLLO	67
3.3.1	CONEXIÓN A BASE DE DATOS FUENTE	67
3.3.2	INSTALACIÓN DE SQL SERVER 2014	68
3.3.3	INSTALACIÓN DE VISUAL STUDIO 2013.....	77
3.3.4	CREACIÓN DE BASE DE DATOS MULTIDIMENSIONAL.....	84
3.3.5	CONSTRUCCIÓN DE PROCESOS ETL.....	87
3.3.6	CONSTRUCCIÓN DEL CUBO CON MODELO MULTIDIMENSIONAL	104
3.3.7	CONSTRUCCIÓN DEL MODELO TABULAR	113
3.4	PRUEBAS	122
	CAPITULO 4.....	131
	CONCLUSIONES.....	132
	RECOMENDACIONES.....	133
	ANÁLISIS DE IMPACTOS	134
	BIBLIOGRAFÍA.....	136
	GLOSARIO.....	138

ÍNDICE DE ILUSTRACIONES

Ilustración 1: Preguntas para la Toma de Decisiones	6
Ilustración 2: Proceso de Inteligencia de Negocios	10
Ilustración 3: Niveles de toma de decisiones	11
Ilustración 4: Arquitectura de Inteligencia de Negocios.....	14
Ilustración 5: Esquema Datamart Dependientes.....	21
Ilustración 6: Esquema Datamart Independientes	22
Ilustración 7: Diferencias entre DWH y Datamart	22
Ilustración 8: Esquema Metodología Kimball	24
Ilustración 9: Fases para la construcción DWH según Kimball.	24
Ilustración 10: Esquema Metodología Bill Inmon.....	25
Ilustración 11: Representación Tabla de Hechos.....	30
Ilustración 12: Representación de Dimensiones	30
Ilustración 13: Representación gráfica de un CUBO.....	31
Ilustración 14: Modelamiento de un DATAWAREHOUSE	32
Ilustración 15: Esquema Estrella.....	33
Ilustración 16: Esquema Copo de Nieve	34
Ilustración 17: Arquitectura de Inteligencia en Negocios en Microsoft.....	37
Ilustración 18: Modelo Semántico de Analysis Services.....	39
Ilustración 19: Fases de diseño del DataMart	48
Ilustración 20: Diagrama de Proceso de ventas	51
Ilustración 21: Variables de Tiempo	52
Ilustración 22: Variables Vendedor	52
Ilustración 23: Variables Comerciales.....	53
Ilustración 24: Variables de Producto.....	53
Ilustración 25: Variables Numéricas	54
Ilustración 26: Mapa de orígenes de datos	55
Ilustración 27: Origen Dimensión cliente	56
Ilustración 28: Origen Dimensión Artículos	57
Ilustración 29: Origen Dimensión Vendedor	57
Ilustración 30: Origen Dimensión Fecha y Tabla de Hechos.....	58
Ilustración 31: Modelo de Datos	66
Ilustración 32: Instalación de ORACLE DATA ACCES COMPONENT	67
Ilustración 33: Archivo tnsnames de conexión a ORACLE.....	68
Ilustración 34: Paso 1 Instalación SQL Server 2014	68
Ilustración 35: Paso 2 Instalación SQL Server 2014	69
Ilustración 36: Paso 3 Instalación SQL Server 2014	69
Ilustración 37: Paso 4 Instalación SQL Server 2014	70
Ilustración 38: Paso 5 Instalación SQL Server 2014	70
Ilustración 39: Paso 6 Instalación SQL Server 2014	71
Ilustración 40: Paso 7 Instalación SQL Server 2014	71
Ilustración 41: Paso 8 Instalación SQL Server 2014	72
Ilustración 42: Paso 9 Instalación SQL Server 2014	72

Ilustración 43: Paso 10 Instalación SQL Server 2014	73
Ilustración 44: Paso 11 Instalación SQL Server 2014	73
Ilustración 45: Paso 12 Instalación SQL Server 2014	74
Ilustración 46: Paso 11 Instalación SQL Server 2014	74
Ilustración 47: Paso 14 Instalación SQL Server 2014	75
Ilustración 48: Paso 15 Instalación SQL Server 2014	75
Ilustración 49: Paso 16 Instalación SQL Server 2014	76
Ilustración 50: Paso 17 Instalación SQL Server 2014	76
Ilustración 51: Paso 18 Instalación SQL Server 2014	77
Ilustración 52: Paso 1 Instalación Visual Studio 2013	78
Ilustración 53: Paso 2 Instalación Visual Studio 2013	78
Ilustración 54: Paso 3 Instalación Visual Studio 2013	79
Ilustración 55: Paso 4 Instalación Visual Studio 2013	79
Ilustración 56: Paso 5 Instalación Visual Studio 2013	80
Ilustración 57: Paso 5 Instalación Visual Studio 2013	81
Ilustración 58: Paso 7 Instalación Visual Studio 2013	81
Ilustración 59: Paso 8 Instalación Visual Studio 2013	82
Ilustración 60: Paso 9 Instalación Visual Studio 2013	82
Ilustración 61: Paso 10 Instalación Visual Studio 2013	83
Ilustración 62: Paso 11 Instalación Visual Studio 2013	83
Ilustración 63: Paso 12 Instalación Visual Studio 2013	84
Ilustración 64: Código Creación Base de datos multidimensional	84
Ilustración 65: Código Creación Tabla DIM_ARTICULOS	85
Ilustración 66: Código Creación Tabla DIM_CLIENTES	85
Ilustración 67: Código Creación Tabla DIM_FECHA.....	86
Ilustración 68: Código Creación Tabla DIM_VENDEDOR	86
Ilustración 69: Código Creación Tabla FACT_VENTAS	86
Ilustración 70: Diagrama de Base de Datos Multidimensional.....	87
Ilustración 71: Selección del tipo de administrador de conexión	88
Ilustración 72: Conexión a base de datos fuente en Oracle	89
Ilustración 73: Conexión a base de Datos SQL destino.....	90
Ilustración 74: Selección del tipo de conexión para Excel.....	90
Ilustración 75: Seleccionar el archivo de Excel para conexión	92
Ilustración 76: Conexiones creadas	92
Ilustración 77: Selección de conexión y tabla de origen	93
Ilustración 78. Selección de campos de origen Cliente	93
Ilustración 79: Selección de campos de origen Cliente	94
Ilustración 80. Selección del campo de navegación	94
Ilustración 81: Selección de campos de la combinación de tablas.....	95
Ilustración 82: Selección del archivo de origen Excel	96
Ilustración 83: Selección de campos del origen Excel	96
Ilustración 84: Transformación de datos para Excel.....	97
Ilustración 85: Selección de la tabla destino	97
Ilustración 86: Mapeo de datos.....	98

Ilustración 87: Procesamiento de tareas	99
Ilustración 88: Información cargada en tabla destino	100
Ilustración 89: Código de Vista para Fecha.....	101
Ilustración 90: Mapeo de datos para Fecha	101
Ilustración 91: Información cargada en tabla DIM_FECHA	102
Ilustración 92: Código para extraer y transformar datos para la tabla de Hechos	103
Ilustración 93: Selección origen de datos para tabla de hechos	103
Ilustración 94: Información cargada en Tabla de hechos.....	104
Ilustración 95: Selección del Origen de datos	105
Ilustración 96: Información de Suplantación.....	105
Ilustración 97: Selección de Vistas de origen de datos.....	106
Ilustración 98: Vistas de Origen de datos	106
Ilustración 99: Especificación de tabla origen	107
Ilustración 100: Especificación de atributos	107
Ilustración 101: Estructura de la dimensión	108
Ilustración 102: Selección método de creación Cubo	108
Ilustración 103: Selección de tabla de hechos.....	109
Ilustración 104: Selección de medidas	109
Ilustración 105: Selección de Dimensiones	110
Ilustración 106: Visualización del Cubo	110
Ilustración 107: Correspondencia de dimensiones con grupo de medidas	111
Ilustración 108: Implementación del Cubo.....	111
Ilustración 109: Exportación datos a Excel	112
Ilustración 110: Visualización de datos en Excel	112
Ilustración 111: Menú de POWER PIVOT.....	113
Ilustración 112: Selección de origen de datos.....	114
Ilustración 113: Conexión al origen de datos	114
Ilustración 114: Administrador para la importación de tablas.....	115
Ilustración 115: Selección de tablas de origen	115
Ilustración 116: Vista previa de tablas seleccionadas	116
Ilustración 117: Importación de tablas.....	116
Ilustración 118: Visualización de la información	117
Ilustración 119: Visualización de cálculos.....	118
Ilustración 120: Vista de diagramas.....	118
Ilustración 121: Jerarquía Cliente	119
Ilustración 122: Jerarquía Tiempo	119
Ilustración 123: Jerarquía Clasificación Artículos	120
Ilustración 124: Opciones de Tabla dinámica	120
Ilustración 125: Ambiente de Excel para tablas dinámicas con Power Pivot	122
Ilustración 126: Reporte Compras Clientes por época	123
Ilustración 127: Reporte de Ventas y Utilidad por Artículo.....	124
Ilustración 128: Ventas y Utilidad por año	125
Ilustración 129: Reporte con el mismo filtro de dos puntos de vista	125
Ilustración 130: Ventas por Clase de Cliente al año	126

Ilustración 131: Ambiente de Excel para Power View	126
Ilustración 132: Menú de opciones de Power View	127
Ilustración 133: Informe con secciones en Power View	128
Ilustración 134: Informe interactivo de Ventas	129
Ilustración 135: Informe interactivo con línea de tiempo	129

ÍNDICE DE TABLAS

Tabla 1: Definiciones de Inteligencia de Negocios	10
Tabla 2: Diferencias entre las bases de datos transaccionales y almacén de datos.	13
Tabla 3: Principales herramientas de Inteligencia de Negocios	15
Tabla 4: Comparación entre Metodologías de DWH.....	27
Tabla 5: Tipos de herramientas OLAP.....	28
Tabla 6: Ediciones de SQL Server 2014	36
Tabla 7: Tipos de soluciones de Analysis Services	39
Tabla 8: Tecnologías de SQL Server	45
Tabla 9: Tabla de actividades.....	49
Tabla 10: Definición de recursos	50
Tabla 11: Matriz de proceso de Ventas	52
Tabla 12: Mapeo Dimensión Cliente.....	59
Tabla 13: Mapeo Dimensión Artículos.....	60
Tabla 14: Mapeo Dimensión Vendedor	61
Tabla 15: Mapeo Dimensión Fecha	62
Tabla 16: Mapeo Tabla de Hechos.....	63
Tabla 17: Dimensión Clientes	64
Tabla 18: Dimensión Artículo.....	65
Tabla 19: Dimensión Vendedor	65
Tabla 20: Dimensión Fecha.....	65
Tabla 21: Tabla de hechos	66
Tabla 22: Descripción estados Cliente	91
Tabla 23. Descripción indicadores Cliente.....	91

RESUMEN

El presente proyecto de tesis tiene como finalidad explicar el proceso a seguir para diseñar un DATAMART del área de ventas de una entidad comercial, para lo cual se pone en práctica los conceptos de Inteligencia de Negocio y se hace uso de las herramientas que permiten llevar a cabo este proceso.

En el capítulo 1 se detalla una pequeña introducción a los aspectos generales de Inteligencia de Negocios, así como también se hace énfasis en la importancia y los beneficios que pueden obtener las empresas con la adaptación de estas herramientas.

En el capítulo 2 se describe el marco teórico en donde se explica la arquitectura y las herramientas que permiten realizar los procesos de Inteligencia de negocios.

En el capítulo 3 se realiza la documentación de todo el proceso a seguir para la construcción del DATAMART, para lo cual se aplica metodología de construcción ágil.

En el cuarto capítulo finalmente se detallan las conclusiones y recomendaciones planteadas de acuerdo a la experiencia obtenida.

SUMMARY

This thesis project aims to explain the process to follow in designing a DataMart the sales area of a commercial entity, for which we practice the concepts of Business Intelligence and make use of the tools to carry out this process.

In Chapter 1 outlined a brief introduction to the general aspects of Business Intelligence and is also emphasis on the importance and benefits that companies can obtain the adaptation of these tools.

In chapter 2 the theoretical framework in which the architecture and the tools that allow us to make the Business Intelligence processes described explained.

Chapter 3 documents the entire process is performed to continue to build the mart, for which construction agile methodology is applied.

In the fourth chapter finally the conclusions and recommendations made in accordance with the experience gained you can itemize.

CAPITULO 1

INTRODUCCIÓN



- ✓ **INTRODUCCIÓN**
- ✓ **ANTECEDENTES**
- ✓ **DEFINICIÓN DEL PROBLEMA**
- ✓ **OBJETIVOS**
- ✓ **ALCANCE**

1. INTRODUCCIÓN

1.1 ANTECEDENTES

Las empresas están compuestas por varios procesos, que generan gran cantidad de información diariamente, misma que se almacena en sistemas transaccionales que se encargan de administrarla. Dicha información es lo más valioso de una empresa, ya que de ella se puede obtener conocimientos que permiten plantear estrategias competitivas.

Esta información generada, por lo general se encuentra desorganizada por lo cual su análisis se vuelve complejo, en este contexto, las soluciones de Inteligencia de Negocios buscan organizar esta información a fin de dar ventajas a las empresas para lograr mayor visibilidad sobre el mercado y sus actores, ayudando a los ejecutivos a entender los datos en menor tiempo y de esta manera puedan tomar mejores decisiones e incrementar así su eficiencia organizacional, a través de la obtención de reportes óptimos, sobre información estructurada.

1.2 DEFINICIÓN DEL PROBLEMA

Actualmente en muchas empresas se realiza una manipulación compleja de la información, para creación de reportes gerenciales, de cualquier área; reportes que son necesarios para proyectarse con estrategias de ventas competitivas; es decir para la toma de decisiones.

El problema se presenta a la hora de la toma de decisiones estratégicas, ya que toda la información que generan las empresas como resultado de sus procesos, se encuentra dispersa y desorganizada por lo que su análisis se vuelve lento, y de bajo rendimiento, provocando así que las decisiones se tomen sin toda la información relevante o se tomen malas decisiones.

La información en los sistemas transaccionales es estática, por lo tanto no se puede acceder a ella de manera eficiente, para obtener los reportes de apoyo a

la toma de decisiones. Es decir limita a tener una visión a corto plazo mas no ha tener una planificación o proyección a largo plazo.

La información con la que cuentan las empresas para ser competitivas debe ser rápida y oportuna, para que sus procesos no sean lentos ya que esto hace que la gestión y administración se vuelva obsoleta.

En si las Bases de Datos están diseñadas para procesar transacciones, mas no para dar soporte a decisiones. Mientras que en un DataWarehouse (almacén de datos) la información está preparada para el análisis ya que cuenta con un modelamiento multidimensional.

Por dicha razón, se plantea diseñar una solución Inteligencia de Negocios para tener toda la información de ventas, estructurada y consolidada en un solo repositorio, permitiendo explotarla, para la generación de reportes y que permitan plantear estrategias de ventas eficaces, con menor tiempo y esfuerzo.

1.3 OBJETIVOS

1.3.1 OBJETIVO GENERAL

Diseñar un DATAMART¹ orientado al proceso de Ventas utilizando la herramienta de Inteligencia de Negocios SQL Server 2014².

1.3.2 OBJETIVOS ESPECÍFICOS

- Obtener la información de los sistemas transaccionales que manejan los procesos en área de ventas.
- Analizar la información recolectada para definir los requerimientos e indicadores más importantes que ayuden al diseño del DATAMART.

¹ DATAMART: almacén de información de un tema específico.

² SQL SERVER 2014: sistema de gestión de datos.

- Investigar las herramientas de Inteligencia de Negocios a usar para el diseño del DATA MART, como lo es el SQL Server 2014 y Excel como herramienta de presentación de reportes.
- Estudiar los componentes de inteligencia de negocios así como varias de sus metodologías de desarrollo.

1.4 ALCANCE

En el presente proyecto se estudiará los componentes de Inteligencia de negocios, los cuales permiten diseñar un DATAMART de información, entre los componentes están: los procesos de Extracción, Transformación y Carga, herramientas OLAP³, herramientas de análisis y visualización, metodologías de desarrollo de un repositorio de datos, entre otras.

Usando las herramientas de Inteligencia de Negocios contenidas en Microsoft SQL Server 2014 se diseñará un DATAMART para el área de ventas de una empresa comercial, el mismo que conforma modelos dimensionales y tabulares de información, y que son la fuente de los reportes en los que se muestra los conocimientos para la toma de decisiones estratégicas del área mencionada.

Se utilizará las herramientas de SQL Server 2014 para el diseño multidimensional y los procesos ETL (extracción, transformación y carga), para la explotación del cubo y la presentación de los reportes se usa herramientas Microsoft para Excel.

A través del diseño del presente DATAMART se podrá identificar como se está manejando la historia de ventas y proyectarse que acciones aplicar para mejorar, así como también conocer información, que no había estado disponible para el análisis.

Para lograrlo intervienen varias dimensiones y métricas que al interrelacionarse brindan información al detalle sobre las ventas.

³ OLAP: procesamiento analítico en línea.

Las dimensiones que se analizan son las siguientes:

- **Clientes:** Registro de información completa de clientes
- **Vendedor:** Registro de la información de cada vendedor
- **Tiempo:** Información de fechas al detalle por mes, día y año
- **Artículo:** Registro completo de cada artículo

1.5 JUSTIFICACIÓN

Debido al gran impacto que han generado las soluciones de Inteligencia de Negocios en las empresas, por su capacidad de convertir los datos en conocimiento, y así generar un nivel alto de competitividad, se ha propuesto estudiar en este caso las herramientas Microsoft para Inteligencia de Negocios.

La mayoría de las grandes empresas utilizan los sistemas transaccionales como fuente de su información, a la cual se puede acceder de manera compleja para determinar ciertas consultas, ya que hay que recorrer varias tablas de la Base de Datos para sacar la información requerida, mientras que con la información contenida de manera organizada en un solo repositorio de datos se puede acceder fácilmente a la información ya consolidada; esto se logra gracias a la aplicación de soluciones de Inteligencia de negocios. Generando así grandes beneficios para las empresas en cuanto a recursos y tiempo, ya que les permite proyectarse a futuro con tácticas de operación, con las cuales podrán generar una mayor rentabilidad.

A través de las soluciones de Inteligencia de Negocios las empresas logran tomar una decisión óptima y realista basada y fundamentada en información. Esta información sirve de apoyo para poder realizar, proyecciones y hasta prevenir posibles problemas.

Este tipo de herramientas son fundamentales para resolver los problemas que comúnmente se presentan en las empresas, como por ejemplo:

- Obtención de reportes gerenciales, sin depender del departamento de sistemas.
- Optimización de tiempo en acceso a la información requerida en los reportes.
- Calidad de datos e información comprensible para el usuario final, entre otros.

Todos estos factores solucionados, influyen directamente en términos económicos en beneficio de las entidades que usan estas herramientas.

En cuanto al software disponible actualmente para desarrollar soluciones de BI, existen varias herramientas muy óptimas que permiten obtener excelentes resultados, en este contexto para el desarrollo del presente proyecto se utiliza SQL Server Management Studio como motor de base de datos, SQL Server Data Tools⁴ para el modelado dimensional, Power Pivot⁵ en Excel para el modelo tabular y el análisis de datos así como varias herramientas Microsoft en Excel para la presentación de informes, la combinación de estas herramientas ofrecen una plataforma para crear e implementar soluciones de Inteligencia de Negocios escalables, seguras y dinámicas en cuanto a su administración, además de ser herramientas bastante amigables con el usuario final. En conclusión con la aplicación de las soluciones de Inteligencia de Negocios se busca brindar beneficios a las empresas proporcionando información y conocimiento necesario que les permita responder a preguntas planteadas sobre su futuro las cuales les permitirá mejorar su desempeño para crecer en el mundo competitivo.

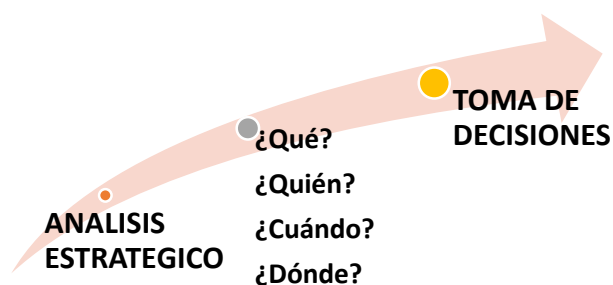


Ilustración 1: Preguntas para la Toma de Decisiones
Fuente: Propia

⁴ SQL Server DataTools: conjunto de herramientas para trabajar datos.

⁵ Power Pivot: Herramienta para procesar y analizar datos en ambiente de Excel

CAPITULO 2

MARCO TEÓRICO



- ✓ INTELIGENCIA DE NEGOCIOS
- ✓ INTRODUCCIÓN
- ✓ DEFINICIÓN Y CARACTERÍSTICAS
- ✓ PRINCIPALES CARACTERÍSTICAS
- ✓ ARQUITECTURA
- ✓ HERRAMIENTAS DE INTELIGENCIA DE NEGOCIOS
- ✓ ELEMENTOS DE INTELIGENCIA DE NEGOCIOS

2. INTELIGENCIA DE NEGOCIOS

2.1 INTRODUCCIÓN

Toda empresa en la actualidad automatiza sus procesos de gestión a través del uso de sistemas informáticos que les permite la gestión de sus datos y su explotación diaria, pero hasta ahí solo se consigue tener un respaldo de los datos, el siguiente paso es consolidar, refinar y analizar estos datos extrayendo información que apoye a la toma de decisiones para beneficio de las empresas.

La inteligencia de negocios tiene como finalidad proporcionar la información para ser analizada y tomar decisiones, a través de la aplicación de ciertos procesos que permiten obtener dicha información, en este capítulo se analiza la base teórica que permite llevar a cabo una solución de Inteligencia de Negocios para una empresa comercial a través del diseño de un DATAMART para su departamento de ventas.

2.2 DEFINICIÓN Y CARACTERÍSTICAS

Kevin Roebuck (2012)

“La inteligencia de negocios (BI) se refiere principalmente a las técnicas informáticas utilizadas en la identificación, la extracción y el análisis de los datos de negocio, tales como los ingresos por ventas de productos y / o departamentos, o por los costos y los ingresos asociados. Las tecnologías de Business Intelligence proporcionan histórico, actual y puntos de vista de predicción de las operaciones comerciales. Las funciones de Inteligencia de Negocios son: procesamiento analítico en línea, análisis, minería de datos, minería de procesos, gestión de rendimiento empresarial, la evaluación comparativa, la minería de textos y análisis predictivo. La inteligencia de negocios tiene como objetivo apoyar una mejor toma de decisiones (DSS).”

Libros Científicos (2015)

“Podemos definir la inteligencia de negocio o BI (Business Intelligence) como el conjunto de estrategias enfocadas a la administración y creación de conocimiento sobre el medio, a través del análisis de los datos existentes en una organización o empresa. El objetivo de la inteligencia de negocios es ofrecer conocimientos para respaldar las decisiones empresariales. “

Las soluciones de Inteligencia de Negocios proporcionan información que permiten competitividad, en la actualidad se puede ver que muchas empresas compiten en similares condiciones, por ejemplo: tienen productos similares o tienen los mismos proveedores, siguen los mismos procesos, pero su objetivo debe ser el de lograr una ventaja competitiva usando la información para optimizar procesos o mejorar la oferta de productos y servicios.

En este contexto se puede decir que la Inteligencia de Negocios es un proceso orientado a la información el cual parte de los **datos** digitalizados en una BDD o fuentes de datos externa, y permite pasarlos a **información**, misma que tiene un significado y proporciona **conocimiento**, en base al cual se puede actuar para tomar las mejores decisiones que al final brindarán resultados que proporcionan rentabilidad a la empresa.

Definiendo mejor estos conceptos:

DEFINICIÓN	EJEMPLOS
Datos: Conjunto de hechos sin procesar acerca de un evento.	<ul style="list-style-type: none">• Notas de Andy Becerra: Matemáticas 9, Lengua 10, Ingles 8.• Costo de un producto, Gasto mensual.

<p>Información: Datos con significado, que han sido procesados. Ejemplos:</p>	<ul style="list-style-type: none"> • El promedio de Andy Becerra es 9. • Ventas por producto, ganancias de Enero.
<p>Conocimiento: Patrones de comportamiento. Ejemplos:</p>	<ul style="list-style-type: none"> • El promedio de Andy Becerra con respecto al resto de la clase es alto. • Las ganancias de enero son malas, causas de las malas ganancias.

Tabla 1: Definiciones de Inteligencia de Negocios
Fuente: Propia

En general las soluciones de Inteligencia de Negocios en cualquier ambito al que se la aplique permite:

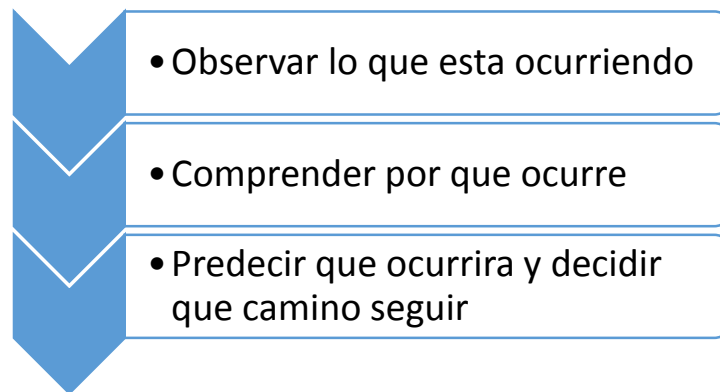


Ilustración 2: Proceso de Inteligencia de Negocios
Fuente: Propia

La intligencia de negocios para cumplir con sus objetivos permite acceder y analizar la información para la toma de desiciones a nivel operativo, táctico y estrategico:



Ilustración 3: Niveles de toma de decisiones

Fuente: Propia

NIVEL OPERATIVO: en este nivel la Inteligencia de Negocios permite que el personal que trabajan con información operativa pueda recibirla de manera oportuna y a tiempo, la información operativa es la que se genera diariamente de los procesos de gestión empresarial, que pueden ser reportes predefinidos.

NIVEL TÁCTICO: en este nivel se permite que los analistas de información y mandos medios de la empresa puedan interactuar con la información de manera directa, y acceder a consultas para su análisis.

NIVEL ESTRATÉGICO: en este nivel se encuentra a los ejecutivos de alta gerencia los cuales pueden analizar, monitorear y evaluar la información generada como tendencias, patrones, anticipar problemas, establecer correctivos etc. En este nivel se aplican los cuadros de mando o tableros de control.

2.3 PRINCIPALES CARACTERÍSTICAS DE INTELIGENCIA DE NEGOCIOS

Las características y beneficios de la Inteligencia de negocios son muchas por todas las ventajas que brinda su aplicación, pero entre las principales se destacan las siguientes:

- Posee un solo repositorio de datos, con información consolidada y estructurada.
- Manejo de datos no volátiles.
- Proporciona datos preparados para la el análisis.
- Apoya a la toma de decisiones usando herramientas de análisis.
- Permite conocer mejor la evolución de los negocios.
- Ahorro de tiempo y dinero.
- Brinda la posibilidad de revisar los datos al detalle.
- Ayuda a optimizar los procesos de negocio y permite actuar para crear ventajas competitivas.
- Proporciona técnicas de análisis de información para el apoyo a la toma de decisiones. Entre otras.

En base a todos los beneficios que brinda la aplicación de soluciones de Inteligencia de Negocios, se puede realizar una comparación con los sistemas transaccionales o aplicaciones estándar:

BASE DE DATOS TRANSACCIONAL VS. ALMACÉN DE DATOS

BASE DE DATOS TRANSACCIONAL	ALMACÉN DE DATOS
Necesidad del negocio	Oportunidad de negocio
Consultas estáticas, Datos Volátiles	Consultas dinámicas
Nuevos informes necesitan desarrollo	El usuario diseña con facilidad y flexibilidad sus diseños
Orientado a niveles bajos y altos de administración	Principalmente orientado a altos niveles jerárquicos

Información operacional	Datos integrados y estandarizados
Datos actuales	Datos actuales e históricos
Manejo de pequeños conjuntos de datos	Maneja un conjunto importante de datos
Datos para lectura y escritura	Datos solo para lectura

Tabla 2: Diferencias entre las bases de datos transaccionales y almacén de datos.

Fuente: Propia

2.4 ARQUITECTURA

La Inteligencia de negocios posee una arquitectura global muy flexible, capaz de proporcionar al diseñador varias opciones para crear una solución de BI, en general se basa en el siguiente proceso:

1. Obtiene información de las fuentes de datos, que pueden ser: sistemas operacionales, fuentes externas, ERP, entre otras.
2. A estas fuentes de datos aplica un proceso ETL (Extracción, Transformación y Carga), para optimizar su proceso analítico.
3. A la información resultante, ya unificada, depurada y consolidada, se la almacena en los repositorios de información, que pueden ser Data Ware House o Data Mart.
4. Se aplica las herramientas OLAP, que permiten construir el cubo de datos.
5. Finalmente la información está lista, sobre la cual se puede aplicar herramientas de gestión y análisis que permitan su explotación y así obtener los reportes necesarios para la toma de decisiones.

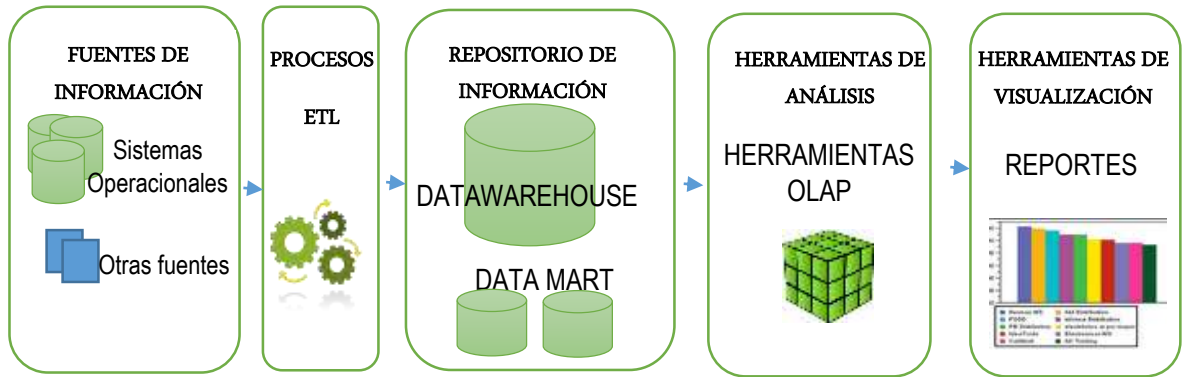


Ilustración 4: Arquitectura de Inteligencia de Negocios
Fuente: Propia

2.5 HERRAMIENTAS DE INTELIGENCIA DE NEGOCIOS

2.5.1 CARACTERÍSTICAS, VENTAJAS Y DESVENTAJAS

Para aplicar la Inteligencia de Negocios se cuenta con una amplia gama de herramientas con arquitectura escalable, que permiten obtener información depurada y acceder a la información para explotarla. La gran mayoría de estas herramientas permite satisfacer las necesidades de los usuarios finales en cuanto a análisis y reportes de los datos que se encuentran en los repositorios, brindando varias opciones de visualización, búsqueda y acceso como pueden ser cuadros de mando entre otros.

Herramienta	Descripción
Generadores de informes	Utilizadas por desarrolladores profesionales para crear informes estándar para grupos, departamentos o la organización
Herramientas de usuario final de consultas e informes	Empleadas por usuarios finales para crear informes para ellos mismos o para otros; no requieren programación.

Herramientas OLAP	Permiten a los usuarios finales tratar la información de forma multidimensional para explorarla desde distintas perspectivas y periodos de tiempo.
Herramientas de Dashboard	Permiten a los usuarios finales ver información crítica para el rendimiento con un simple vistazo utilizando iconos gráficos y con la posibilidad de ver más detalle para analizar información detallada e informes, si lo desean.
Herramientas de planificación, modelización y consolidación	Permite a los analistas y a los usuarios finales crear planes de negocio y simulaciones con la información de Business Intelligence. Pueden ser para elaborar la planificación, los presupuestos, las previsiones. Estas herramientas proveen a los dashboards y los scorecards con los objetivos y los umbrales de las métricas.
Herramientas Datamining	Permiten a estadísticos o analistas de negocio crear modelos estadísticos de las actividades de los negocios. Datamining es el proceso para descubrir e interpretar patrones desconocidos en la información mediante los cuales resolver problemas de negocio. Los usos más habituales del datamining son: segmentación, venta cruzada, sendas de consumo, clasificación, previsiones, optimizaciones, etc.”

Tabla 3: Principales herramientas de Inteligencia de Negocios

Fuente: Josep Lluís Cano (2007)

Para el presente caso de estudio se usaron varias de estas herramientas, las mismas que servirán de apoyo para el diseño del DATAMART, basada en la arquitectura propuesta anteriormente.

Como se ha descrito anteriormente los beneficios de las Soluciones de Inteligencia de Negocios son muy amplios y útiles para el desarrollo empresarial.

Sin embargo al llevar a cabo los procesos para su aplicación se pueden presentar varios riesgos o desventajas como:

- La mala predisposición de ciertos departamentos dentro de la empresa ante los cambios, ya que aplicar soluciones de Business Intelligence implican llevar a cabo procesos de reingeniería que conllevan un esfuerzo extra, y varios procesos tediosos a los cuales los usuarios ponen resistencia ya que no comprenden todavía las ventajas de una solución de BI.
- En cuanto a las fuentes de datos, si la empresa no cuenta con un registro organizado de la información, por ejemplo si la base transaccional no tiene adecuada integridad referencial, será más difícil cumplir con la totalidad de requerimientos, solicitados por el usuario final.
- Otro problema es que al no haber información necesaria para el análisis, se deba implementar un sistema de gestión que registre correctamente los datos que se necesitan para el análisis, implicando así que la información histórica a analizar no sea suficiente para sacar mejores resultados.
- La calidad de los datos es otro factor que hay que tomar en cuenta a la hora de implementar estas soluciones ya que estos deben estar validados, para tener mejores resultados, caso contrario no se podrá obtener los beneficios esperados.
- Para su implementación se requiere una inversión económica fuerte, ya que su implementación no es una tarea fácil y requiere de muchos

recursos tanto humanos como materiales. Pero se lo puede considerar como una inversión a mediano o largo plazo.

2.6 ELEMENTOS DE INTELIGENCIA DE NEGOCIOS

Dentro de los principales componentes de Inteligencia de Negocios, partiendo de su arquitectura se analizar los siguientes:

2.6.1 FUENTES DE INFORMACIÓN

Las fuentes de información son con las que cuenta la entidad, sobre las cuales se trabaja para sacar la información que alimentara al repositorio de datos, se puede usar multiples fuentes, por lo general las principales fuentes de información son:

- Fuentes internas o generadas por los sistemas operacionales o transaccionales como ERP, CRM,⁶ SCM,⁷ Bases de datos, archivos planos, archivos de excel, entre otros.
- Fuentes externas que pueden ser fuentes que permitan complementar la información disponible, como por ejemplo poblacion, numero de habitantes, etc.
- Fuentes de informacion generadas por los departamentos de las empresas que no se registren en el sistema transaccional.

En esta fase el objetivo principal implica la realizacion de un estudio minucioso de la informacion que se tiene disponible acerca de su formato, disponibilidad y calidad, para asi poder capturarla, porcesarla y dejarla lista para usarla.

2.6.2 PROCESOS ETL

Este conjunto de procesos permiten obtener los datos de las fuentes de información para transformarlos y adaptarlos al modelo definido y luego

⁶ CRM: Customer Relationship Management, modelo de gestión enfocado principalmente en los clientes

⁷ SCM: Supply Chain Management, modelo de gestión o administración de la cadena de suministro

depositarlos en las bases de datos de destino o repositorios. En si se encarga de la gestión de datos, agregando calidad a los mismos.

Esta fase es una de las más importantes de los proyectos de Inteligencia de Negocios, e incluso uno de los más largos ya que consume por lo menos el 50% de tiempo del ciclo de vida del proyecto, además que debe estar bien diseñado para tener mejores resultados a la hora de la toma de decisiones.

Cabe destacar también que en ciertos casos en esta fase se suele necesitar herramientas de apoyo, debido a su complejidad e importancia, por esto muchas veces se utiliza el apoyo de bases de datos intermedias. Es decir que la información antes de entrar a los repositorios se almacena en un área de Staging o un ODS(Operational Data Store).

Staging: es un área temporal que se encuentra en el flujo de datos entre las fuentes y los repositorios, su objetivo es el de facilitar la extracción de datos, de llevar a cabo tareas de depuración y limpieza de datos, o ser también una memoria temporal para los datos operacionales que no estén registrados en los repositorios.

ODS: es un almacén operacional de datos, que se alimenta de los sistemas transaccionales, y al mismo tiempo les da soporte, para que estos tengan un menor carga de trabajo, sirve también como base de datos de consulta para las herramientas de reporting. Cuando existen ODS es mejor que los repositorios se alimenten de ellos para quitar las lecturas al sistema transaccional.

Entre las herramientas más populares de aplicaciones para procesos ETL (Extracción, transformación y carga) se destaca las siguientes:

- Pentaho Data Integration
- Oracle Warehouse Builder
- Cognos Decisionstream
- Microsoft SQL Server Integration Services (SSIS), entre otras.

Los principales procesos que se lleva a cabo son:

EXTRACCIÓN:

Este proceso se realiza para capturar la información desde las fuentes de datos o sistemas transaccionales. Para ello se utiliza herramientas especializadas de ETL (extracción, transformación y carga) contenidas en los motores de datos, que permitirán visualizar o detectar errores durante el proceso de carga.

El principal objetivo de la extracción es tomar tan solo los datos que sean necesarios y prepararlos para el resto de los procesos, realizando una depuración y validación de los datos obtenidos.

TRANSFORMACIÓN:

En esta fase el objetivo es usar los datos que se validaron en la extracción, para adaptarlos al formato de los datos de destino que se han definido en los repositorios. A parte de los cambios de formato en los datos la transformación incluye cambios en los valores de origen y la creación de nuevos valores en el destino, es decir valores derivados y agregados, por ejemplo valores totales de sumatorias, etc.

CARGA:

En este proceso se carga los nuevos datos a los repositorios DWH (Data Ware House) o DATA MART, de acuerdo a los formatos allí definidos y la periodicidad de actualización, también es importante en esta fase verificar que los datos coincidan con la información de la fuente.

2.6.3 REPOSITORIO DE INFORMACIÓN DATAWAREHOUSE

En esta parte de la arquitectura de Inteligencia de negocios se encuentra a los **DATAWAREHOUSE** como almacén de datos, en este almacén es en donde se depositara toda la información que se haya tratado con los procesos ETL (Extracción, transformación y carga), aquí la información estará lista para ser analizada para la toma de decisiones.

Kevin Roebuck (2012):

“Un almacén de datos (DW) es una base de datos utilizada para la presentación de informes. Los datos se cargan desde los sistemas operacionales para la presentación de informes. Los datos pueden pasar a través de un almacén de datos operativo para operaciones adicionales antes de que se utilicen en el DW para la presentación de informes.

Un almacén de datos mantiene sus funciones en tres capas: puesta en escena, la integración y el acceso. La extracción se utiliza para almacenar datos en bruto para su uso por los desarrolladores. La capa de integración se utiliza para integrar los datos y tener un nivel de abstracción para los usuarios. La capa de acceso es para la obtención de datos por los usuarios.”

DATAWAREHOUSE conocido por sus siglas (DWH) es un repositorio central que contiene la información más valiosa de la empresa. Los datos que aquí se almacenan han pasado por un proceso de calidad que asegura su consistencia. Además el repositorio está construido de tal manera que el acceso sea lo más rápido posible.

Entre los principales objetivos que debe cumplir un DWH (data ware house) se destacan los siguientes según:

Ralph Kimbal, Margy Ross (2013):

- El DWH (data ware house) debe permitir que la información de la organización sea accesible.
- Debe presentar la información de la organización consistentemente.
- Debe ser adaptable y resistente a cambios.
- Debe ser bastante seguro, que proteja la información contenida.
- La información contenida en el DWH (data ware house) debe ser de calidad para que pueda servir como base para la toma de decisiones.

2.6.4 DATAMART

De la definición de DWH (data ware house) se deriva también el concepto de DATAMART que consiste en un repositorio de datos orientado a un tema específico.

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. (Sinnexus, 2014).

Los Datamart pueden ser dependientes o independientes de acuerdo a como se vayan a diseñar:

2.6.4.1 DATAMART DEPENDIENTES

Estos son alimentados desde los datos del Datawarehouse.

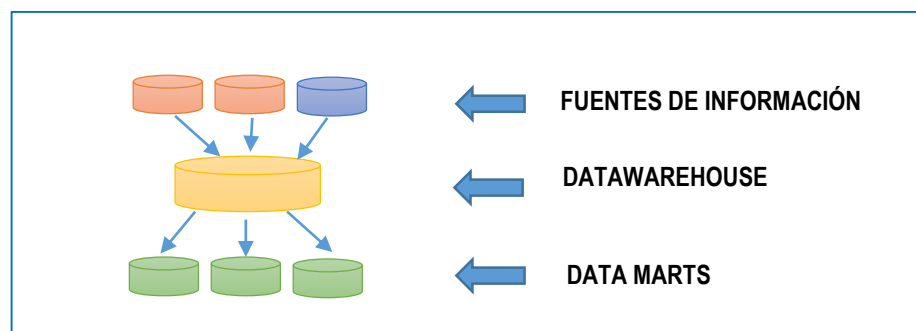


Ilustración 5: Esquema Datamart Dependientes

Fuente: Propia

2.6.4.2 DATAMART INDEPENDIENTES

Estos son alimentados directamente desde los orígenes de datos.

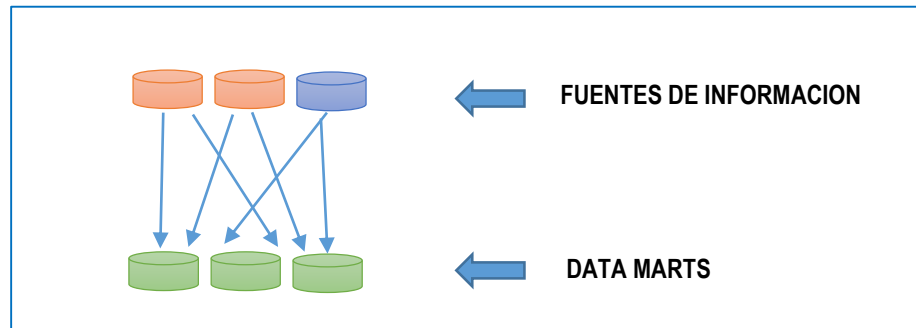


Ilustración 6: Esquema Datamart Independientes
Fuente: Propia

La diferencia entre estos dos conceptos de Datawarehouse y Datamart reside en el alcance que tiene cada uno, ya que mientras que el primero contiene información centralizada con datos globales de todas las áreas, el otro contiene datos específicos de un área de negocios.

DIFERENCIAS ENTRE DWH Y DATAMART

DATAWAREHOUSE	DATAMART
<ul style="list-style-type: none"> • Un solo proyecto de creación a largo plazo. • Etapas largas de análisis, diseño y construcción • Proyectos de 1 a 2 años 	<ul style="list-style-type: none"> • Varios proyectos cortos • Proyecto de 3 o 4 meses • Etapas de análisis, diseño construcción más rápida por su focalización.

Ilustración 7: Diferencias entre DWH y Datamart
Fuente: Propia

En este mismo contexto cabe describir las metodologías de Inteligencia de Negocios para la construcción de un Datawarehouse.

2.6.5 METODOLOGÍAS DE DATAWAREHOUSE

En el desarrollo de una solución de Inteligencia de Negocios, existen varias metodologías que permiten llevar a cabo los procesos de construcción, de una manera organizada, para optimizar los resultados. Para el presente caso de estudio se detalla la descripción de dos de las metodologías más usadas para el diseño.

2.6.5.1 METODOLOGÍA KIMBALL

Esta metodología es usada en Inteligencia de Negocios para la construcción de un almacén de datos DataWarehouse, fue planteada por Ralph Kimball, se basa principalmente en el ciclo de vida dimensional del negocio, en la cual plantea varias recomendaciones a tomar en cuenta para la creación de DWH (data warehouse):

- Analizar y comprender con claridad los requerimientos del negocio.
- Realizar un modelo flexible de las áreas de datos involucradas.
- Diseñar una arquitectura que responda a los procesos de negocio.
- Realizar entregas en incrementos significativos, basados en los procesos de negocios que conforman la Matriz Bus Empresarial.
- Ofrecer la solución completa.

Para la optimización de consultas y mejorar la facilidad de uso de datamarts, Kimball propone el modelo de datos como esquema estrella. Además sugiere una visión Bottom-Up⁸, es decir una vez extraído los datos de las fuentes se construye varios datamart departamentales e independientes, modelados en forma dimensional; que posteriormente pueden evolucionar para formar un Datawarehouse, que se relacionan entre sí a través de sus dimensiones.

⁸ Bottom-Up: estrategia de procesamiento de información “de abajo arriba”

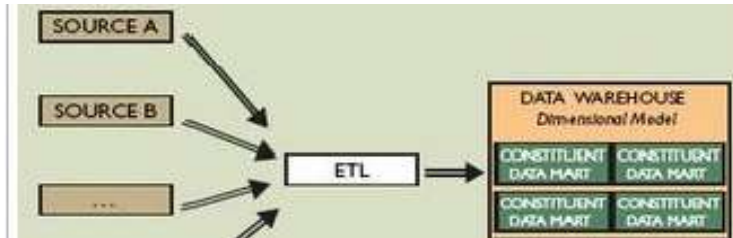


Ilustración 8: Esquema Metodología Kimball

Fuente: (mundobi. 2015) <http://mundobi.com.ar/wp-content/uploads/2012/04/kimball2.jpg>

En resumen prácticamente esta metodología propone 4 fases para la construcción del DWH, las cuales se detalla a continuación:

Fase 1. Selección del proceso de negocio

- En esta etapa lo primero es identificar que proceso o procesos de negocio se va a modelar para combinar y entender los requerimientos con los datos disponibles. Estos procesos son las actividades que desarrolla la empresa como rutina diaria y que están soportados por sistemas transaccionales.

Fase 2. Declarar la granularidad de la información

- Una vez identificado los procesos es necesario definir la granularidad, es decir: ¿Qué nivel de detalle debería estar disponible en el modelo dimensional?. Este paso es sumamente importante ya que de él dependerá el modelo de datos y la implementación del datamart.

Fase 3. Elección de las dimensiones de análisis

- Se deben elegir las dimensiones que se aplicarán para cada tabla de hechos, las dimensiones deben responder a la pregunta: ¿Cómo los usuarios describen los datos que resultan del proceso de negocios?.

Fase 4. Identificar los hechos o métricas

- En esta parte se debe definir los hechos numéricos o métricas que formarán parte de la tabla de hechos. Los hechos son determinados por esta pregunta: “¿Qué estamos midiendo?”. Los usuarios están muy interesados en el análisis de estas medidas de la ejecución de procesos de negocio.

Ilustración 9: Fases para la construcción DWH según Kimball.

Fuente: Propia

2.6.5.2 METODOLOGÍA BILL INNMON

Fue propuesta por Bill Inmon en el libro “Building the Data Warehouse”, en la cual propone varias características que un Data Warehouse debe cumplir, partiendo de su concepto este propone una visión Top-Down⁹, es decir construir un almacén de datos corporativo y a partir de él ir construyendo los modelos específicos para los distintos departamentos, parte de la estrategia a lo más operativo. Los datamart departamentales son tratados como subconjuntos del Data Warehouse corporativo.

El proceso que sigue, es primero extraer los datos de los sistemas operacionales a través de los procesos ETL (extracción, transformación y carga) para luego cargarlos en el área de stage, donde se los valida y se los consolida en el Data Warehouse corporativo. Una vez realizado este proceso, los procesos de refresco de los Data mart departamentales obtienen la información de este, y con las consiguientes transformaciones, organizan los datos en las estructuras particulares requeridas por cada uno de ellos, refrescando su contenido.

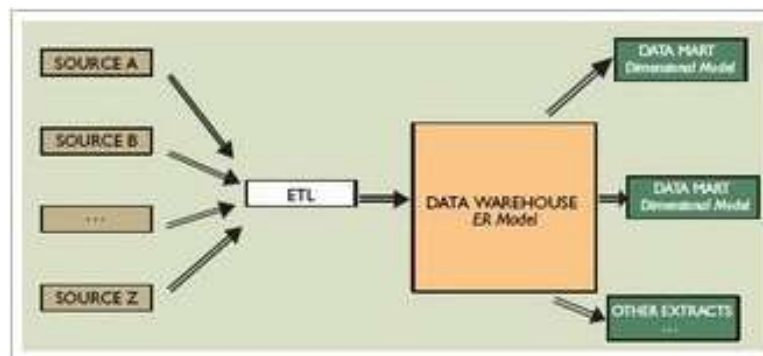


Ilustración 10: Esquema Metodología Bill Inmon

Fuente: (mundobi. 2015) <http://mundobi.com.ar/wp-content/uploads/2012/04/Kimball1.jpg>

⁹ Top-Down: estrategia de procesamiento de información “de arriba abajo”

Bill Inmon propone varias características que debe cumplir un *datawarehouse*, que son:

- **Orientado a temas.**- Los datos en la base de datos están organizados de manera que todos los elementos de datos relativos al mismo evento u objeto del mundo real queden unidos entre sí.
- **Integrado.**- La base de datos contiene los datos de todos los sistemas operacionales de la organización, y dichos datos deben ser consistentes.
- **No volátil.**- La información no se modifica ni se elimina, una vez almacenado un dato, éste se convierte en información de sólo lectura, y se mantiene para futuras consultas.
- **Variante en el tiempo.**- Los cambios producidos en los datos a lo largo del tiempo quedan registrados para que los informes que se puedan generar reflejen esas variaciones.

Las dos metodologías descritas coinciden en usar un repositorio de datos para poder integrarlos y explotarlos tomados de varias fuentes. Más sin embargo cada una tiene sus características y procedimientos. A continuación se detalla una comparación entre estas dos metodologías:

Metodología	Kimball	Inmon
Definición Data Mart	Un data mart mantiene los datos al menor nivel de detalle, los cuales se refieren a un proceso de negocio	Un data mart mantiene los datos agregados que se relacionan a la unidad de negocio
	Un data mart se construye mediante la extracción de datos directamente desde los sistemas operacionales	Un data mart se construye mediante la extracción de datos del data warehouse de la empresa (también llamados data marts dependientes).
	Los data marts están vinculados entre sí.	Los data marts no están vinculados entre sí.
Enfoque de desarrollo	Se basan en procesos específicos del negocio y están vinculadas a las dimensiones, que forman la arquitectura de bus data warehouse.	El diseño de un Data Warehouse para toda la empresa se basa en su modelo de datos. Es una aplicación progresiva de Las áreas temáticas, de acuerdo con las prioridades establecidas

Tabla 4: Comparación entre Metodologías de DWH

Fuente: Roberto Villareal, (2013). Estudio de metodologías de Data Warehouse para la implementación de repositorios de información para la toma de decisiones gerenciales. UTN.

2.6.6 HERRAMIENTAS OLAP

Las herramientas OLAP (OnLine Analytical Processing) permiten analizar la información que reside en los Datawarehouse o los Datamart. Su principal objetivo es agilizar las consultas de grandes volúmenes de información para lo cual utiliza estructuras multidimensionales conocidas como Cubos.

Las herramientas OLAP cuentan con varios operadores que permiten manipular la información:

DRILL: “más detalle” nivel de detalle.

ROLL: “agregación” nivel de agregación.

SLICE : “rotar” cambiar de orden las dimensiones.

DICE: “seleccionar” solo las celdas requeridas.

Existen distintos tipos de herramientas OLAP, con las cuales se puede construir una DWH (data ware house), la única diferencia entra ellas, básicamente, depende de cómo acceden a los datos, a continuación se detallan los 3 tipos:

<p>ROLAP: Relational OLAP</p>	<p>Las capacidades OLAP acceden directamente a la base de datos relacional. Se accede por tanto a una base de datos relacional. Accede habitualmente sobre un modelo “estrella”. La principal ventaja es que no tiene limitaciones en cuanto al tamaño, pero es más lento que el MOLAP, aunque algunos productos comerciales permiten cargar cubos virtuales para acelerar los tiempos de acceso.</p>
<p>MOLAP: Multidimensional OLAP</p>	<p>La implementación OLAP accede directamente sobre una base de datos multidimensional (MDDDB). La ventaja principal de esta alternativa es que es muy rápida en los tiempos de respuesta y la principal desventaja es que, si se desea cambiar las dimensiones, se debe cargar de nuevo el cubo.</p>
<p>HOLAP: Hybrid OLAP</p>	<p>Accede a los datos de alto nivel en una base de datos multidimensional y a los atómicos directamente sobre la base de datos relacional. En esencia utiliza las ventajas del ROLAP y del MOLAP. “</p>

Tabla 5: Tipos de herramientas OLAP
Fuente: Josep Lluís Cano Deposito 2007.

En este mismo contexto cabe definir también las Bases de Datos Multidimensionales:

2.6.7 BASE DE DATOS MULTIDIMENSIONALES

Las Bases de datos dimensionales tienen la capacidad de almacenar datos con varias dimensiones, es decir en vez de un valor se encuentran varios dependiendo de los ejes definidos, está orientada a consultas complejas y de alto rendimiento, con grandes volúmenes de datos. Esta base de datos está compuesta por varias estructuras dimensionales llamadas cubos. Para el almacenamiento de los procesos de negocio utiliza tablas de hechos y tablas de dimensiones.

- **Tabla de Hechos:** contiene hechos o datos numéricos de indicadores a analizar que proporcionan información histórica, es decir contiene información cuantitativa. Los valores o campos de esta tabla se los conoce como **medidas**.

Una **medida** es lo que está siendo sujeto de análisis a través de las múltiples dimensiones, un ejemplo puede ser el importe de ventas.

Las medidas deben vincularse con los niveles de granularidad requeridos por el diseño que se plantee, cada medida se toma mediante la intersección de las dimensiones que la definen.

Además una tabla de hechos debe poseer una clave que está compuesta por las claves primarias de las tablas de dimensiones relacionadas a esta; es decir contiene las claves subrogadas de las dimensiones que definen su nivel de detalle. Ya que como característica más importante de las tablas de hechos se tiene el **nivel de detalle** de la información que almacena, por ejemplo en el presente proyecto las ventas están organizadas a nivel de producto, clientes, vendedores y la fecha.

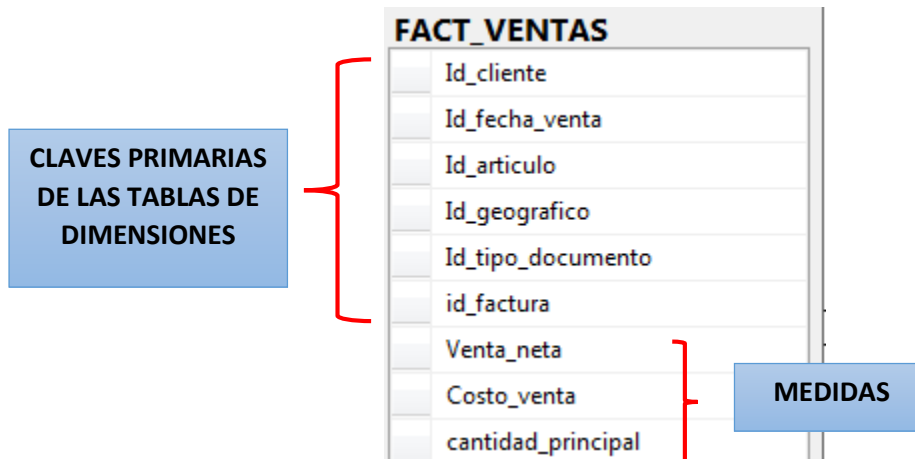


Ilustración 11: Representación Tabla de Hechos
Fuente: Propia

- Dimensiones:** son tablas relacionadas con la tabla de hechos, contiene información acerca de un evento o perspectiva de análisis. Contiene información cualitativa de los indicadores, sus campos se llaman **atributos**. Las dimensiones se relacionan con la tabla de hechos por su clave primaria. Todos los atributos de una dimensión pertenecen a un mismo grupo. Cuando más dimensiones se agregan mayor es la profundidad de análisis. La relación entre las dimensiones y las medidas brindan información como por ejemplo “ventas por producto” en un determinado tiempo, aquí se están relacionando la dimensión PRODUCTO y TIEMPO a través de la medida cantidad.



Ilustración 12: Representación de Dimensiones
Fuente: Propia

Continuando con las Bases de datos multidimensionales cabe recalcar la siguiente definición:

2.6.8 CUBOS

La representación gráfica de OLAP es lo que se conoce como “**Cubos**”. Un cubo OLAP es un vector multidimensional de N dimensiones en donde se almacena la información de manera ordenada facilitando un análisis rápido de sus contenido. Una base de datos multidimensional puede estar formada por varios cubos OLAP, dependiendo de sus requerimientos a analizar.

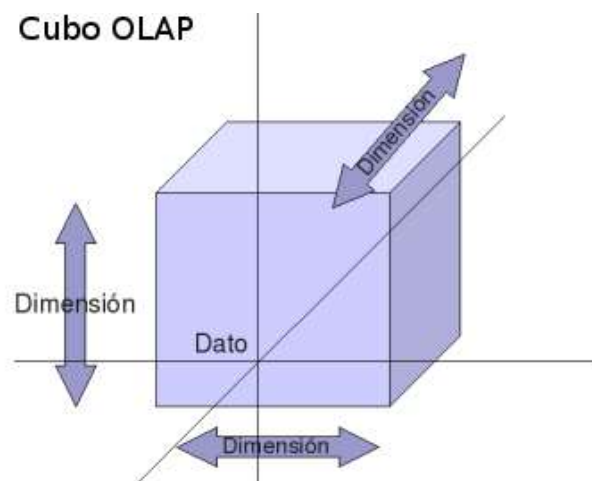


Ilustración 13: Representación gráfica de un CUBO

Fuente: zisko (2015). <https://zisko.files.wordpress.com/2011/02/cuboolap.png>

Los cubos tienen varios componentes que permiten trabajar con la información, los más importantes son:

Atributos: son los campos de las dimensiones, estos campos se forman de acuerdo a criterios de análisis, describen características que comparten los miembros de la dimensiones, por ejemplo en una dimensión Producto el atributo marca, permite buscar todos los productos de la misma marca.

Jerarquías: la jerarquía de dimensiones describe la relación lógica o jerárquica entre dos o más atributos miembros de la dimensión, utiliza niveles ordenados como medio de organización y agregación de datos, por ejemplo en la dimensión tiempo se podría tener una jerarquía para agregar datos desde el nivel mes al

nivel trimestre y al nivel año. Además una dimensión puede tener varias jerarquías por ejemplo en la dimensión tiempo también se puede analizar la jerarquía Día, mes y año.

Indicadores: son operaciones o sumalizaciones que se efectúan sobre algún hecho de la tabla de hechos.

Agregaciones

Así se le llama al proceso de pre calcular sumas de datos, para ayudar a disminuir los tiempos de respuestas, en los procesos de búsquedas de información.

2.6.9 MODELAMIENTO DE UN DATAWAREHOUSE

Los cubos OLAP se generan mediante esquemas sobre el Datawarehouse de la siguiente manera:

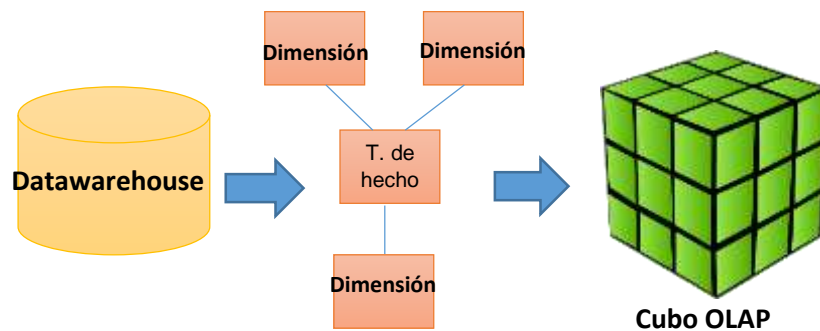


Ilustración 14: Modelamiento de un DATAWAREHOUSE

Fuente: Propia

Para modelar o diseñar un DWH (data ware house) existen varios esquemas que permiten organizar la información de acuerdo a los requerimientos de modelado, usados para modelar son:

2.6.9.1 ESQUEMA ESTRELLA

Este es uno de los favoritos a la hora de modelar un DWH (data ware house), el cual consiste en una tabla de hechos rodeada de dimensiones dando la

aparición de una estrella, de ahí se deduce su nombre. Lo característico de la arquitectura de estrella es que sólo existe una tabla de dimensiones para cada dimensión, esto quiere decir que la única tabla que tiene relación con otra es la de hechos, lo que significa que toda la información relacionada con una dimensión debe estar en una sola tabla. Se debe tomar en cuenta también el incluir siempre una dimensión tiempo ya que la información se mide en función del tiempo.

En la tabla de hechos se encuentran los atributos destinados a medir o cuantificar el hecho es decir las 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). Las tablas de dimensión se encuentran además totalmente desnormalizadas, es decir, toda la información referente a una dimensión se almacena en la misma tabla.

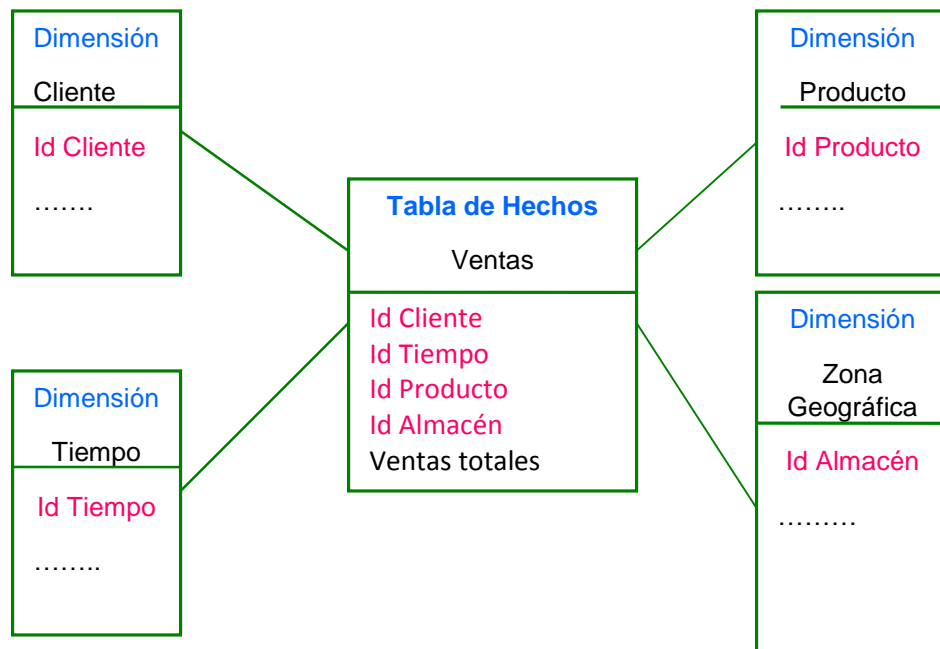


Ilustración 15: Esquema Estrella
Fuente: Propia

2.6.9.2 ESQUEMA COPO DE NIEVE

Este esquema se deriva del esquema estrella, en el que las tablas de dimensiones se normalizan en múltiples tablas, y la tabla de hechos deja de ser la única relacionada con otras tablas ya que existen otras tablas que se relacionan con las dimensiones y que no tienen relación directa con la tabla de hechos. Este modelo fue creado para facilitar el mantenimiento de las dimensiones. En la estructura dimensional normalizada, la tabla que representa el nivel base de la dimensión es la que hace relación directamente con la tabla de hechos. La diferencia entre ambos esquemas reside en la estructura de las tablas de dimensión. Para conseguir un esquema en copo de nieve se debe 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 desnormalizadas, ahora se dividen en subtablas tras un proceso de normalización. Es posible distinguir dos tipos de esquemas en copo de nieve, un esquema completo en el que todas las tablas de dimensión en el esquema en estrella aparecen ahora normalizadas o un esquema parcial donde sólo se lleva a cabo la normalización de algunas de las dimensiones.

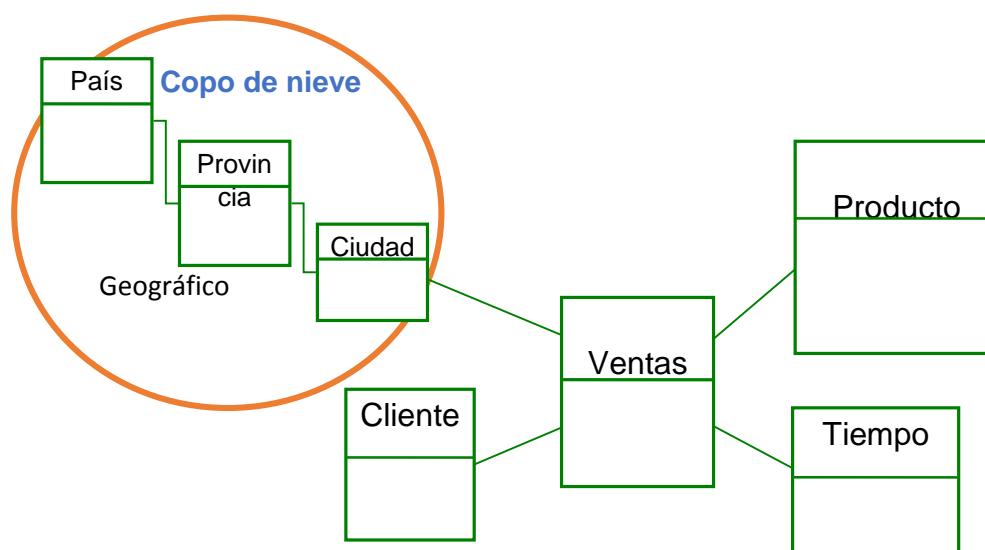


Ilustración 16: Esquema Copo de Nieve

Fuente: Propia

2.6.10 HERRAMIENTAS DE VISUALIZACIÓN

Existen varias herramientas que permiten trabajar sobre la información contenida en los Datawarehouse y los Datamart, su objetivo principal es permitir acceder, analizar y visualizar dicha información facilitando la elaboración de informes tanto en detalle como sobre información agregada.

Principalmente estas herramientas proporcionan una interfaz de usuario, en la que se puede encontrar varias opciones de explotar la información, va desde el uso de Excel hasta por ejemplo la construcción de Dashboards, Cuadros de mando, etc. En la actualidad existen varias herramientas muy robustas que permiten el análisis y la visualización, de la información, varias de ellas son:

- Power Pivot, Power View y Power Map para Excel: comúnmente es usado como herramienta de análisis y visualización de datos por muchas personas ya que la mayoría están familiarizados con su entorno.
- Microsoft SQL Reporting Services.
- Pentaho Report Designer.
- Oracle Administration Tools.
- Qlik View.
- Power BI.

2.7 SQL SERVER BUSINESS INTELLIGENCE 2014

2.7.1 DEFINICIÓN Y CARACTERÍSTICAS

SQL Server es un sistema de administración y análisis de bases de datos, que permite la creación de Datawarehouse y la implementación de soluciones de Inteligencia de Negocios, para lo cual dispone de una amplia gama de herramientas. La versión de SQL Server Business Intelligence 2014 trabaja con la ayuda de SQL Server Data Tools Business Intelligence el cual permite hacer una administración sencilla de SQL Server en lo que respecta al ambiente de desarrollo, ya que se integra con Microsoft Visual Studio 2012 o también el 2013.

Este componente en su modulo de Inteligencia de negocios permite crear los proyectos de Analisis Services, Integration Services, y Reporting Services con el origen de datos del DWH que se haya diseñado en el motor de base de datos SQL Server Management Studio.

Esta versión de SQL Server tiene varias características nuevas en relación a sus versiones anteriores, principalmente en lo que se refiere al motor de base de datos con pequeños cambios, en el Analysis Services se ha añadido un soporte para Power Pivot y en su instalación también pequeños cambios, todas estas nuevas características tienen el objetivo de aumentar la eficacia y productividad de SQL Server 2014. Entre las principales ediciones que presenta el SQL Server 2014 son:

Edición de SQL Server	Definición
Enterprise (64 bits y 32 bits)	Proporciona capacidades de centro de datos de tecnología avanzada completas con un rendimiento ultrarrápido, virtualización ilimitada y Business Intelligence integral, que habilita los mayores niveles de servicio para las cargas de trabajo de gran importancia y el acceso del usuario final a ideas claras de los datos.
Business Intelligence (64 bits y 32 bits)	SQL Server 2014 Business Intelligence Edition ofrece una plataforma completa que capacita a las organizaciones para crear e implementar soluciones de BI seguras, escalables y fáciles de administrar. Proporciona funcionalidad emocionante, como exploración y visualización de datos en un explorador; funciones eficaces de mezcla de datos y administración de integración mejorada.
Standard (64 bits y 32 bits)	SQL Server 2014 Standard Edition proporciona de administración básica de bases de datos y base de datos de Business Intelligence para que los departamentos y pequeñas organizaciones ejecuten sus aplicaciones y admite las herramientas de desarrollo comunes, tanto locales como en la nube, que habilitan la administración eficaz de bases de datos con recursos de TI (tecnologías de información) mínimos.

Tabla 6: Ediciones de SQL Server 2014

Fuente: Microsoft(2015). <https://msdn.microsoft.com/es-es/library/ms144275.aspx>

La arquitectura general de Inteligencia de Negocios en Microsoft es la siguiente:

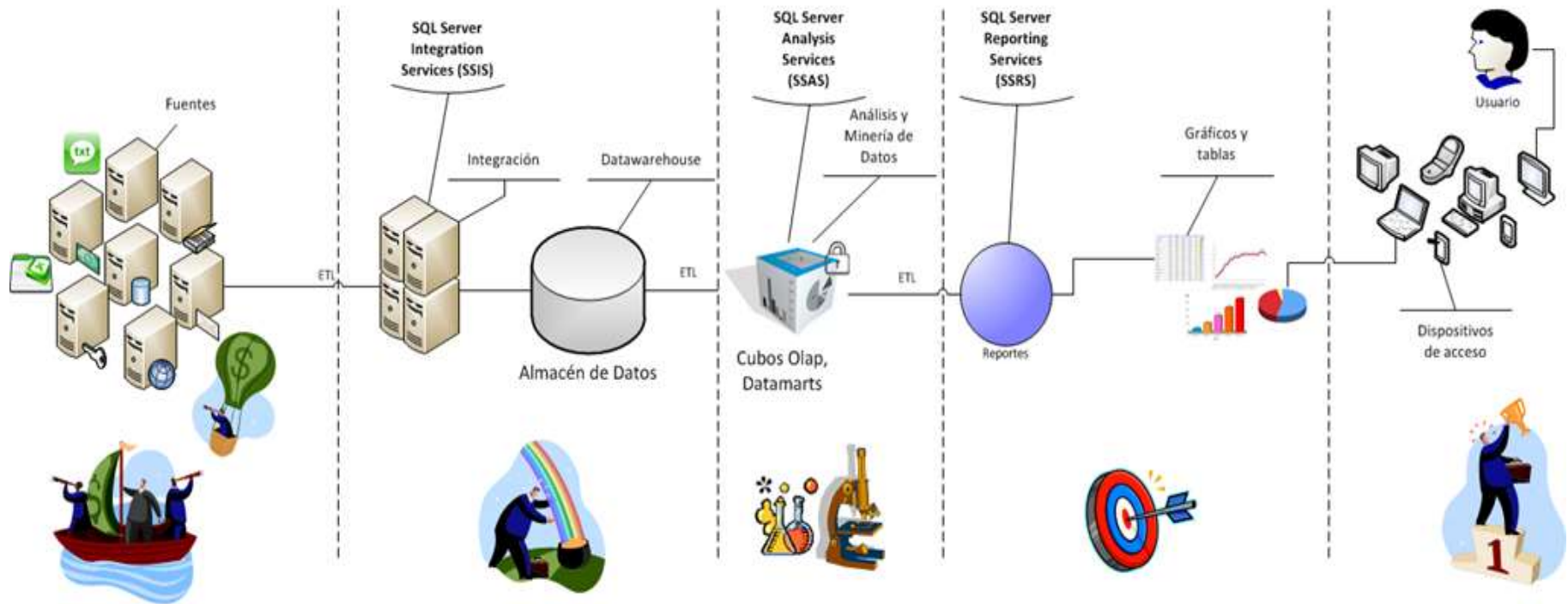


Ilustración 17: Arquitectura de Inteligencia en Negocios en Microsoft.

Fuente: blogs.itsynergy (2015). <http://blogs.itsynergy.co/jvelezc/2011/02/18/arquitectura-de-inteligencia-de-negocios/>

2.7.2 SQL SERVER DATA TOOLS

Conocido por sus siglas (SSDT) SQL Server Data Tools, para Visual Studio, como se indica anteriormente incluye herramientas para la creación de proyectos de inteligencia empresarial y plantillas de proyectos para SQL Server como son: Analysis Services, Reporting Services e Integration Services integrados en el shell de Visual Studio. Para la instalación de este paquete se dispone la opción de instalarlo de manera integrada al SQL Server, solo los componentes para inteligencia de negocios o por separado con todos sus componentes, pero es más recomendable para proyectos de Inteligencia de Negocios optar por la primera opción.

2.7.3 ANALYSIS SERVICES

Conocido por sus siglas (SSAS) SQL Server Analysis Services, permite diseñar, crear y administrar estructuras multidimensionales que contienen datos agregados desde varios orígenes de datos, como bases de datos transaccionales. Incluye las funciones OLAP(Procesamiento Analítico en línea) y también minería de datos (Data Mining) .

Una vez que este diseñado y poblada la base de datos multidimensional se procede a referenciarla como origen de datos desde Analysis Server, y se crea los cubos con la estructura requerida para el análisis de datos.

Además proporciona dos enfoques para modelar los datos como son el modelo de datos Tabular y el modelo de datos multidimensional.

Las soluciones multidimensionales: y de minería de datos usan construcciones de modelado OLAP (cubos y dimensiones) y almacenamiento MOLAP, ROLAP u HOLAP que usa el disco como almacenamiento de datos principal para los datos previamente agregados.

Las soluciones tabulares: usan construcciones de modelado relacional como tablas y relaciones para modelar los datos y el motor de análisis en memoria para almacenarlos y calcularlos. La mayor parte del modelo, si no en su totalidad, se almacena en RAM y suele ser mucho más rápido que su homólogo multidimensional. Admiten una gama más amplia de orígenes de datos, incluidos archivos sin formato, fuentes de datos y orígenes de datos a los que se obtiene acceso a través de proveedores de datos ODBC.

Tabla 7: Tipos de soluciones de Analysis Services

Fuente: microsoft (2015). <https://msdn.microsoft.com/es-ec/library/hh212940.aspx>

En el siguiente grafico se puede apreciar varias características en cuanto a los dos modelos, es decir el modelo semántico de BI en Analysis Services 2014.

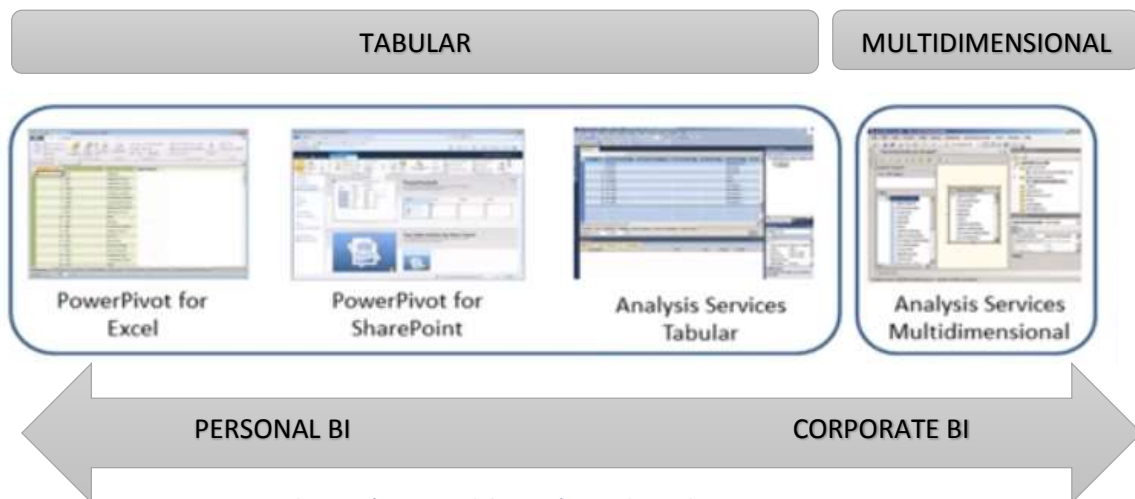


Ilustración 18: Modelo Semántico de Analysis Services

Fuente: Propia

De entre los dos tipos de modelos, se puede decir que los modelos tabulares están más sugeridos para análisis de tipo personal, departamental y de servidor, mientras que el modelo multidimensional está sugerido para el análisis corporativo, por el tipo de herramientas que contiene.

Para el presente caso de estudio se realiza un diseño para cada modelo, con el propósito de realizar una comparación entre el procedimiento que utilizan estos dos enfoques, pero finalmente se realiza la demostración del análisis y presentación de reportes de datos con el modelo tabular que viene incrustado en Excel a través de Power Pivot y Power View.

2.7.3.1 POWER PIVOT

Las herramientas de análisis de datos para Inteligencia de Negocios se han manejado casi en su mayoría de manera reservada para la parte técnica a través del uso de herramientas como Analysis Services, entre otras, pero Microsoft en este aspecto y conociendo de las ventajas poderosas que brinda Excel ve la necesidad de agregarle nuevos componentes, para que sea el usuario final que necesita la información quien la analice y además lo realice desde un ambiente familiar que es el que proporciona Excel, creando así un autoservicio. Por este motivo agrega entre otros componentes a Power Pivot y Power View a Excel, los cuales se usan en el presente proyecto para generar informes.

Power Pivot es el motor de Analysis Services incorporado dentro de Excel, que proporciona un desempeño superior al que maneja Excel normalmente, ya que desde ese entorno se puede realizar el análisis de datos para Inteligencia de Negocios, entre sus principales características se detalla las siguientes:

- Incorpora el poder de las bases de datos columnares comprimidas ya que permite en un libro de Excel guardar millones de registros y con respuestas rápidas.
- Esta nueva tecnología permite la conexión desde Excel a múltiples fuentes de datos, y permite trabajar con varias fuentes a la vez.

- Proporciona la posibilidad de tener un mini cubo en Excel que se lo puede llevar donde quiera ya que no necesita estar conectado a la fuente, puesto que Power Pivot entrega una copia comprimida de los datos que se están consultando.
- Utiliza un motor de memoria llamado **XVelocity**¹⁰ que aprovecha la memoria RAM¹¹ para con un modelo tabular mejorar la velocidad con respecto al modelo multidimensional.
- Al igual que un modelo multidimensional permite la posibilidad de crear KPI, jerarquías, cálculos, relaciones etc, la información se la puede observar también en diagramas relacionados.
- Utiliza el código DAX¹² para los cálculos o consultas.

En cuanto a los límites de información que maneja Power Pivot se detallan los siguientes:

- Para 32 bits el tamaño máximo que puede tener en una hoja de Excel es de 2 gigabyte.
- Para 64 bits no tiene límites en el tamaño del archivo pero si se limita por el uso de memoria.

2.7.3.2 POWER VIEW

Esta herramienta incorporada a Microsoft Excel 2013, permite la visualización, exploración y presentación de reportes de datos o informes ad hoc, es decir permite interactuar con los datos para presentarlos de manera interactiva en cuadros, tableros, filtros, tipo cuadros de mando, incorpora varias herramientas que permiten por ejemplo clasificar datos, realizar gráficos de burbuja interactivos, escalas de tiempo, vistas desplegadas entre otros; en si Power View es una herramienta intuitiva para representaciones graficas de datos.

¹⁰ XVelocity: tecnología de administración de datos optimizada para memoria.

¹¹ RAM: memoria principal

¹² DAX: data analysis expresion, lenguaje de fórmulas para modelos tabulares

Para poder referenciar la pestaña de Power View dentro de Excel se debe tener instalado el componente Silverlight, el mismo que se encuentra disponible en la página de Microsoft.

Se lo puede usar con dos versiones, la una para Excel en donde Power View forma parte del archivo Excel XLSX como una hoja más, y la otra desde Share Point como archivos RDLX. En el presente proyecto los datos de origen para los informes con Power View se los toma del modelo de datos que se crea con Power Pivot.

En cuanto al acceso a los datos, Power View permite interactuar con los datos de varias maneras: en el mismo libro de Excel que la hoja de Power View, en los modelos de datos de los libros de Excel publicados por Power Pivot, o en los modelos tabulares y multidimensionales implementados en Analysis Services.

En el presente proyecto se realiza la elaboración de un reporte dinámico, tomando los datos cargados con el Power Pivot para analizarlos desde varios puntos de vista en una hoja de Power View.

2.7.4 INTEGRATION SERVICES

Conocido por sus siglas (SSIS) SQL Server Integration Services, permite llevar a cabo los procesos de Extracción y Transformación desde varios orígenes para luego cargar estos datos a su destino o repositorio de datos. Incluye herramientas gráficas y asistentes para crear, administrar y depurar paquetes, pero también brinda la posibilidad de referenciar o programar los procesos de manera personalizada, de acuerdo a la necesidad de cada proyecto.

Al trabajar en SQL Server Data Tools, se puede realizar las siguientes tareas:

- Ejecutar el Asistente para importación y exportación de SQL Server para crear paquetes básicos que copian datos de un origen en un destino.
- Crear paquetes que incluyan flujo de control complejo, flujo de datos, lógica controlada por eventos y registro.

- Probar y depurar paquetes mediante las características de solución de problemas y supervisión en el Diseñador SSIS, y las características de depuración en SQL Server Data Tools (SSDT).
- Crear configuraciones que actualizan las propiedades de los paquetes y los objetos de paquete en el tiempo de ejecución.
- Guardar copias de paquetes en la base de datos de SQL Server, el Almacén de paquetes de SSIS y el sistema de archivos.

2.7.5 REPORTING SERVICES

Conocido por sus siglas SSRS (SQL Server Reporting Services) es la herramienta que permite la ilustración de la información que se encuentra en los cubos de una forma amigable con el usuario final; es decir permite la creación de informes en donde se debe especificar el origen de datos, el tipo de datos y como se van a mostrar. Al ejecutar el informe, el procesador de informes toma toda la información especificada, recupera los datos y los combina con el diseño del informe para generar el informe. Puede obtener una vista previa de los informes en el generador de informes o publicar el informe en un servidor de informes, donde otros usuarios podrán ejecutarlo.

Algunas características de este paquete son:

- Permite exportar los reportes a distintos formatos
- Cuenta con un servicio para la distribución de reportes
- Permite explotar los reportes desde aplicaciones personalizadas.
- La administración y consumo de reportes también puede ser a través de webservices con los que cuenta.
- Permite mantener historial de reportes ejecutados, además de almacenar información en cache para la rápida ejecución de los reportes.

Como referencia al tema y según la fuente de Microsoft.com, SQL Server 2014 a parte de las tecnologías ya mencionadas SQL Server incluye también las siguientes tecnologías:

	<p>Motor de base de datos</p> <p>El Motor de base de datos es el servicio principal para almacenar, procesar y proteger los datos. El Motor de base de datos proporciona acceso controlado y procesamiento rápido de transacciones para satisfacer los requerimientos de las aplicaciones consumidoras de datos más exigentes de su empresa. El Motor de base de datos también proporciona un amplio soporte para mantener una alta disponibilidad.</p>
	<p>Data Quality Services</p> <p>SQL Server Data Quality Services (DQS) le proporciona una solución de limpieza de datos basada en el conocimiento. DQS permite generar una base de conocimiento y usarla para realizar tareas de corrección de datos y eliminación de datos duplicados, usando medios asistidos por ordenador e interactivos. Puede usar servicios de consulta de datos basados en la nube y puede generar una solución de administración de datos que integra DQS con SQL Server Integration Services y Master Data Services.</p>
	<p>Master Data Services</p> <p>Master Data Services es la solución de SQL Server para la administración de datos maestros. Una solución basada en Master Data Services ayuda a asegurarse de que los informes y los análisis se basan en la información correcta. Con Master Data Services, se crea un repositorio central de los datos maestros y se mantiene un registro auditable y protegible de los mismos a medida que van cambiando con el tiempo.</p>


	<p>Replicación</p> <p>La replicación es un conjunto de tecnologías destinadas a la copia y distribución de datos y objetos de base de datos de una base de datos a otra, para luego sincronizar ambas bases de datos con el fin de mantener su coherencia. La replicación permite distribuir datos a diferentes ubicaciones y a usuarios remotos o móviles mediante redes de área local y de área extensa, conexiones de acceso telefónico, conexiones inalámbricas e Internet.</p>
---	--

Tabla 8: Tecnologías de SQL Server

Fuente: Microsoft(2015). <https://msdn.microsoft.com/es-es/library/ms130214.aspx>

CAPITULO 3

DISEÑO DEL DATAMART DE VENTAS



- ✓ PLANIFICACIÓN
- ✓ ANÁLISIS Y DISEÑO
- ✓ DEFINICIÓN Y CARACTERÍSTICAS
- ✓ CONSTRUCCIÓN Y DESARROLLO
- ✓ PRUEBAS

3. DISEÑO DEL DATAMART DE VENTAS

Para el diseño del Data Mart del área de ventas, en el presente proyecto se lo realiza en base a las etapas del ciclo de vida del proyecto, propuestas por Ralph Kimball, ya que como se ha analizado antes en esta metodología es en donde se propone en primera instancia la creación del Data Mart, partiendo directamente de las fuentes de datos. Por lo tanto se plantea las siguientes fases de desarrollo:



Ilustración 19: Fases de diseño del DataMart

Fuente: Propia

3.1 PLANIFICACIÓN

El objetivo de esta fase es planear las **actividades** a ser ejecutadas, definir los **recursos** a emplear e identificar los **requerimientos**.

Para el presente proyecto cabe recalcar que se está representando un caso de estudio de un modelo de gestión del departamento de ventas de una empresa real y en producción de la cual se obtuvo la base de datos fuente.

3.1.1 ACTIVIDADES Y ALCANCE

En el presente caso de estudio el objetivo es definir y diseñar un repositorio de datos empresarial para consulta y análisis, que integre los procesos de ventas. Es por esto que se propone diseñar un Data Mart con los requerimientos del área de ventas para presentar los reportes gerenciales más importantes de ventas y análisis de rentabilidad, con la información que se encuentren almacenada en la base de datos origen.

Entre las actividades principales a realizar se detallan las siguientes:

1.	Identificación y análisis de requerimientos de la funcionalidad del área de ventas basándose en la información de análisis existente en los datos fuente.
2.	Diseño y Construcción de un almacén de datos o Data Mart que resida en una base de datos de SQL Server 2014.
3.	Brindar la posibilidad de generar análisis de información y reportes de ventas desde la bodega de datos.
4.	Desarrollar un proceso ETL (Extracción, Transformación y Carga) para la carga de información desde la base fuente hasta la nueva bodega de datos.
5.	Obtención de conocimiento en manejo de herramientas de desarrollo utilizadas.
6.	Construcción de reportes a través de Excel.

Tabla 9: Tabla de actividades

Fuente Propia

3.1.2 DEFINICIÓN DE RECURSOS

Entre los recursos para el diseño del presente proyecto se detallan los más importantes, tanto en lo tecnológicos como externos:

Recursos de Información	Lo más importante para iniciar un proyecto de Inteligencia de negocios es la base de datos fuente.
-------------------------	--

Recursos Tecnológicos de Software	La herramienta para construcción del proyecto, en este caso SQL Server 2014, SQL Server Data Tools, Internet, Excel 2013, Motor de base de datos ORACLE, entre otras.
Recursos Tecnológicos de Hardware	Computador con las siguientes especificaciones: Memoria de 4GB DDR3, Procesador Intel Core i5 – 2430M 2.4GHz, Disco duro de 500 GB, entre otros.
Varios	Materiales de oficina.

Tabla 10: Definición de recursos
Fuente: Propia

3.1.3 IDENTIFICACIÓN DE REQUERIMIENTOS

Los requerimientos del departamento de ventas sujeto del presente análisis abarca temas con respecto a:

Análisis de clientes: ¿Quiénes son los mejores clientes? ¿Cómo ha sido su movimiento durante el último año?, por ejemplo para dependiendo de esto dar un trato especial a los mejores clientes.

Análisis de territorio: ¿Dónde están localizados los mejores clientes? ¿En qué zona están los clientes?

Análisis de artículos: ¿cuáles fueron los artículos más vendidos? ¿Qué artículos tienen menos acogida?

Análisis de vendedores: ¿Cuáles fueron los mejores vendedores?

Todos los análisis que se realizan son en función del **Tiempo**, por lo cual es una dimensión obligatoria.

En este contexto cabe indicar que la empresa de la cual se ha obtenido la base fuente que se analiza en este proyecto, nombrada en adelante como XX, se dedica principalmente a la comercialización al por mayor de varias líneas de productos de consumo masivo en la zona norte del país. Dicha empresa para garantizar una gestión de ventas más óptima cuenta con un grupo capacitado de

vendedores encargados de contactar al cliente para promocionar los productos y dar un servicio personalizado en el proceso de ventas y distribución, el mismo que se orienta a mantener las relaciones comerciales con los clientes y satisfacer sus necesidades.

El proceso de ventas se lo realiza de acuerdo a la clase de clientes, por ejemplo, tienda, bazares, abarrotes, entre otros, los cuales recibirán una proforma de productos y precios por parte de los vendedores, para ingresar su pedido, luego tendrán la posibilidad de llenar una solicitud para acceder al crédito de pago, máximo de 15 días, finalmente se le entrega la factura como comprobante de la transacción realizada. El proceso de ventas que se lleva a cabo se lo puede apreciar a manera general en el siguiente diagrama:

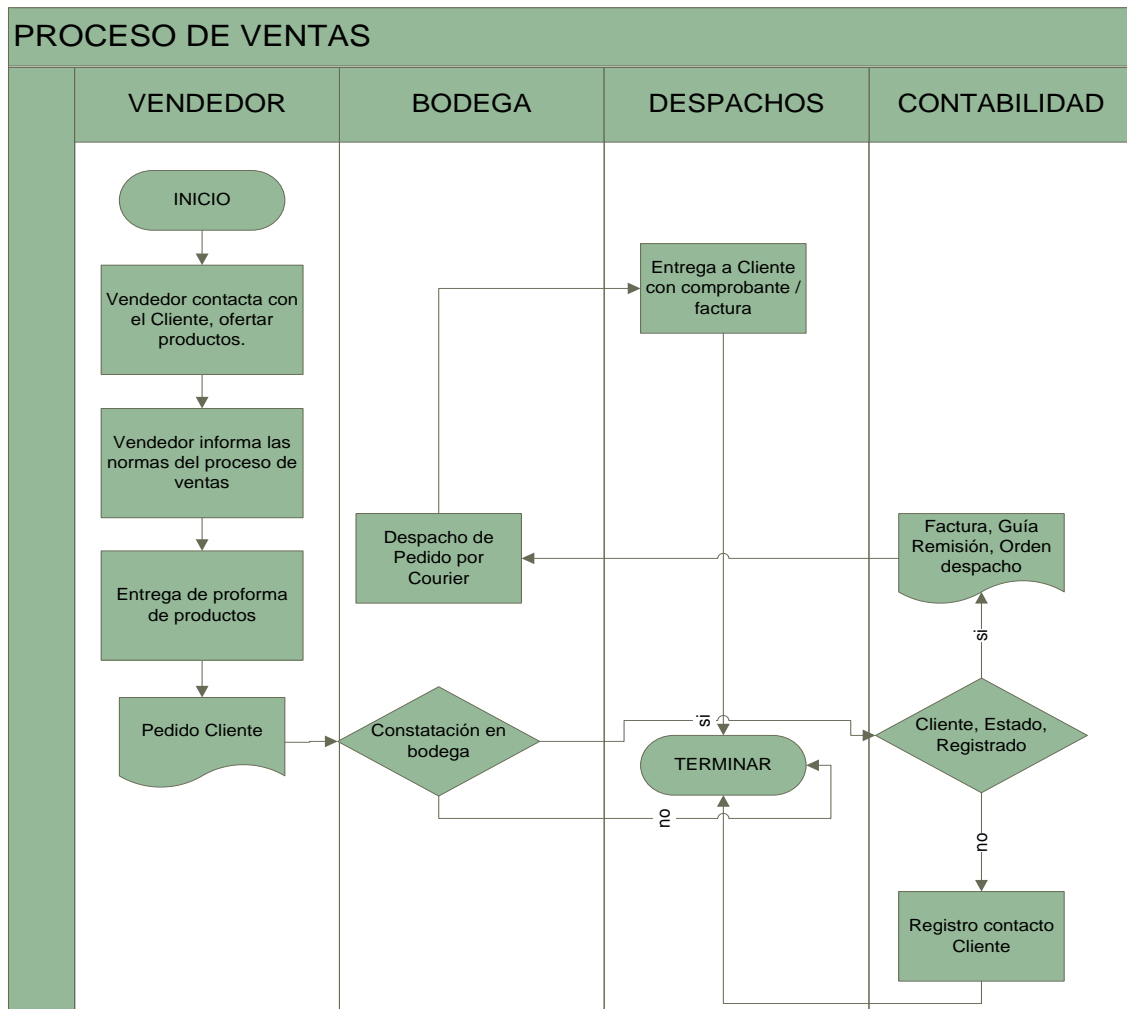


Ilustración 20: Diagrama de Proceso de ventas

Fuente: Propia

Del análisis realizado hasta este punto, se ha podido determinar que el área de ventas implica la relación entre las siguientes variables.

PROCESO DE NEGOCIO	Clientes	Productos	Vendedor	Fecha
Proceso de Ventas	X	X	X	X

Tabla 11: Matriz de proceso de Ventas

Fuente: Propia

En cuanto a la definición de requerimientos, después de haber analizado la fuente de datos y haber realizado un primer análisis de proceso del departamento de ventas, se encontraron las siguientes **variables** generales de análisis de que satisfacen las necesidades actuales y que permiten la organización y construcción del DataMart de ventas.

TIEMPO	Año	Para el análisis de ventas en función del tiempo
	Meses	
	Días	

Ilustración 21: Variables de Tiempo

Fuente: Propia

VENDEDOR	Información general	Para análisis de ventas por vendedor
	Clientes a cargo	
	Información Comisiones	

Ilustración 22: Variables Vendedor

Fuente: Propia

COMERCIALES	Pago	Crédito, Efectivo
	Tipo Cliente	Tienda, Abarrotes, Minimarket, Bazares entre otros.
	Bodegas	Ubicación
	Cliente	Información General
	Estado	Cliente Activo / Inactivo
	Indicador	Crédito Abierto / Cerrado
	Zonas Clientes	

Ilustración 23: Variables Comerciales

Fuente: Propia

DESCRIPCIÓN PRODUCTO	Línea	Árbol de Producto
	Familia o Grupo	Árbol de Producto
	Genérico o Subgrupo	Árbol de Producto
	Unidad medida	Prefijos, nombre
	Proveedor	Nombre
	Bodega	Nombre
	Descripción	Nombre Producto

Ilustración 24: Variables de Producto

Fuente: Propia

NUMÉRICO PRODUCTO	Cantidad Vendida	En unidades
	Número de Factura	
	Valor Total	En USD
	Costo Unitario	En USD
	Utilidad	En USD
	Número de Facturas	

Ilustración 25: Variables Numéricas

Fuente: Propia

3.2 ANÁLISIS Y DISEÑO

El objetivo de esta fase es definir el alcance de cada una de las áreas de análisis en sus medidas, dimensiones y atributos. Así mismo la creación del modelo de datos lógico y físico, que soporta el área de análisis.

Para definir correctamente las dimensiones y tabla de hechos, se analiza los datos disponibles de la base de datos fuente los mismos que son procesados para ser cargados a la base de datos multidimensional.

3.2.1 ORÍGENES DE DATOS

Como origen de datos se dispone de la base de datos que soporta los procesos comerciales de la empresa XX sujeto del presente análisis, de la cual se ha identificado en base a los requerimientos del área de ventas, la información representada por el siguiente diagrama en el cual se detalla las tablas de origen que son necesarias para cada dimensión:

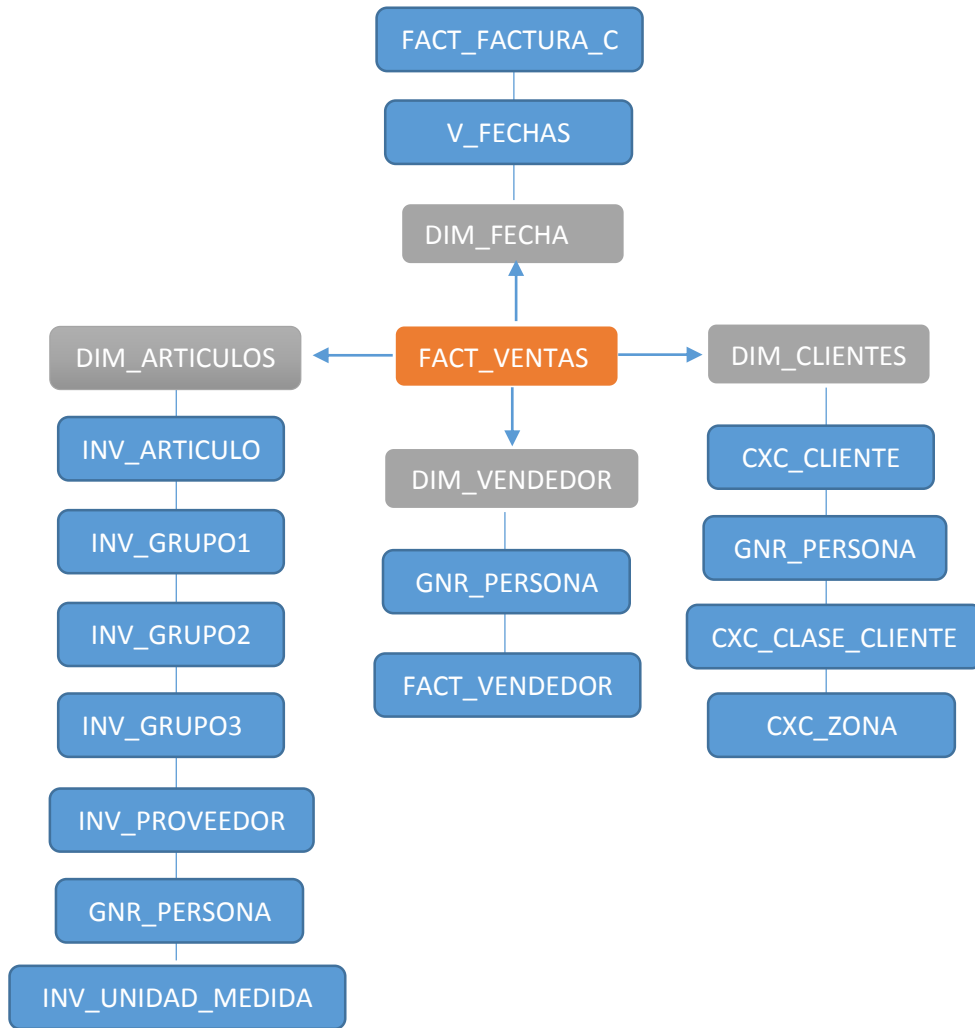


Ilustración 26: Mapa de orígenes de datos
Fuente: Propia

Las tablas de la base de datos transaccional fuente originales enunciadas en el diagrama anterior, con sus campos se detallan a continuación:

- **ORIGEN DE DATOS PARA LA DIMENSIÓN CLIENTES**

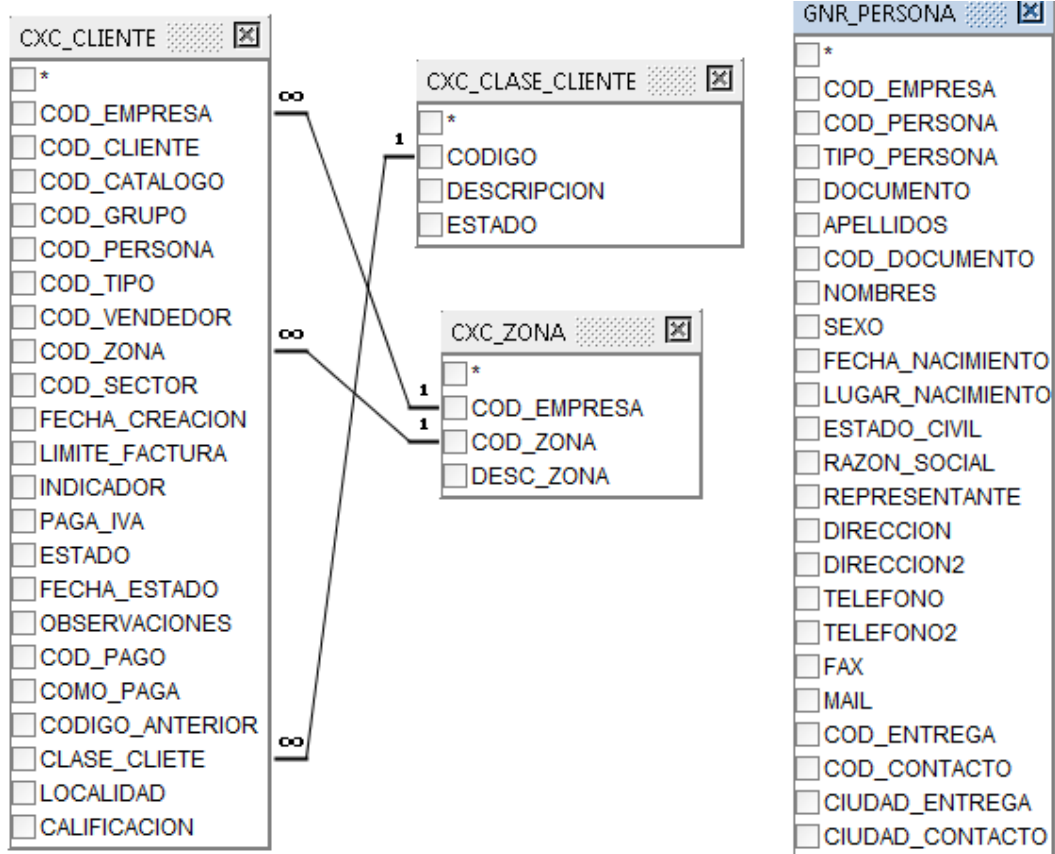


Ilustración 27: Origen Dimensión cliente

Fuente: Propia

- **ORIGEN DE DATOS PARA LA DIMENSIÓN ARTÍCULOS**

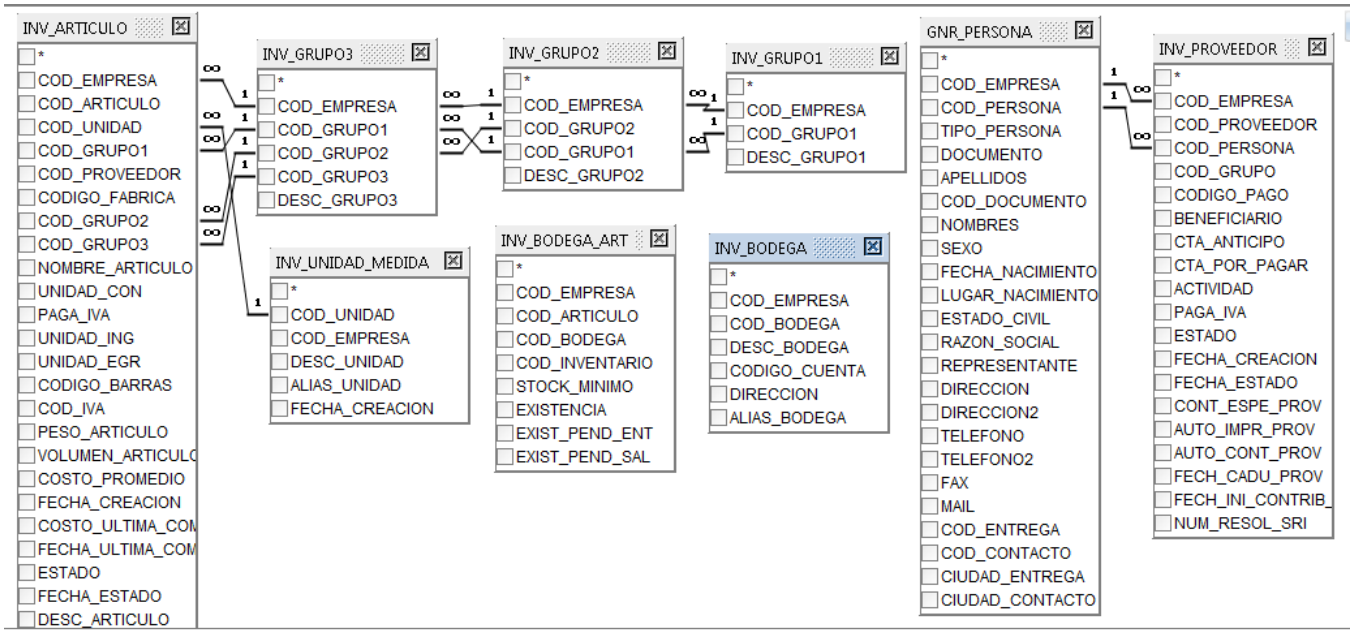


Ilustración 28: Origen Dimensión Artículos

Fuente: Propia

- **ORIGEN DE DATOS PARA LA DIMENSIÓN VENDEDOR**

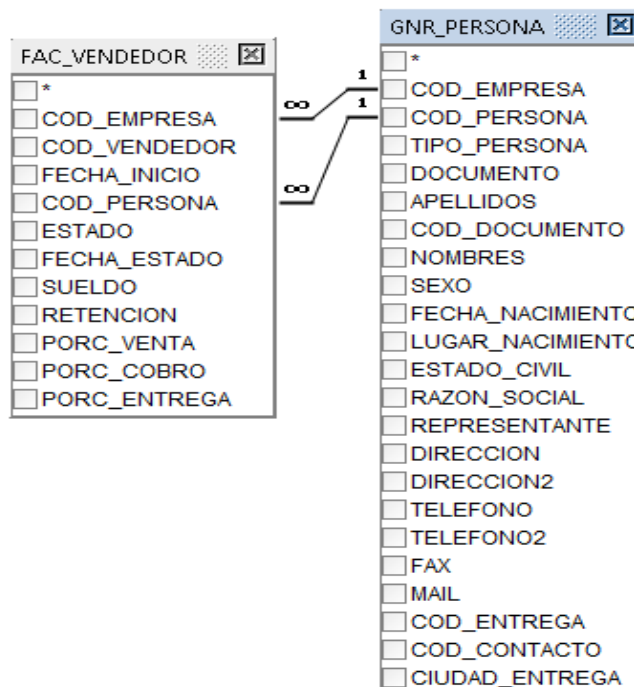


Ilustración 29: Origen Dimensión Vendedor

Fuente: Propia

- **ORIGEN DE DATOS PARA LA DIMENSIÓN FECHA Y LA TABLA DE HECHOS**

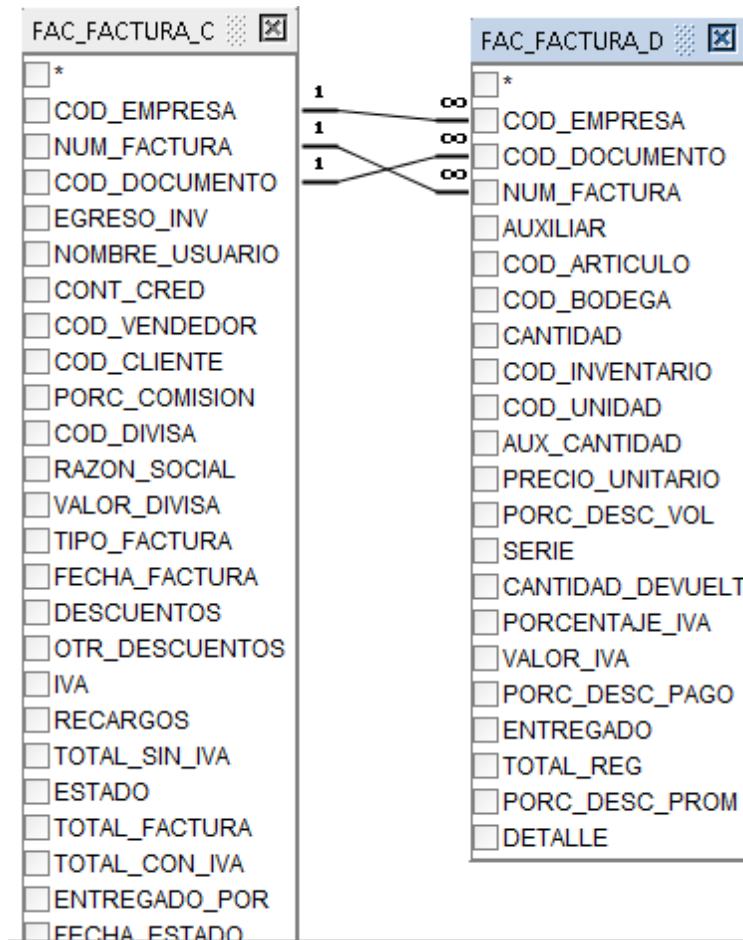


Ilustración 30: Origen Dimensión Fecha y Tabla de Hechos
Fuente: Propia

3.2.2 MAPEO DE LA INFORMACIÓN

Una vez identificado el grupo de información necesaria para alimentar las dimensiones y tabla de hechos, se puede describir de manera detallada la información que se toma del origen y a que tablas de destino que forman las dimensiones corresponden, así como también la tabla de hechos. Esta información detallada sirve fundamentalmente para construir de manera correcta el modelo de datos multidimensional. A continuación se detalla el mapeo de datos correspondiente a cada dimensión y tabla de hechos:

DIMENSIÓN CLIENTE

DESCRIPCIÓN Almacena todos los datos referentes a los Clientes

DESTINO			CAMPO	ORIGEN		
NOMBRE TABLA	CAMPO	TIPO DATO	NAVEGACION	NOMBRE TABLA	CAMPO	TIPO DATO
DIM_CLIENTE	ID_CLIENTE	NUMERIC(15,0)	COD_CLIENTE	CXC_CLIENTE	COD_CLIENTE	NUMBER
	APELLIDOS	VARCHAR(100)	COD_PERSONA	GNR_PERSONA	APELLIDOS	VARCHAR2(100 BYTE)
	NOMBRES	VARCHAR(100)		GNR_PERSONA	NOMBRES	VARCHAR2(100 BYTE)
	RAZON_SOCIAL	VARCHAR(100)		GNR_PERSONA	RAZON_SOCIAL	VARCHAR2(100 BYTE)
	DIRECCION	VARCHAR(100)		GNR_PERSONA	DIRECCION	VARCHAR2(100 BYTE)
	TELEFONO	VARCHAR(50)		GNR_PERSONA	TELFONO	VARCHAR2(20 BYTE)
	CIUDAD_ENTREGA	VARCHAR(100)		GNR_PERSONA	CIUDAD_ENTREGA	VARCHAR2(100 BYTE)
	CREDITO	VARCHAR(10)		INDICADOR	CXC_CLIENTE	INDICADOR
	ESTADO	VARCHAR(10)	ESTADO	CXC_CLIENTE	ESTADO	VARCHAR2(1BYTE)
				EXCEL, INDICDORES	DESCRIPCION	
	LIMITE_FACTURA	NUMERIC(18,2)	COD_CLIENTE	CXC_CLIENTE	LIMITE_FACTURA	NUMBER
	CLASE_CLIENTE	VARCHAR(50)	CLASE_CLIENTE	CXC_CLIENTE	CLASE_CLIENTE	VARCHAR2(20 BYTE)
			CODIGO	CXC_CLASE_CLIENTE	DESCRIPCION	VARCHAR2(50 BYTE)
	ZONA	VARCHAR(100)	COD_ZONA	CXC_CLIENTE	COD_ZONA	VARCHAR2(2 BYTE)
				CXC_ZONA	DESC_ZONA	VARCHAR2(100 BYTE)

Tabla 12: Mapeo Dimensión Cliente

Fuente: Propia

DIMENSIÓN ARTÍCULOS

DESCRIPCIÓN Almacena todos los datos referentes a los Artículos

DESTINO			CAMPO	ORIGEN		
NOMBRE TABLA	CAMPO	TIPO DATO	NAVEGACION	NOMBRE TABLA	CAMPO	TIPO DATO
DIM_ARTICULOS	ID_ARTICULO	NUMERIC(15,0)	COD_ARTICULO	INV_ARTICULO	COD_ARTICULO	NUMBER(8,0)
	LINEA	VARCHAR(50)	COD_GRUPO1	INV_ARTICULO	COD_GRUPO1	VARCHAR2(2 BYTE)
				INV_GRUPO1	DESC_GRUPO1	VARCHAR2(40 BYTE)
	GRUPO	VARCHAR(50)	COD_GRUPO1	INV_ARTICULO	COD_GRUPO2	VARCHAR2(2 BYTE)
				INV_GRUPO2	DESC_GRUPO2	VARCHAR2(40 BYTE)
	SUBGRUPO	VARCHAR(50)	COD_GRUPO1	INV_ARTICULO	COD_GRUPO2	VARCHAR2(2 BYTE)
				INV_GRUPO2	COD_GRUPO3	VARCHAR2(2 BYTE)
				INV_GRUPO3	DESC_GRUPO3	VARCHAR2(40 BYTE)
	NOMBRE	VARCHAR(100)	COD_ARTICULO	INV_ARTICULO	NOMBRE_ARTICULO	VARCHAR2(80 BYTE)
	UNIDAD	VARCHAR(50)	COD_UNIDAD	INV_ARTICULO	COD_UNIDAD	VARCHAR2(2 BYTE)
				INV_UNIDAD_MEDIDA	DESC_UNIDAD	VARCHAR2(40 BYTE)
	CODIGO_FABRICA	VARCHAR(50)	COD_ARTICULO	INV_ARTICULO	CODIGO_FABRICA	VARCHAR2(40 BYTE)
	PROVEEDOR	VARCHAR(100)	COD_PERSONA	INV_ARTICULO	COD_PROVEEDOR	NUMBER(9,0)
INV_PROVEEDOR				COD_PROVEEDOR	NUMBER(9,0)	
GNR_PERSONA				RAZON_SOCIAL	VARCHAR2(100 BYTE)	
PESO	NUMERIC(18,2)	COD_ARTICULO	INV_ARTICULO	PESO_ARTICULO	NUMBER	
FECHA_CREACION	DATE	COD_ARTICULO	INV_ARTICULO	FECHA_CREACION	DATE	

Tabla 13: Mapeo Dimensión Artículos

Fuente: Propia

DIMENSIÓN VENDEDOR

DESCRIPCIÓN Almacena todos los datos referentes a los Vendedores

DESTINO			CAMPO	ORIGEN		
NOMBRE TABLA	CAMPO	TIPO DATO	NAVEGACION	NOMBRE TABLA	CAMPO	TIPO DATO
DIM_VENDEDOR	ID_VENDEDOR	NUMERIC(18,0)	COD_VENDEDOR	FAC_VENDEDOR	COD_VENDEDOR	NUMBER(9,0)
	NOMBRE	VARCHAR(100)	COD_PERSONA	GNR_PERSONA	RAZON_SOCIAL	VARCHAR2(100 BYTE)
	DIRECCION	VARCHAR(100)	COD_PERSONA	GNR_PERSONA	DIRECCION	VARCHAR2(100 BYTE)
	TELEFONO	NUMERIC(18,0)	COD_PERSONA	GNR_PERSONA	TELEFONO	VARCHAR2(20 BYTE)
	ESTADO	VARCHAR(50)	COD_VENDEDOR	FAC_VENDEDOR	ESTADO	VARCHAR2(1 BYTE)
				EXCEL,INDICADORES	DESCRIPCION	
COMISION	INT	COD_VENDEDOR	FAC_VENDEDOR	SUELDO	NUMBER	

Tabla 14: Mapeo Dimensión Vendedor

Fuente: Propia

DIMENSIÓN FECHA**DESCRIPCIÓN** Almacena todos los datos referentes a la Fecha de factura

DESTINO			CAMPO NAVEGACION	ORIGEN		
NOMBRE TABLA	CAMPO	TIPO DATO		NOMBRE TABLA O VISTA	CAMPO	TIPO DATO
DIM_FECHA	ID_FECHA	NUMERIC(18,0)		V_FECHAS	ID	NUMBER
	FECHA	DATE	FECHA_FACTURA	FACT_FACTURA_C	FECHA_FACTURA	DATE
	AÑO	INT		V_FECHAS	ANIO	VARCHAR2(4 BYTE)
	MES	INT		V_FECHAS	MES	VARCHAR2(2 BYTE)
	DIA	INT		V_FECHAS	DIA	VARCHAR2(2 BYTE)
	MES_LETRAS	VARCHAR(50)		V_FECHAS	MES_LETRAS	VARCHAR2(40 BYTE)

Tabla 15: Mapeo Dimensión Fecha

Fuente Propia

TABLA DE HECHOS

DESCRIPCIÓN Almacena todos los datos numéricos y los ID de las dimensiones

DESTINO			CAMPO NAVEGACION	ORIGEN		
NOMBRE TABLA	CAMPO	TIPO DATO		NOMBRE TABLA O VISTA	CAMPO	TIPO DATO
FAC_VENTAS	ID_CLIENTE	NUMERIC(15,0)	COD_CLIENTE	CXC_CLIENTE	COD_CLIENTE	NUMBER
	ID_ARTICULO	NUMERIC(15,0)	COD_ARTICULO	INV_ARTICULO	COD_ARTICULO	NUMBER(8,0)
	ID_VENDEDOR	NUMERIC(18,0)	COD_VENDEDOR	FAC_VENDEDOR	COD_VENDEDOR	NUMERIC(9,0)
	ID_FECHA	NUMERIC(18,0)	ID	V_FECHAS	ID	NUMBER
	NUMERO_FACTURA	NUMERIC(18,0)	NUM_FACTURA	FAC_FACTURA_C	NUM_FACTURA	NUMBER
	PRECIO_UNITARIO	NUMERIC(18,2)		CALCULADO		
	TOTAL	NUMERIC(18,2)		CALCULADO		
	CANTIDAD	NUMERIC(18,0)	CANTIDAD	FACT_FACTURA_D	CANTIDAD	NUMBER

Tabla 16: Mapeo Tabla de Hechos

Fuente Propia

3.2.3 DEFINICIÓN DE DIMENSIONES

Como se ha mencionado anteriormente en este proyecto, las dimensiones contienen la información resumida o consolidada de un tema específico, por lo tanto en base a la definición de requerimientos y el análisis realizado hasta este punto se ha podido identificar las siguientes dimensiones que organizan la información necesaria:

Dimensión Cliente: En esta dimensión se extrae de varias tablas de origen la información básica de los clientes.

DIM_CLIENTE
ID_CLIENTE
APELLIDOS
NOMBRES
RAZON_SOCIAL
DIRECCION
TELEFONO
CIUDAD_ENTREGA
CREDITO
ESTADO
LIMITE_FACTURA
CLASE_CLIENTE
ZONA

Tabla 17: Dimensión Clientes

Fuente: Propia

Dimensión Artículos: en esta dimensión se registra la información que describe y consolida las principales características de los productos.

DIM_ARTICULOS
ID_ARTICULO
LINEA
GRUPO
SUBGRUPO
NOMBRE
UNIDAD
CODIGO_FABICA
PROVEEDOR
FECHA_CREACION

PESO
BODEGA

Tabla 18: Dimensión Artículo

Fuente: Propia

Dimensión Vendedor: En esta dimensión se detalla la información general de los vendedores que contactan a los clientes.

DIM_VENDEDOR
ID_VENDEDOR
RAZON_SOCIAL
TELEFONO
ESTADO
COMISION

Tabla 19: Dimensión Vendedor

Fuente: Propia

Dimensión Fecha: En esta dimensión se detalla el tiempo en año, mes y día en que se analizan cada dimensión.

DIM_FECHA
ID_FECHA
FECHA
AÑO
MES
DIA
MES_LETRA

Tabla 20: Dimensión Fecha

Fuente: Propia

3.2.4 IDENTIFICACIÓN DE TABLA DE HECHOS Y SUS MEDIDAS

Para el diseño del presente proyecto se ha usado el esquema estrella que es el que más se ajusta a las necesidades, por lo cual en la tabla de hechos se refleja las claves primarias de las dimensiones a las cuales se puede navegar para analizar la información, además consta también de varias medidas que resuelven la necesidad de información logrando de esta manera simplicidad en la

comprensión de la información y velocidad en el acceso a la misma. Es decir en esta tabla se registran todos los datos numéricos que se va a analizar o calcular.

FAC_VENTAS
ID_CLIENTE
ID_ARTICULO
ID_VENDEDOR
ID_FECHA
NUMERO_FACTURA
PRECIO_UNITARIO
TOTAL
CANTIDAD
COSTO_PROMEDIO
UTILIDAD

Tabla 21: Tabla de hechos

Fuente: Propia

3.2.5 MODELO DE DATOS

Una vez definidas e identificadas las tablas de dimensiones con su respectivos atributos y la tabla de hechos con sus medidas, se pasa a la construcción del modelo de datos en donde se define como está estructurada la tabla de hechos con las dimensiones.

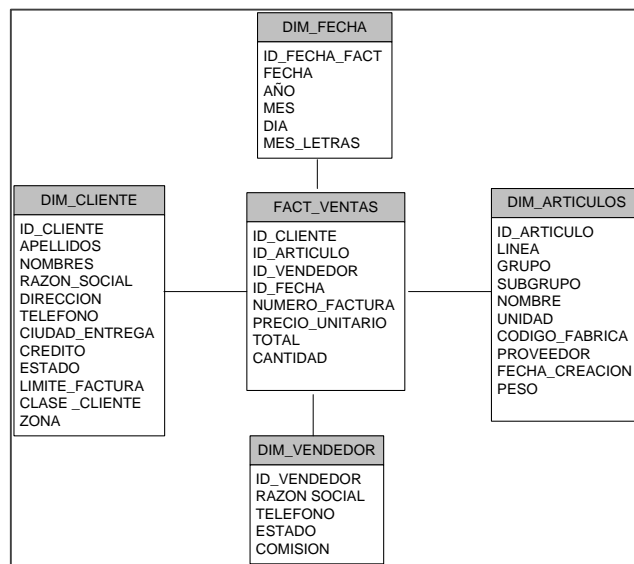


Ilustración 31: Modelo de Datos

Fuente: Propia

3.3 CONSTRUCCIÓN Y DESARROLLO

En esta fase de construcción se diseña todos los procesos necesarios para obtener los reportes finales, empezando con la construcción de la base de datos multidimensional a partir de los datos de origen analizados anteriormente, así como también los procesos ETL mediante el SSIS para poblar el repositorio de datos destino. Y finalmente se construye el cubo mediante el SSAS del cual ya se puede tomar la información necesaria para los reportes requeridos.

Para iniciar la construcción del presente proyecto se desarrolla las siguientes actividades:

3.3.1 CONEXIÓN A BASE DE DATOS FUENTE

Para iniciar con el desarrollo del proyecto lo primero que se necesita es una copia de la base de datos transaccional la cual se la ha cargado en el motor de base de datos ORACLE, posteriormente se procede a realizar la conexión para poder referenciarla.

Para esto es necesario instalar el componente Oracle Data Acces Components, que se encuentra disponible en la página de ORACLE www.oracle.com, el cual permite instalar los archivos necesarios para la conexión desde SQL Server Data Tools al motor de base de datos ORACLE.

Su instalación es sencilla en donde se acepta todos los pasos propuestos:



Ilustración 32: Instalación de ORACLE DATA ACCES COMPONENT

Fuente: Propia

Una vez instalado el ODAC ingresar a C:\odac\network\admin en donde se encuentra el archivo **tnsnames** en el cual se crea la cadena de conexión para acceder a la base de datos de ORACLE.

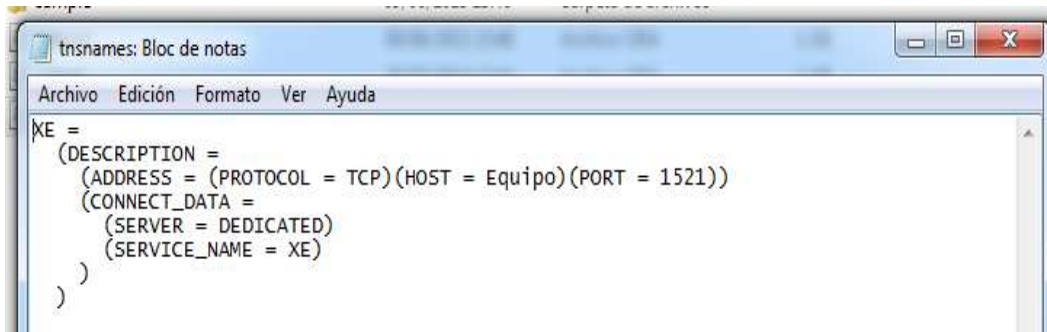


Ilustración 33: Archivo **tnsnames** de conexión a ORACLE
Fuente: Propia

3.3.2 INSTALACIÓN DE SQL SERVER 2014

Los instaladores de SQL Server 2014 y Visual Studio 2013, en su versión Enterprise Edition x 64 bits, para fines académicos del presente proyecto se descargaron del sitio oficial de Microsoft proveedor del producto, su instalación de detalla a continuación:

1. Una vez descargado el paquete de instalación ejecutar setup.exe

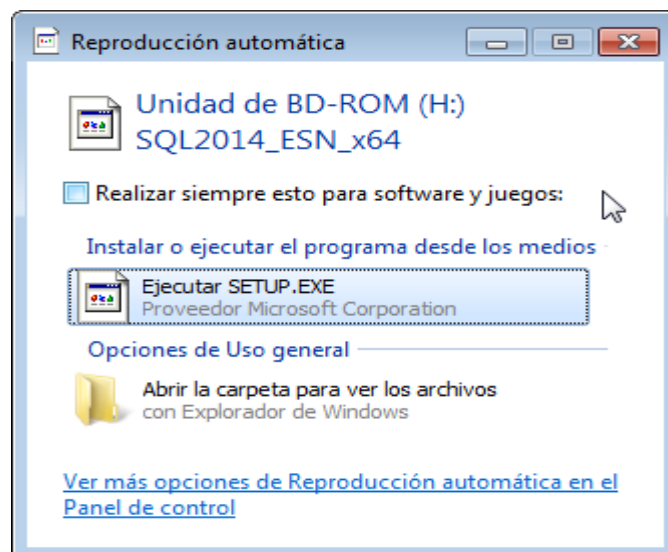


Ilustración 34: Paso 1 Instalación SQL Server 2014
Fuente Propia

2. Hacer clic en el menú Instalación y escoger Nueva instalación independiente de SQL Server o agregar características de una instalación existente.

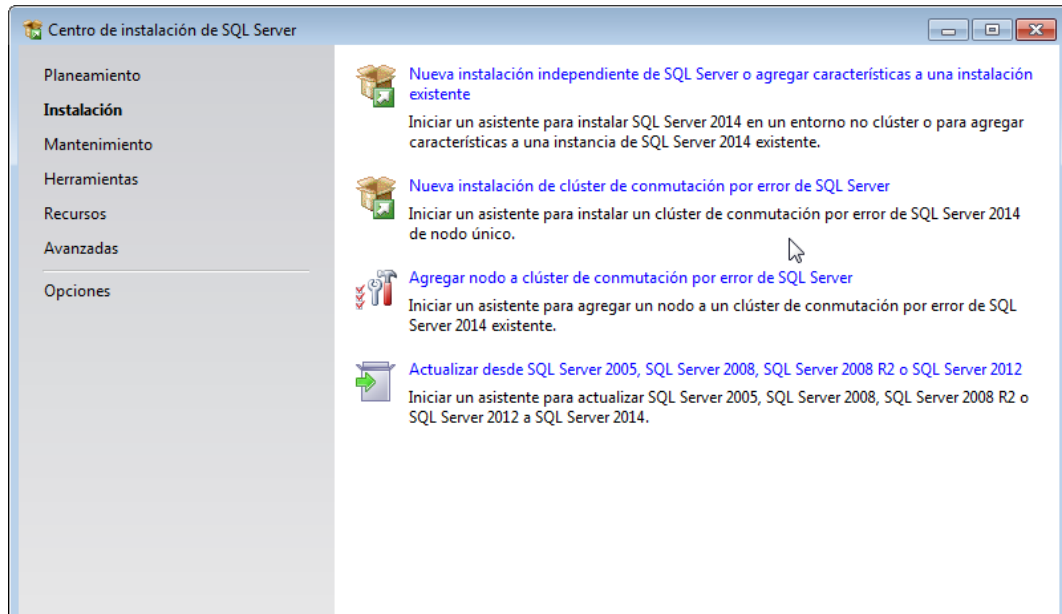


Ilustración 35: Paso 2 Instalación SQL Server 2014
Fuente Propia

3. Ingresar la clave del producto y hacer clic en siguiente

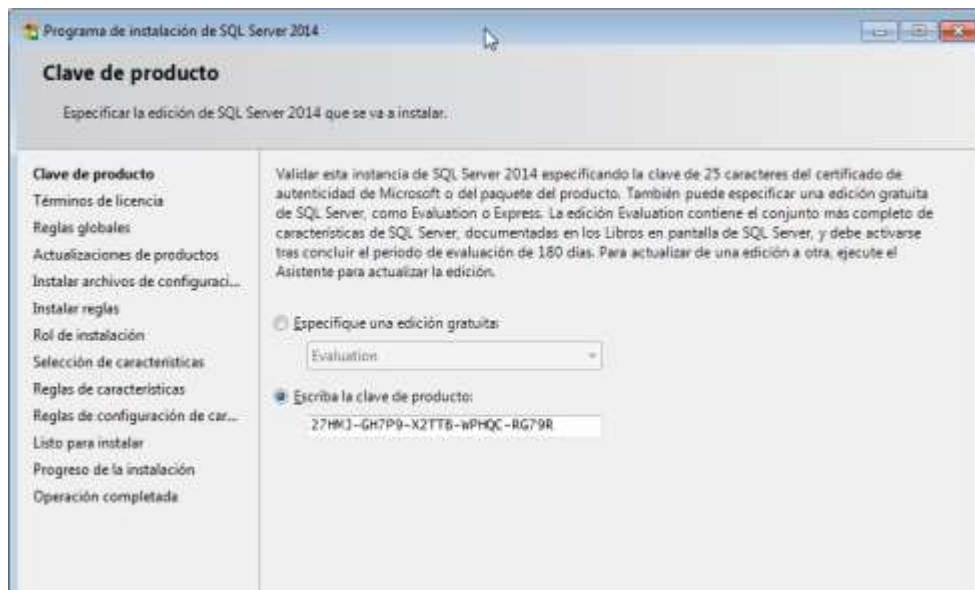


Ilustración 36: Paso 3 Instalación SQL Server 2014
Fuente: Propia

4. Aceptar los términos de licencia y hacer clic en siguiente

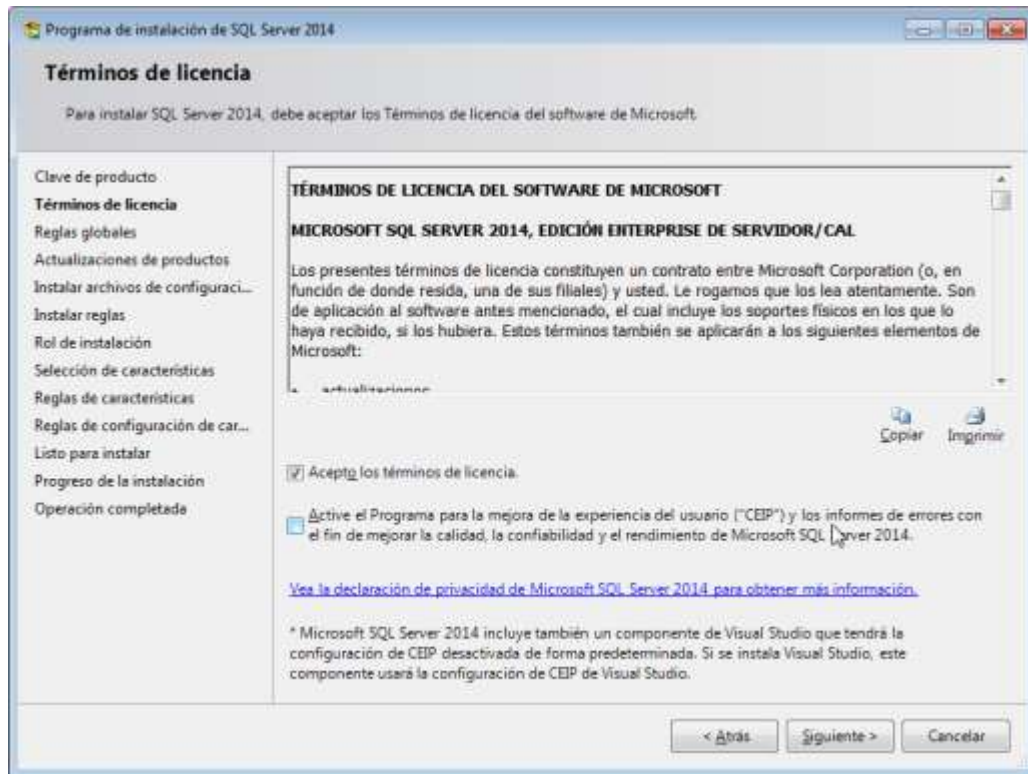


Ilustración 37: Paso 4 Instalación SQL Server 2014

Fuente: Propia

5. Comprobación de reglas globales para la instalación. Hacer clic en siguiente.

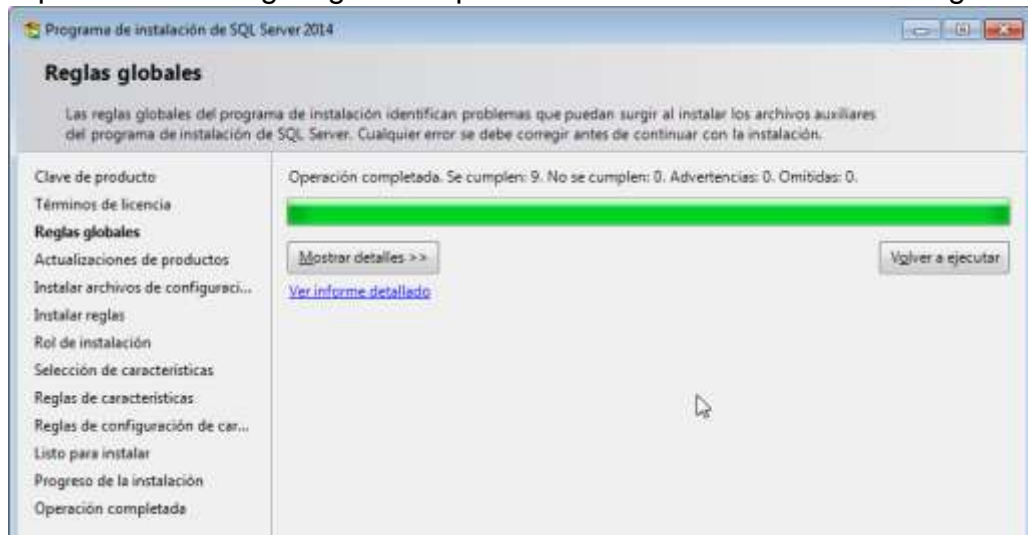


Ilustración 38: Paso 5 Instalación SQL Server 2014

Fuente: Propia

6. Instalación de archivos de configuración. Hacer clic en siguiente

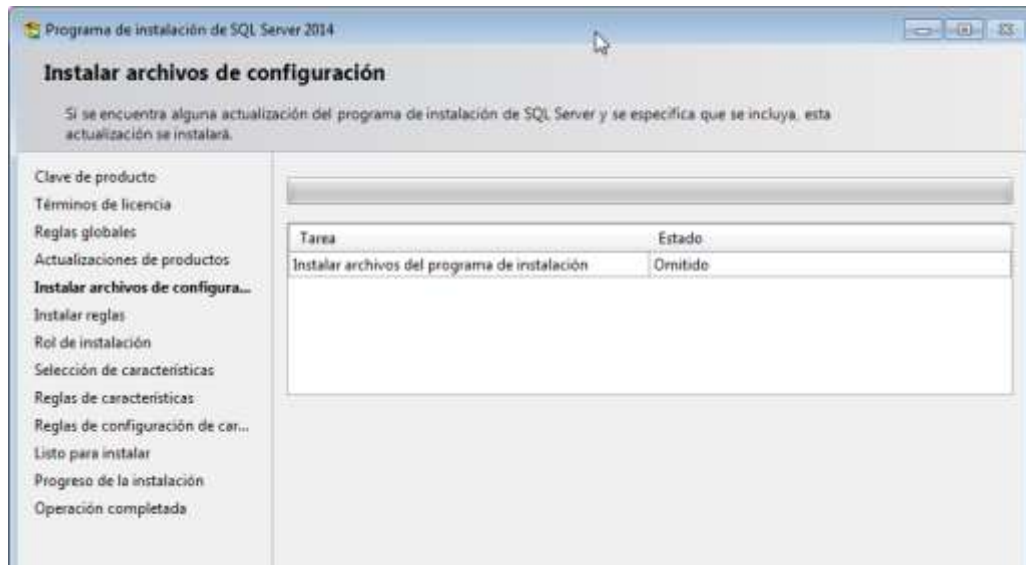


Ilustración 39: Paso 6 Instalación SQL Server 2014

Fuente: Propia

7. Instalar reglas. Hacer clic en siguiente

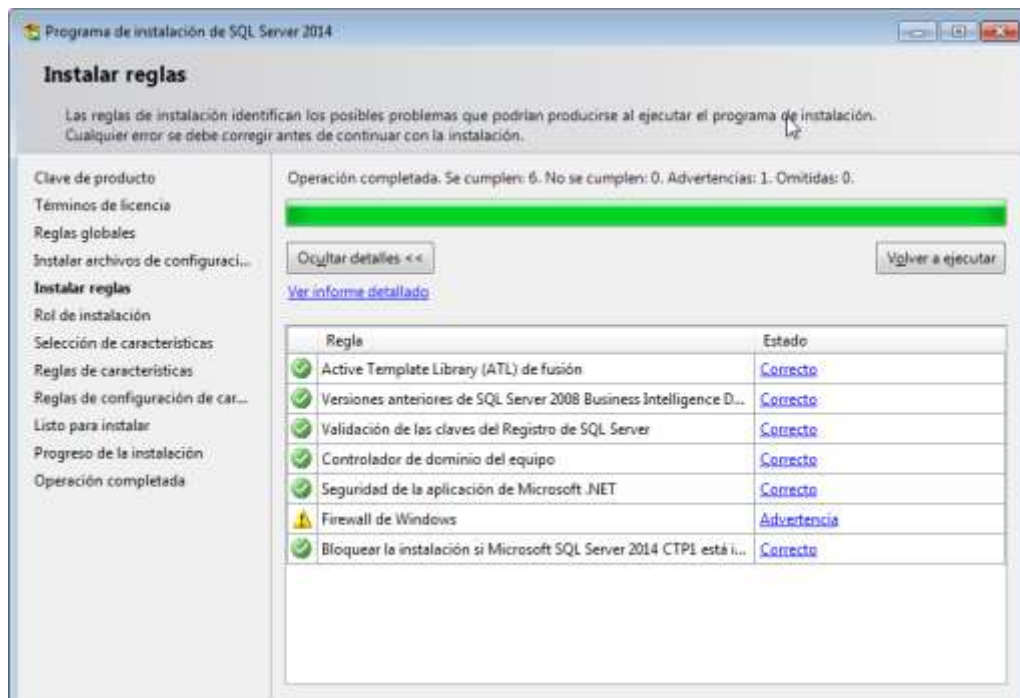


Ilustración 40: Paso 7 Instalación SQL Server 2014

Fuente: Propia

8. Rol de la instalación. Escoger Instalación de características de SQL Server. Hacer clic en siguiente.

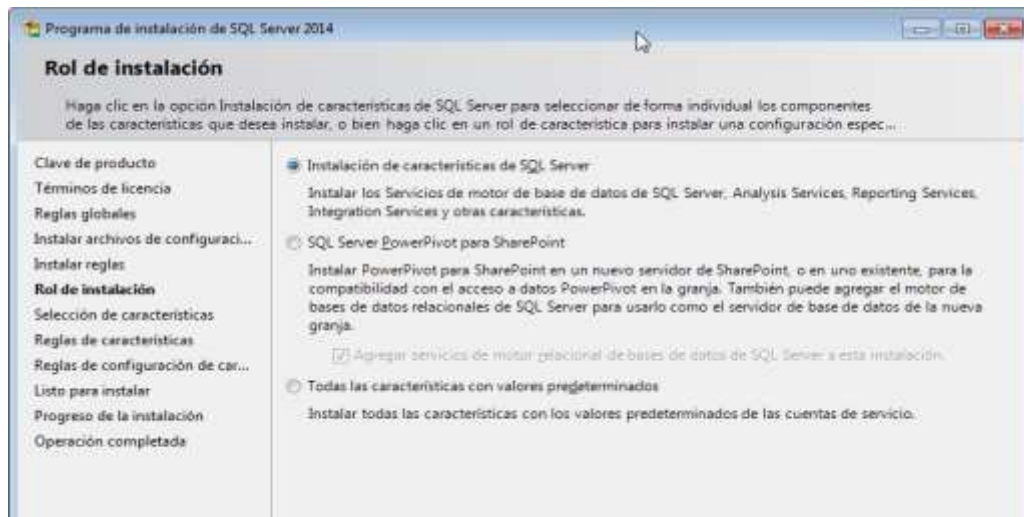


Ilustración 41: Paso 8 Instalación SQL Server 2014
Fuente: Propia

9. Selección de características del Servidor. Escoger todas las características. Hacer clic en siguiente

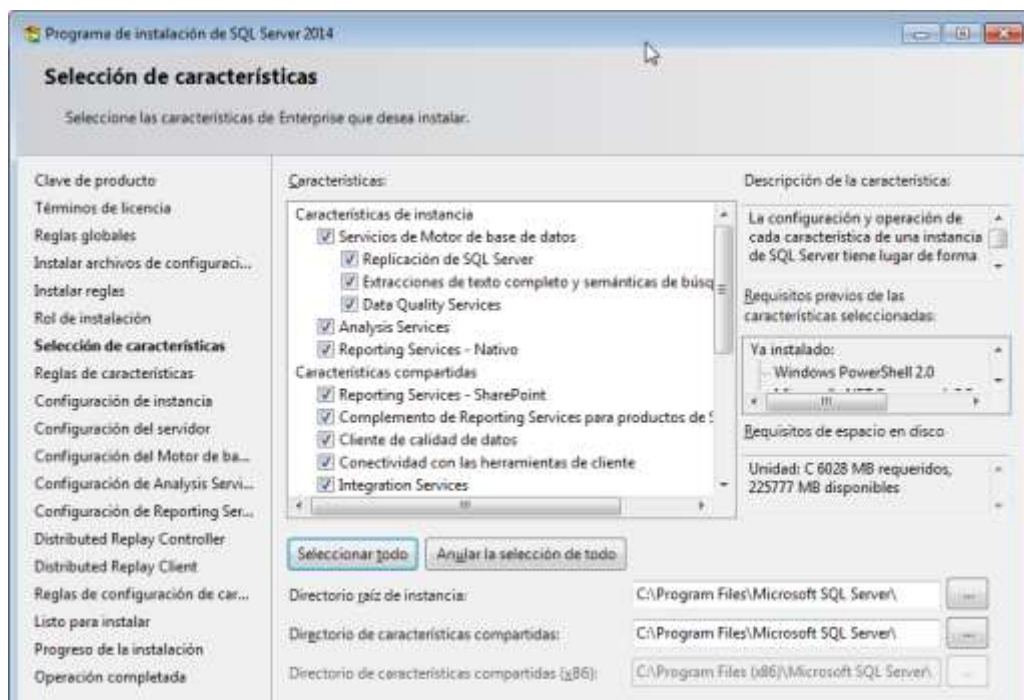


Ilustración 42: Paso 9 Instalación SQL Server 2014
Fuente: Propia

10. Configuración de la Instancia. Dejar por defecto los valores recomendados. Hacer clic en siguiente

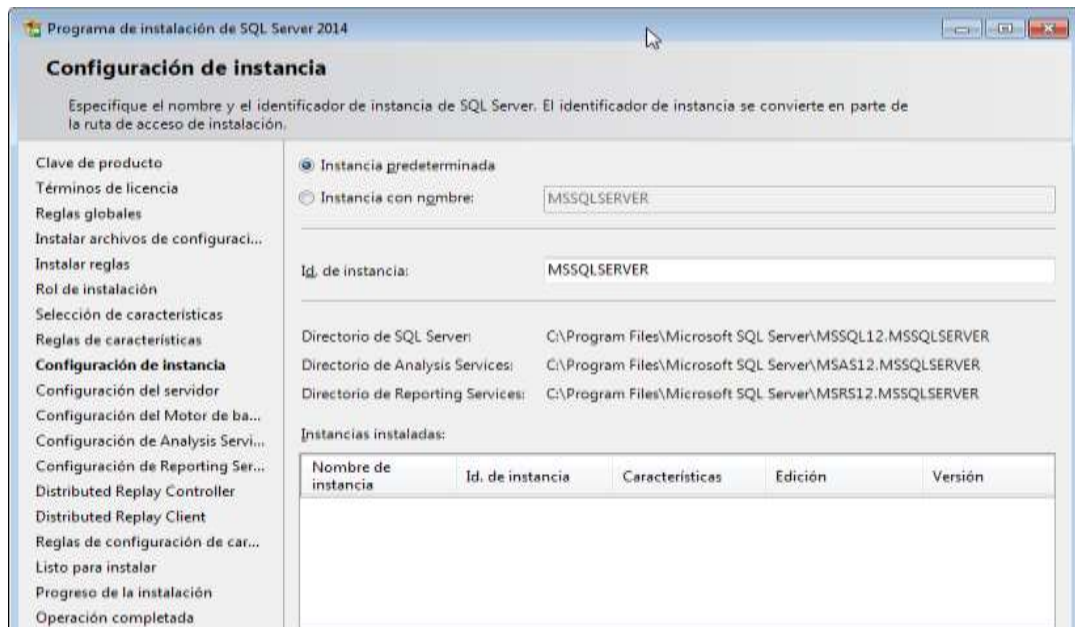


Ilustración 43: Paso 10 Instalación SQL Server 2014

Fuente: Propia

11. Configuración del Servidor y cuentas de servicio. Dejar los valores por defecto verificando que el Agente SQL Server tenga inicio automático y hacer clic en siguiente.

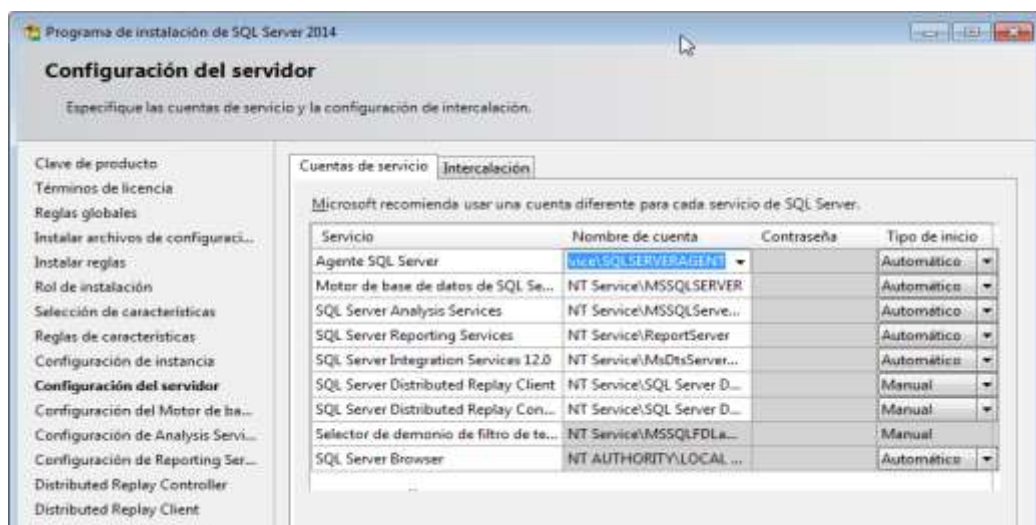


Ilustración 44: Paso 11 Instalación SQL Server 2014

Fuente: Propia

12. Configurar el motor de la base de datos, en el modo de autenticación, estos valores pueden ser cambiados después de la instalación. Agregar los usuarios necesarios. Hacer clic en siguiente.

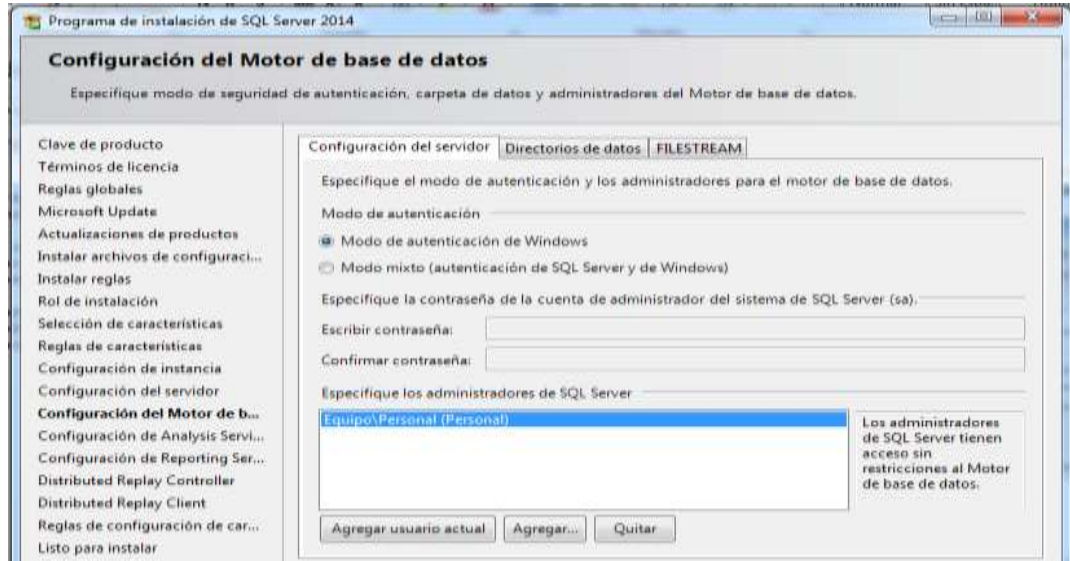


Ilustración 45: Paso 12 Instalación SQL Server 2014

Fuente: Propia

13. Configuración de Analysis Services en Modo Multidimensional y de minería de datos. Agregar los usuarios y hacer clic en siguiente.

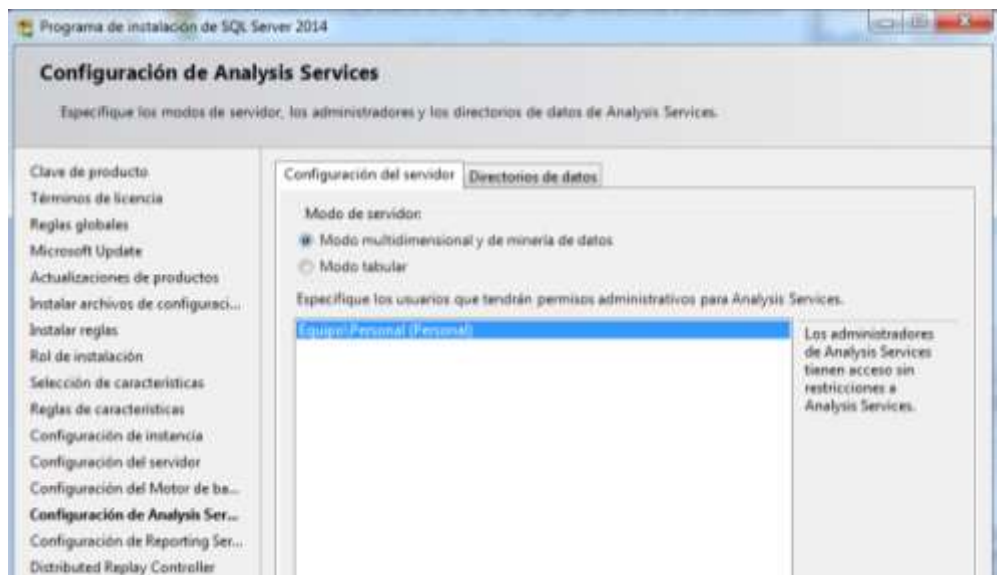


Ilustración 46: Paso 11 Instalación SQL Server 2014

Fuente: Propia

14. Instalación de Reporting Services. Seleccionar Instalar y configurar. Hacer clic en siguiente.



Ilustración 47: Paso 14 Instalación SQL Server 2014

Fuente: Propia

15. Especificar los usuarios con permiso de acceso al servicio Distributed Replay Controller. Agregar los usuarios y hacer clic en siguiente

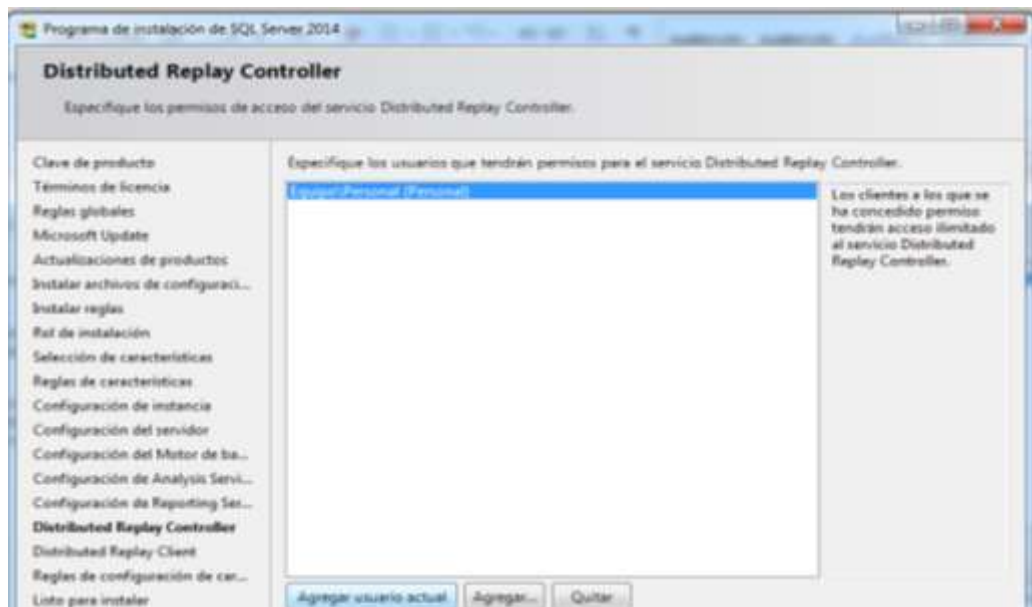


Ilustración 48: Paso 15 Instalación SQL Server 2014

Fuente: Propia

16. Especificar el controlador para Distributed Replay Client. Dejar los valores por defecto. Hacer clic en siguiente

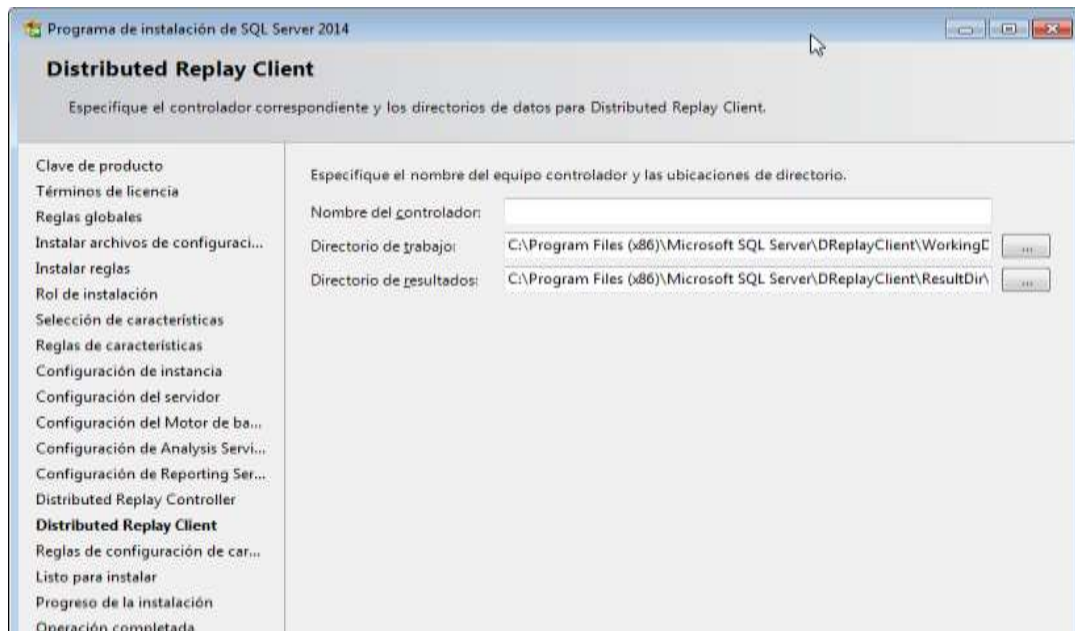


Ilustración 49: Paso 16 Instalación SQL Server 2014

Fuente: Propia

17. Comprobar las características seleccionadas para instalar. Hacer clic en Instalar.

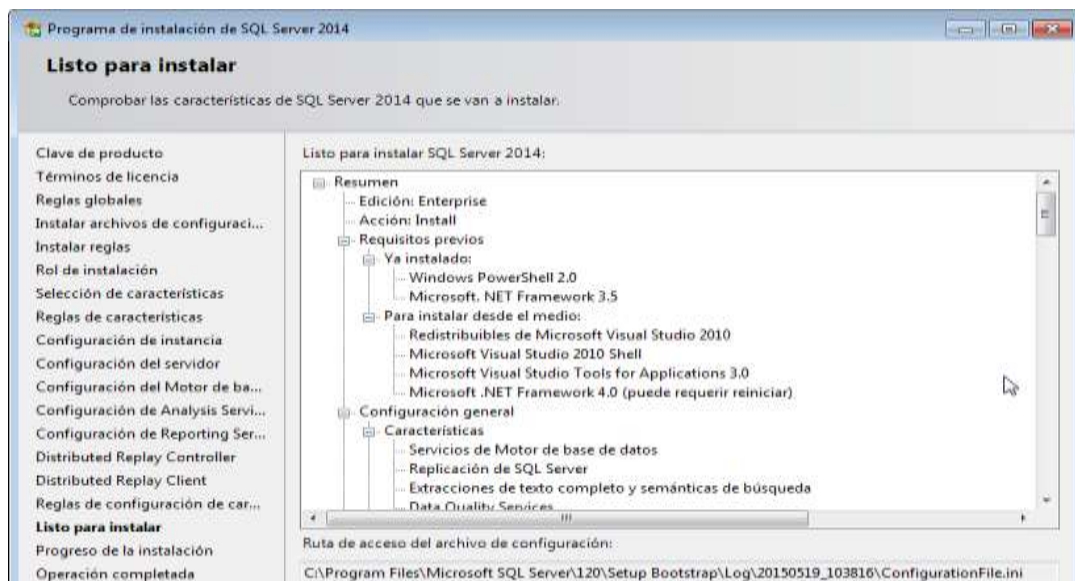


Ilustración 50: Paso 17 Instalación SQL Server 2014

Fuente: propia

18. Después del progreso de la Instalación la operación se completa correctamente.

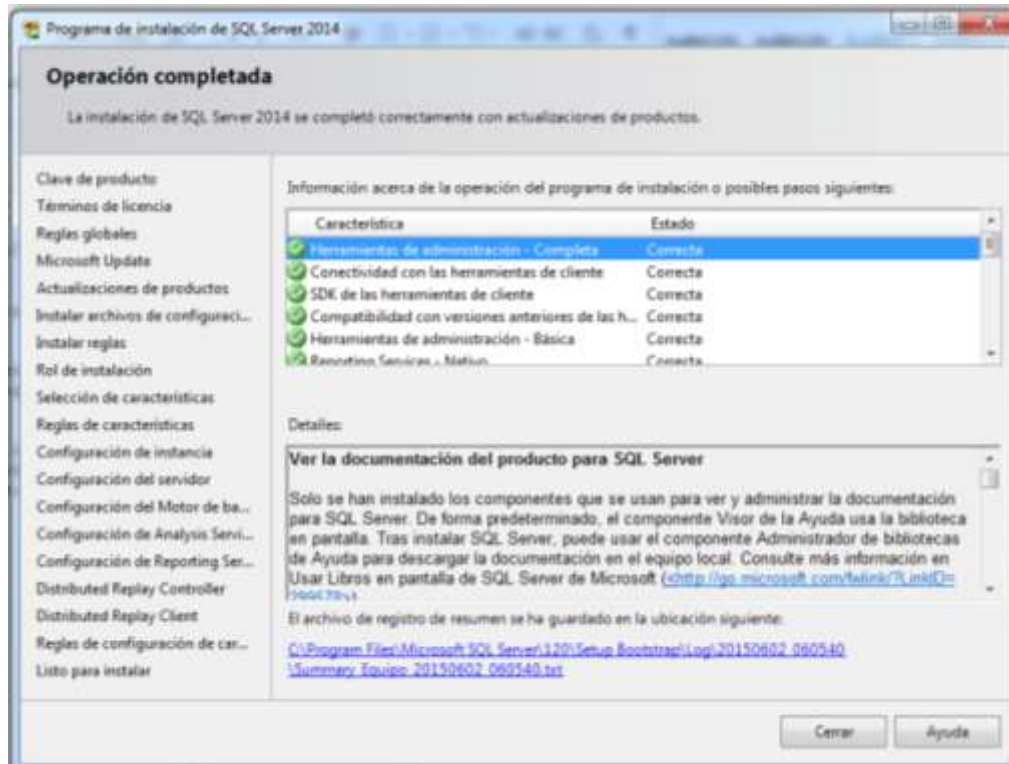


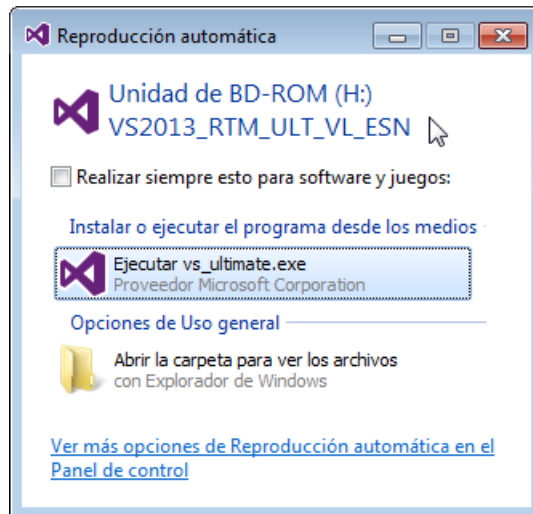
Ilustración 51: Paso 18 Instalación SQL Server 2014

Fuente: Propia

3.3.3 INSTALACIÓN DE VISUAL STUDIO 2013

De igual manera el Visual Studio 2013 que soporta el componente SQL Server Data Tools que permite la implementación de los proyectos tanto de SSIS (SQL Server Integration Services) como SSAS (SQL Server Analysis Services) , está disponible en la página de Microsoft.com.

1. Iniciar la instalación haciendo clic en vs_ultimate.exe



*Ilustración 52: Paso 1 Instalación Visual Studio 2013
Fuente: Propia*

2. Aceptar los términos de licencia. Hacer clic en siguiente.



*Ilustración 53: Paso 2 Instalación Visual Studio 2013
Fuente: Propia*

3. Actualizar Visual Studio 2013. Ejecutar el archivo VS2013.1

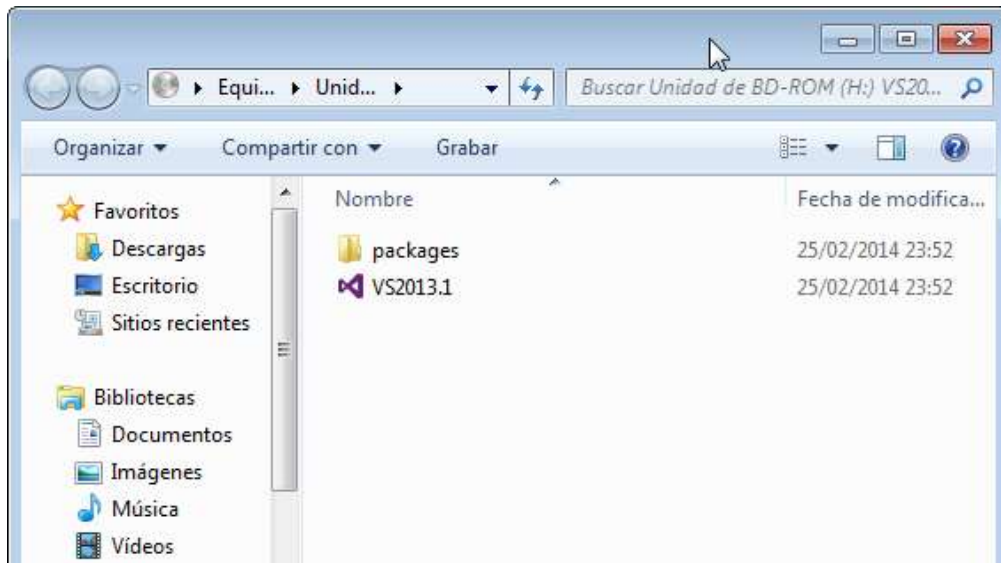


Ilustración 54: Paso 3 Instalación Visual Studio 2013

Fuente: Propia

4. Aceptar términos de licencia y hacer clic en instalar



Ilustración 55: Paso 4 Instalación Visual Studio 2013

Fuente: Propia

5. Cuando se haya instalado los componentes de actualización, hacer clic en Iniciar.

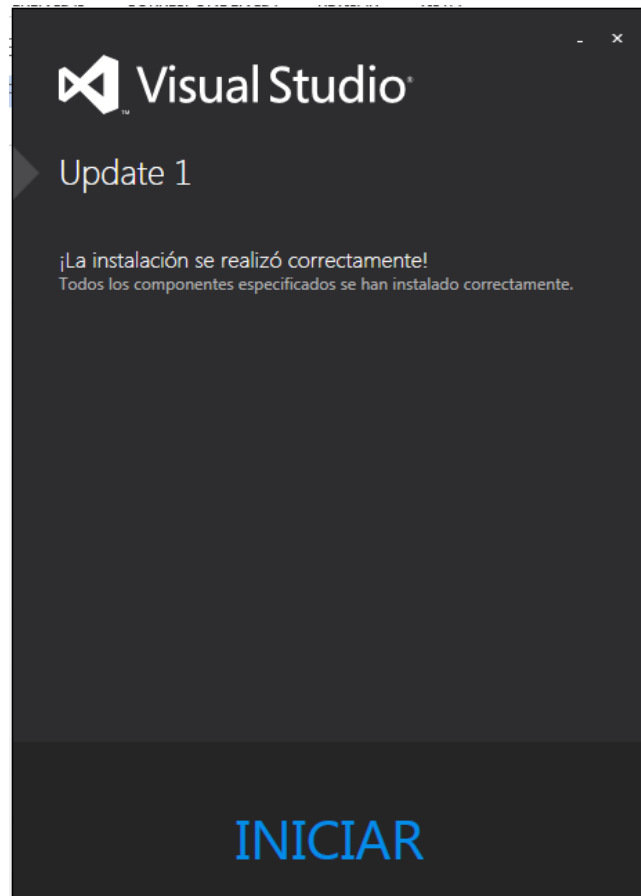
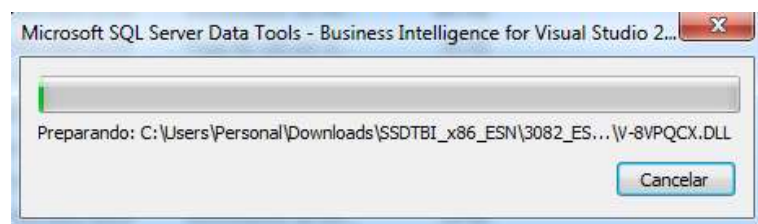


Ilustración 56: Paso 5 Instalación Visual Studio 2013
Fuente: Propia

Progreso de la instalación:



6. Aceptar los términos de licencia y clic en siguiente

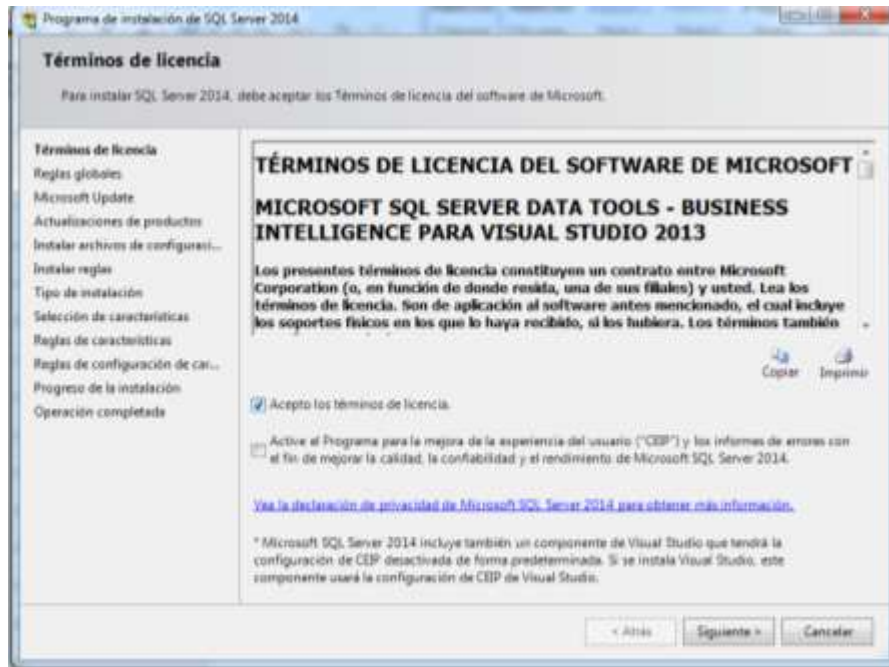


Ilustración 57: Paso 5 Instalación Visual Studio 2013

Fuente: Propia

7. Aceptar las reglas globales y en actualizaciones dejar por defecto y clic en siguiente:



Ilustración 58: Paso 7 Instalación Visual Studio 2013

Fuente: Propia

8. Se instalan los archivos de configuración y clic en siguiente

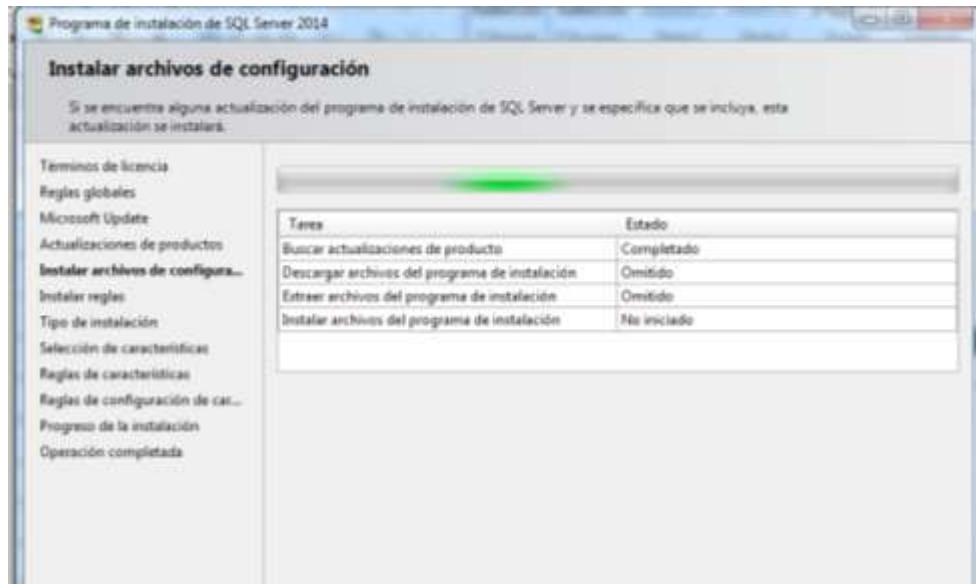


Ilustración 59: Paso 8 Instalación Visual Studio 2013

Fuente: Propia

9. En el tipo de instalación seleccionar la opción agregar características a una instancia existente de SQL, ya que en este caso ya tener instalado el SQL Server.

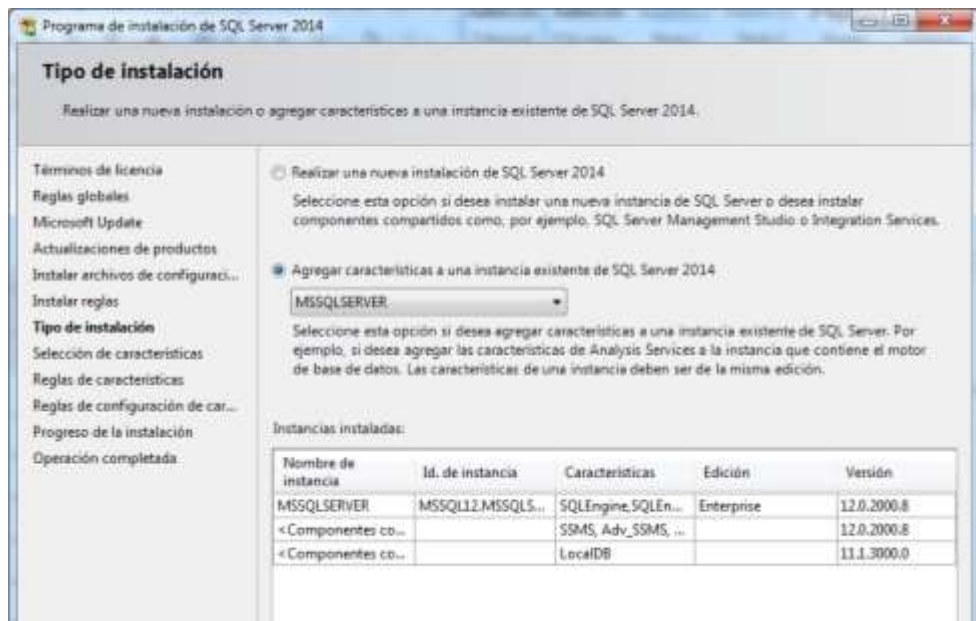


Ilustración 60: Paso 9 Instalación Visual Studio 2013

Fuente: Propia

10. En esta página selección de características, haga clic en SQL Server Data Tools – Business Intelligence para Visual Studio 2012/2013 y después clic en siguiente en esta ventana y las 2 siguientes.

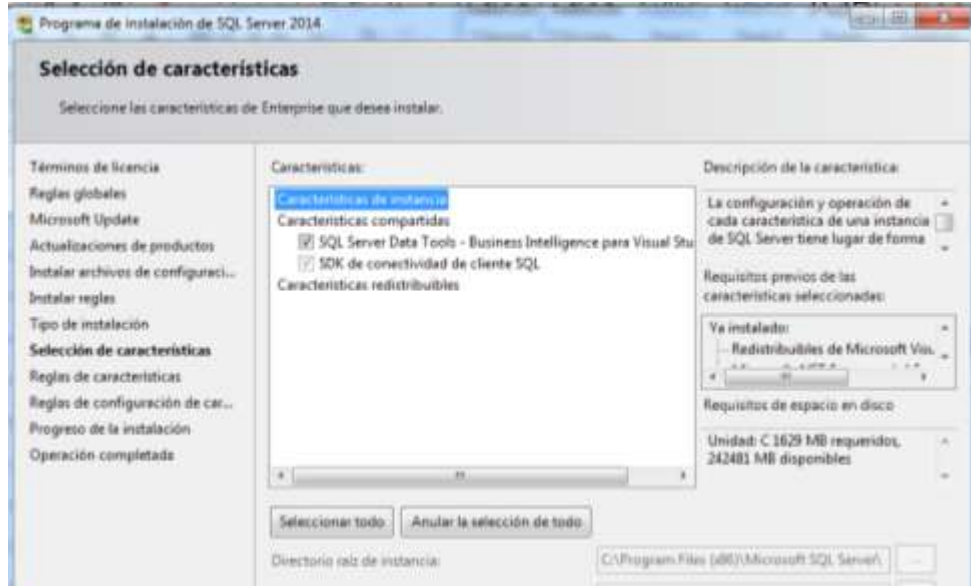


Ilustración 61: Paso 10 Instalación Visual Studio 2013
Fuente: Propia

11. Esperar a que el progreso de instalación se complete

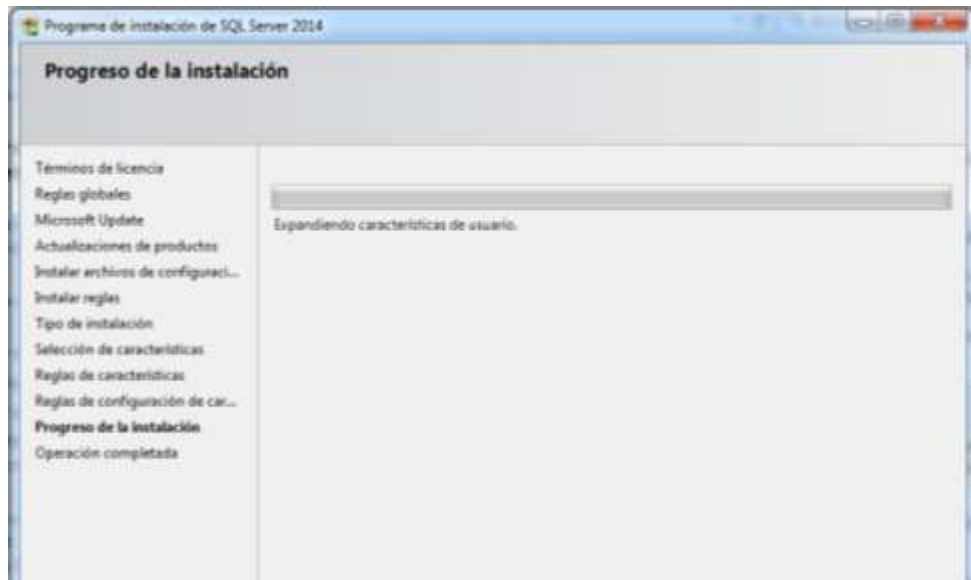


Ilustración 62: Paso 11 Instalación Visual Studio 2013
Fuente: Propia

12. Finalmente se instala y se adjunta al menú de Microsoft SQL Server 2014.

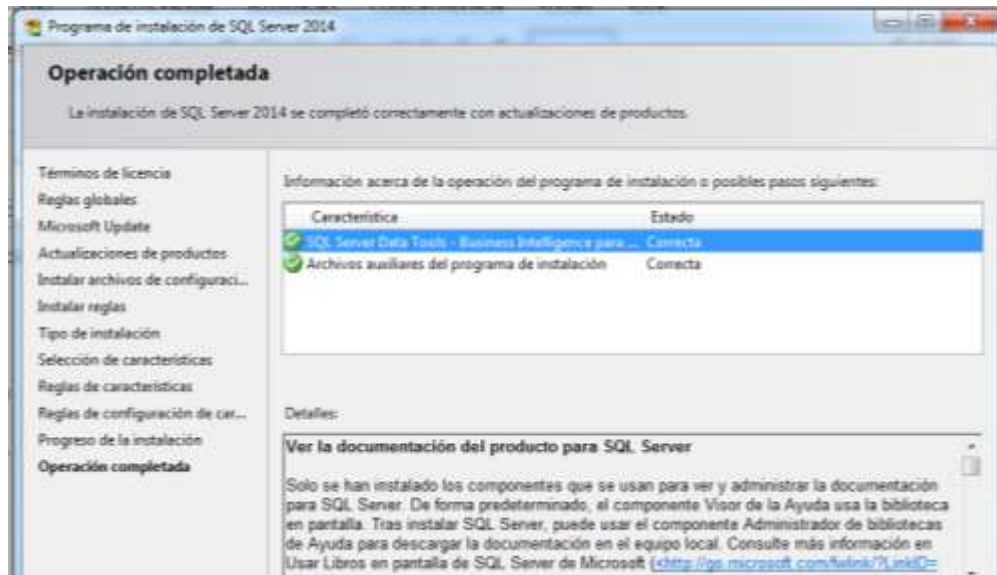


Ilustración 63: Paso 12 Instalación Visual Studio 2013

Fuente: Propia

3.3.4 CREACIÓN DE BASE DE DATOS MULTIDIMENSIONAL

De acuerdo al análisis realizado anteriormente, se diseña la base de datos con las tablas del repositorio que se poblara con la información fuente transformada, para el diseño de la base de datos multidimensional se usa el motor de base de datos SQL Server Management Studio. A continuación el detalle de creación:

Creación de la base de datos:

```
CREATE DATABASE [DataMartVentasDP]
  CONTAINMENT = NONE
  ON PRIMARY
  ( NAME = N'DataMartVentasD',
  FILENAME = N'C:\Program Files\Microsoft SQL Server\MSSQL12.MSSQLSERVER\MSSQL\DATA\DataMartVentasD.mdf' ,
  SIZE = 162816KB , MAXSIZE = UNLIMITED, FILEGROWTH = 1024KB )
  LOG ON
  ( NAME = N'DataMartVentasD_log',
  FILENAME = N'C:\Program Files\Microsoft SQL Server\MSSQL12.MSSQLSERVER\MSSQL\DATA\DataMartVentasD_log.ldf' ,
  SIZE = 568896KB , MAXSIZE = 2048GB , FILEGROWTH = 10%)
```

Ilustración 64: Código Creación Base de datos multidimensional

Fuente: Propia

Creación de la tabla Artículos

```
CREATE TABLE [dbo].[DIM_ARTICULOS](
    [ID_ARTICULO] [numeric](15, 0) NOT NULL,
    [LINEA] [varchar](50) NULL,
    [GRUPO] [varchar](50) NULL,
    [SUBGRUPO] [varchar](50) NULL,
    [NOMBRE] [varchar](100) NULL,
    [UNIDAD] [varchar](50) NULL,
    [CODIGO_FABRICA] [varchar](50) NULL,
    [PROVEEDOR] [varchar](100) NULL,
    [FECHA_CREACION] [date] NULL,
    [PESO] [numeric](18, 2) NULL,
    CONSTRAINT [PK_DIM_ARTICULOS] PRIMARY KEY CLUSTERED
(
    [ID_ARTICULO] ASC
)
WITH (PAD_INDEX = OFF, STATISTICS_NORECOMPUTE = OFF, IGNORE_DUP_KEY = OFF,
    ALLOW_ROW_LOCKS = ON, ALLOW_PAGE_LOCKS = ON)
ON [PRIMARY] ) ON [PRIMARY]
```

Ilustración 65: Código Creación Tabla DIM_ARTICULOS
Fuente: Propia

Creación de la tabla Cliente

```
CREATE TABLE [dbo].[DIM_CLIENTE](
    [ID_CLIENTE] [numeric](15, 0) NOT NULL,
    [RAZON_SOCIAL] [varchar](100) NULL,
    [DIRECCION] [varchar](100) NULL,
    [TELEFONO] [varchar](50) NULL,
    [CIUDADE_ENTREGA] [varchar](100) NULL,
    [CREDITO] [varchar](10) NULL,
    [ESTADO] [varchar](10) NULL,
    [LIMITE_FACTURA] [numeric](18, 2) NULL,
    [CLASE_CLIENTE] [varchar](50) NULL,
    [ZONA] [varchar](150) NULL,
    CONSTRAINT [PK_DIM_CLIENTE] PRIMARY KEY CLUSTERED
(
    [ID_CLIENTE] ASC
)
WITH (PAD_INDEX = OFF, STATISTICS_NORECOMPUTE = OFF, IGNORE_DUP_KEY = OFF,
    ALLOW_ROW_LOCKS = ON, ALLOW_PAGE_LOCKS = ON) ON [PRIMARY]
ON [PRIMARY]
```

Ilustración 66: Código Creación Tabla DIM_CLIENTES
Fuente: Propia

Creación de la tabla Fecha

```
CREATE TABLE [dbo].[DIM_FECHA](
    [ID_FECHA] [numeric](18, 0) NOT NULL,
    [FECHA] [date] NULL,
    [AÑO] [int] NULL,
    [MES] [int] NULL,
    [DIA] [int] NULL,
    [MES_LETRAS] [varchar](50) NULL,
    CONSTRAINT [PK_DIM_FECHA_FACTURA] PRIMARY KEY CLUSTERED
    (
        [ID_FECHA] ASC
    )
) WITH (PAD_INDEX = OFF, STATISTICS_NORECOMPUTE = OFF, IGNORE_DUP_KEY = OFF,
    ALLOW_ROW_LOCKS = ON, ALLOW_PAGE_LOCKS = ON) ON [PRIMARY]
) ON [PRIMARY]
```

Ilustración 67: Código Creación Tabla DIM_FECHA

Fuente: Propia

Creación de la tabla Vendedor

```
CREATE TABLE [dbo].[DIM_VENDEDOR](
    [ID_VENDEDOR] [numeric](18, 0) NOT NULL,
    [RAZON_SOCIAL] [varchar](100) NULL,
    [DIRECCION] [varchar](100) NULL,
    [TELEFONO] [varchar](50) NULL,
    [ESTADO] [varchar](50) NULL,
    [COMISION] [numeric](18, 2) NULL,
    CONSTRAINT [PK_DIM_VENDEDOR] PRIMARY KEY CLUSTERED
    (
        [ID_VENDEDOR] ASC
    )
) WITH (PAD_INDEX = OFF, STATISTICS_NORECOMPUTE = OFF, IGNORE_DUP_KEY = OFF,
    ALLOW_ROW_LOCKS = ON, ALLOW_PAGE_LOCKS = ON) ON [PRIMARY]
) ON [PRIMARY]
```

Ilustración 68: Código Creación Tabla DIM_VENDEDOR

Fuente Propia

Creación de la tabla de hechos Fact_Ventas

```
CREATE TABLE [dbo].[FACT_VENTAS](
    [ID_CLIENTE] [numeric](15, 0) NULL,
    [ID_ARTICULO] [numeric](15, 0) NOT NULL,
    [ID_VENDEDOR] [numeric](18, 0) NOT NULL,
    [ID_FECHA] [numeric](18, 0) NULL,
    [NUMERO_FACTURA] [numeric](18, 0) NOT NULL,
    [PRECIO_UNITARIO] [numeric](18, 2) NULL,
    [TOTAL] [numeric](18, 2) NOT NULL,
    [CANTIDAD] [numeric](18, 0) NULL,
    [COSTO] [numeric](18, 2) NULL,
    [COSTO_TOTAL] [numeric](18, 2) NULL
) ON [PRIMARY]
```

Ilustración 69: Código Creación Tabla FACT_VENTAS

Fuente: Propia

Finalmente el diagrama de base de datos queda de la siguiente manera:

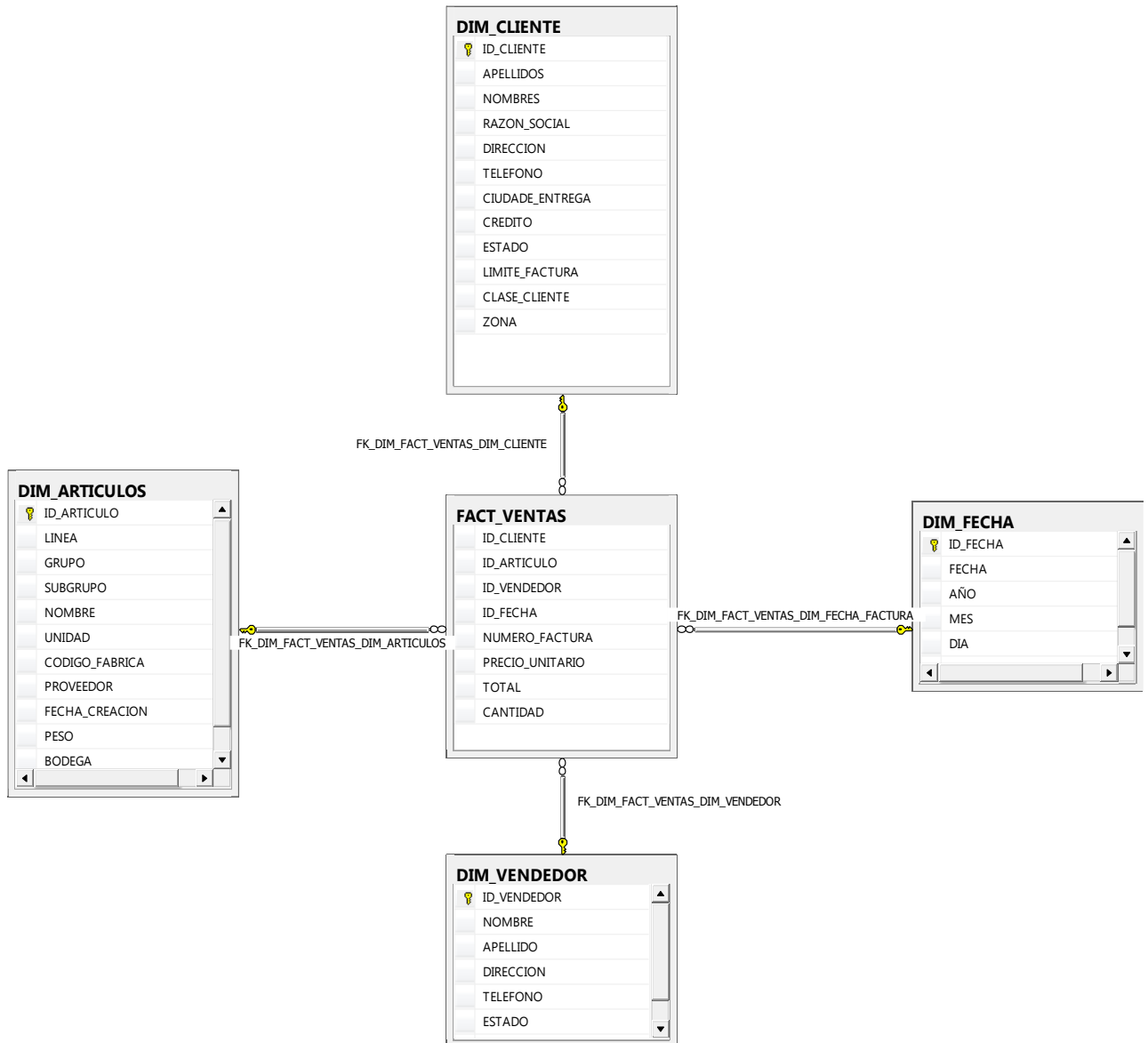


Ilustración 70: Diagrama de Base de Datos Multidimensional

Fuente: Propia

3.3.5 CONSTRUCCIÓN DE PROCESOS ETL

Como se ha especificado anteriormente, para la construcción de los procesos ETL se utiliza la herramienta de SQL Server Data Tools, mediante la creación de un proyecto de Integration Services, a continuación se describe el procedimiento realizado:

En el proyecto de Integration Services, lo primero es definir los orígenes de datos, que en este caso serán: el origen de la base de datos fuente creado anteriormente, el origen de la base de datos multidimensional y también se hace referencia a un origen de Excel.

Para lo cual se da clic derecho en Administrador de conexiones, nuevo origen de datos, usando el asistente seleccionar tipo de administrador OLEDB y clic en agregar:

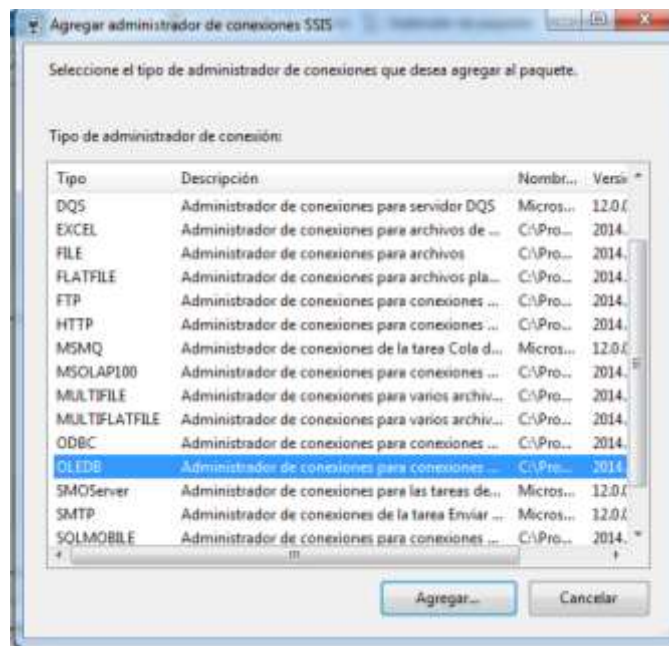


Ilustración 71: Selección del tipo de administrador de conexión
Fuente: Propia

En configurar el administrador de conexiones clic en nueva para ingresar los datos de referencia de la base de datos fuente con proveedor OLE DB Provider for Oracle, aquí se especifica en nombre de Servidor especificado en la configuración de conexión a Oracle explicado anteriormente, y especificar los datos para conexión con la base de datos fuente:

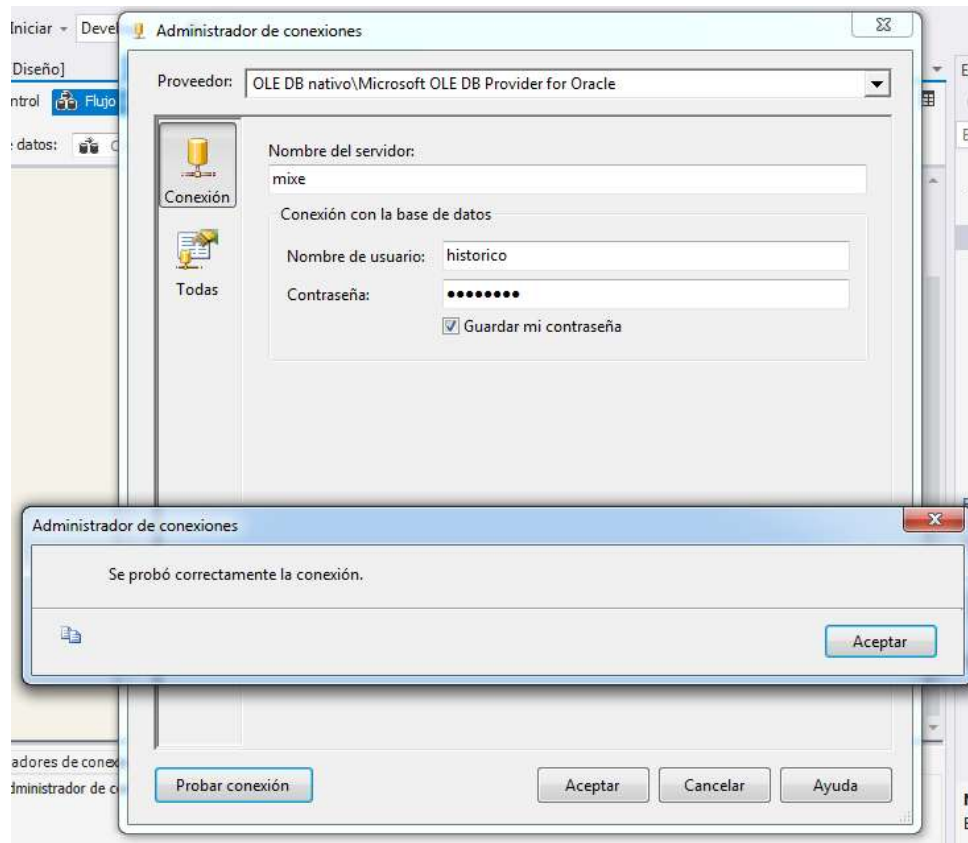


Ilustración 72: Conexión a base de datos fuente en Oracle
Fuente: Propia

De igual manera se crea el origen de la base de datos multidimensional en la que se va a depositar la información consolidada, con proveedor OLE DB Nativo de Microsoft, haciendo referencia al nombre de la base de datos DataMartVentasDP:

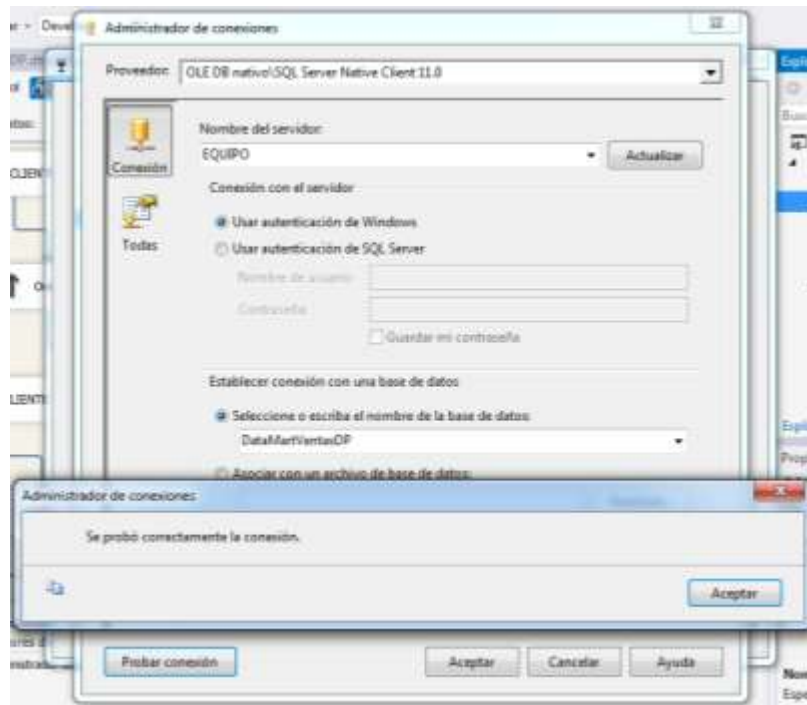


Ilustración 73: Conexión a base de Datos SQL destino
Fuente: Propia

Para la creación del presente proceso ETL también se necesita crear un origen para Excel para lo cual se procede a referenciar con administrador de conexiones de la siguiente manera:

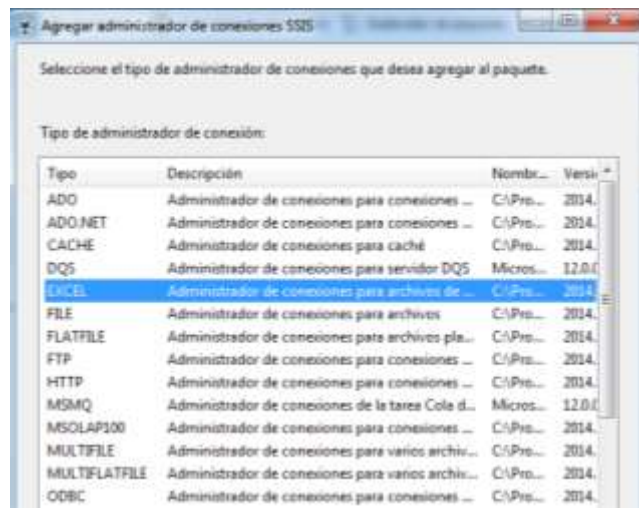


Ilustración 74: Selección del tipo de conexión para Excel
Fuente: Propia

En esta parte se procede a referenciar el archivo de Excel a usar, para el presente proyecto fue necesario crear un archivo que contenga información que no se encontraba detallada de manera comprensible en la base de datos fuente como lo es los estados y los indicadores de los Clientes, a continuación el detalle de la información creada en Excel:

Estados:

CODIGO	DESCRIPCION
A	ACTIVO
NULL	INACTIVO
I	INACTIVO

Tabla 22: Descripción estados Cliente

Fuente: Propia

Indicador:

CODIGO	DESCRIPCION
A	ABIERTO
NULL	CERRADO
C	CERRADO

Tabla 23. Descripción indicadores Cliente

Fuente: Propia

Una vez seleccionado el archivo verificar que la casilla “La primera fila tiene nombres de columna” esté seleccionada:

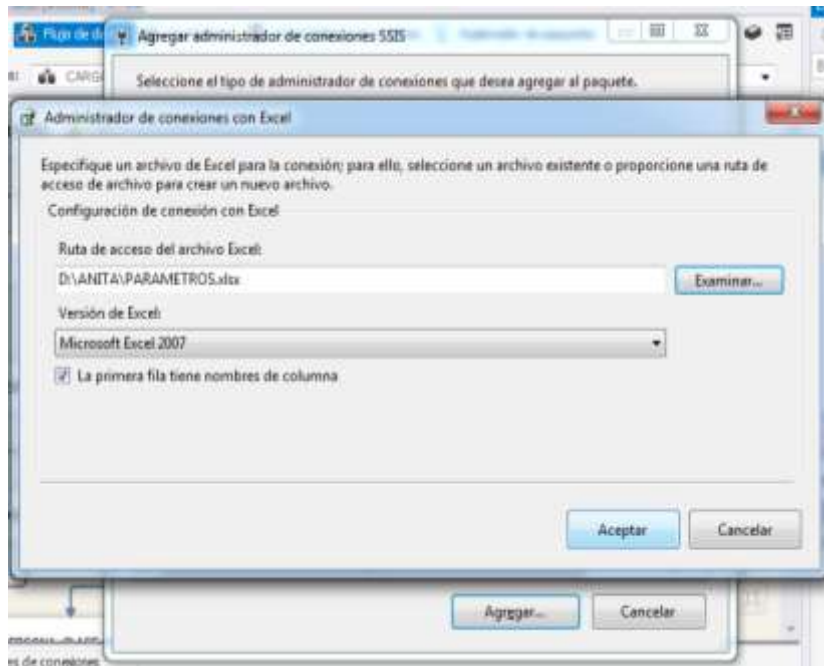


Ilustración 75: Seleccionar el archivo de Excel para conexión
Fuente: Propia

Se puede entonces verificar que los orígenes de datos están cargados:

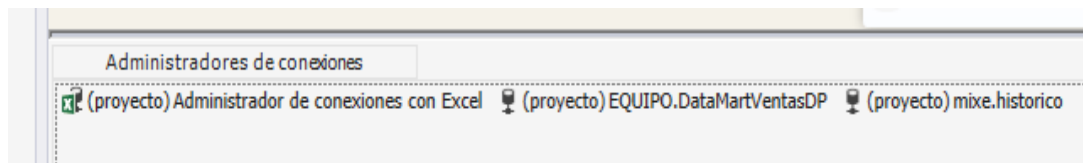


Ilustración 76: Conexiones creadas
Fuente: Propia

Para construir el proceso ETL de cada una de las tablas de la base de datos multidimensional, se utiliza como referencia las tablas de mapeo detalladas anteriormente en donde se especifica de donde se toma la información y en donde va almacenarse.

Para la tabla DIM_CLIENTE, se utiliza el asistente y se realiza la unión de las tablas correspondientes, primero arrastrar al panel de Flujo de datos un Origen de datos OLEDB:

En donde se selecciona la conexión y la tabla de origen CXC_CLIENTE:

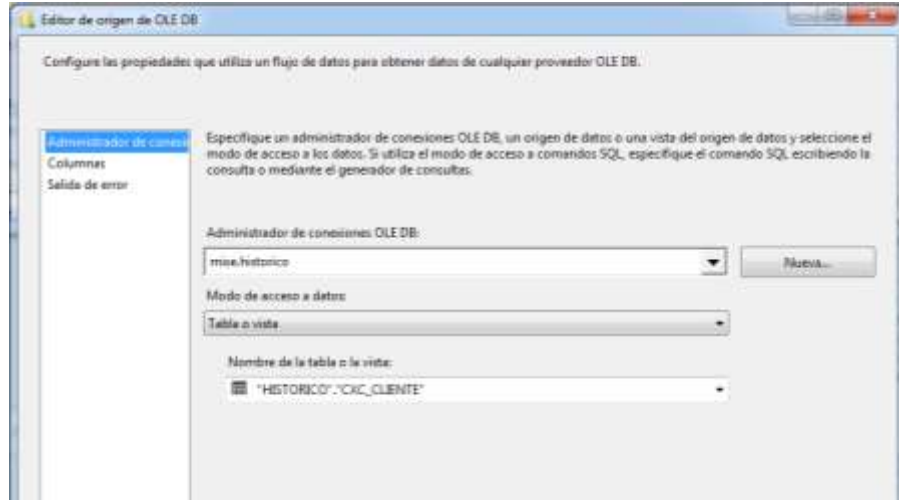


Ilustración 77: Selección de conexión y tabla de origen
Fuente: Propia

En columnas seleccionar solo los campos a usar:

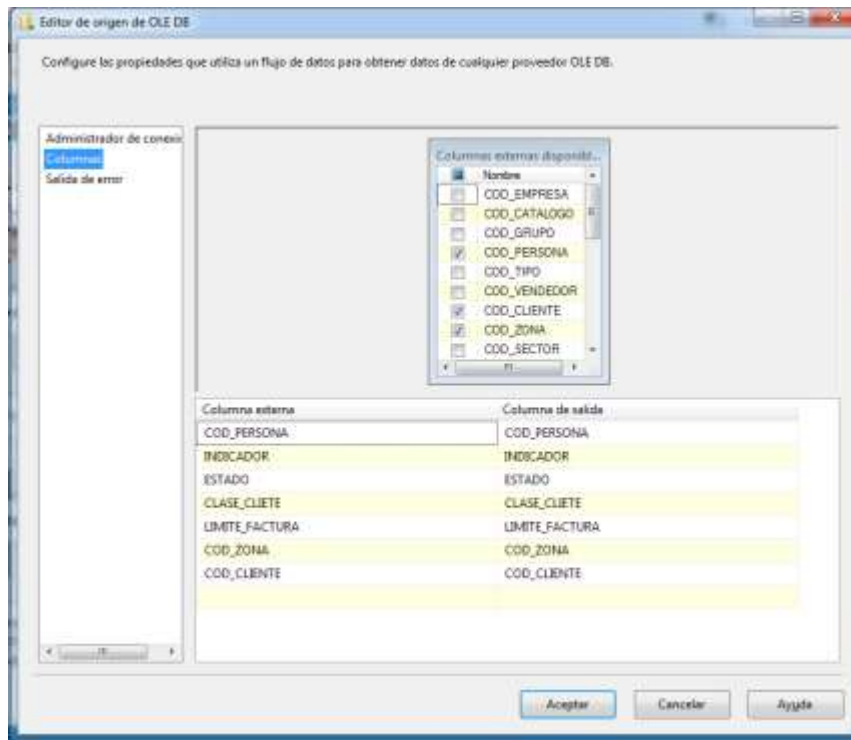


Ilustración 78. Selección de campos de origen Cliente
Fuente: Propia

En un nuevo Origen de OLEDB se selecciona la siguiente tabla de origen GNR_PERSONA con sus campos correspondientes:

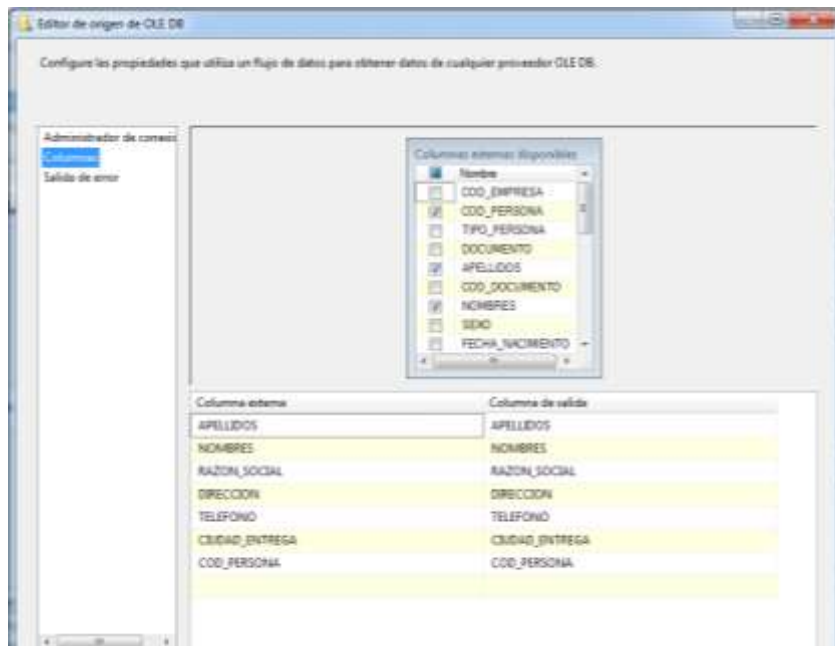


Ilustración 79: Selección de campos de origen Cliente
Fuente: Propia

Para realizar la unión entre las dos tablas es necesario indicar el campo de navegación entre las mismas, a través del componente ORDENAR que se aplica a cada tabla, en este caso el campo es COD_PERSONA.

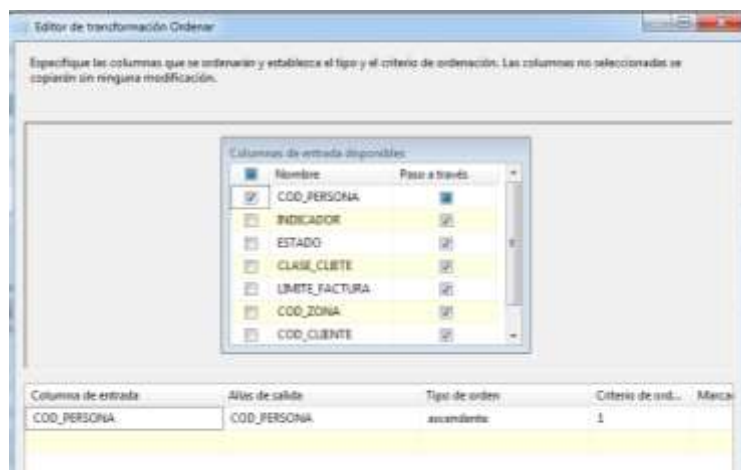


Ilustración 80. Selección del campo de navegación
Fuente: Propia

Luego se procede a unir las dos tablas mediante el componente COMBINACIÓN DE MEZCLA, en el que se puede verificar el campo de navegación y los datos que se mostrarán:

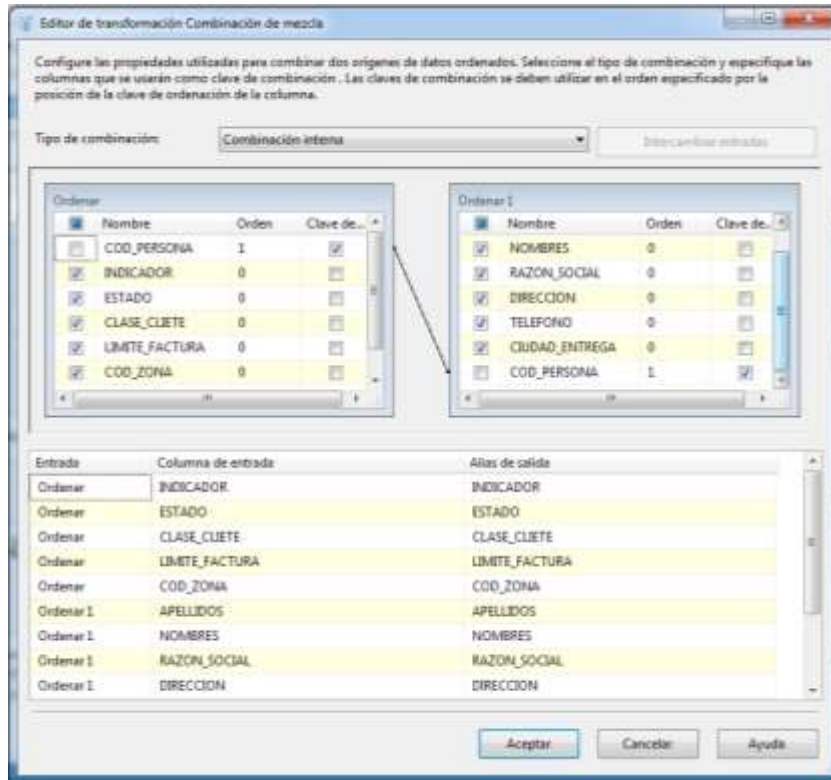


Ilustración 81: Selección de campos de la combinación de tablas

Fuente: Propia

Con este mismo procedimiento se mezcla todas las tablas correspondientes, basándose en el detalle del mapeo definido anteriormente.

Para el caso del origen de Excel se procede a usar el componente ORIGEN DE EXCEL, en el cual se debe especificar el nombre de la hoja del archivo de origen INDICADORES:

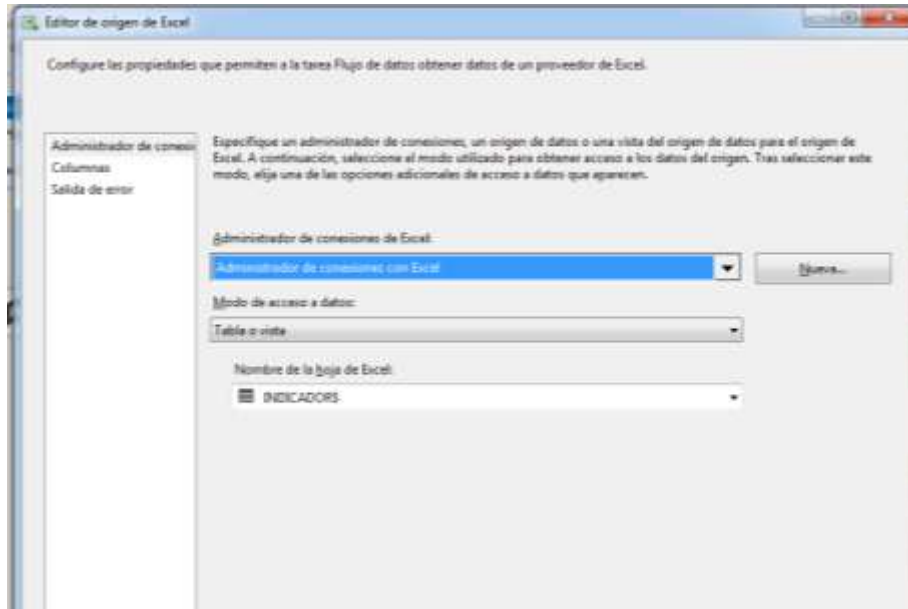


Ilustración 82: Selección del archivo de origen Excel
Fuente: Propia

Y de igual manera seleccionar que campos se va a usar:

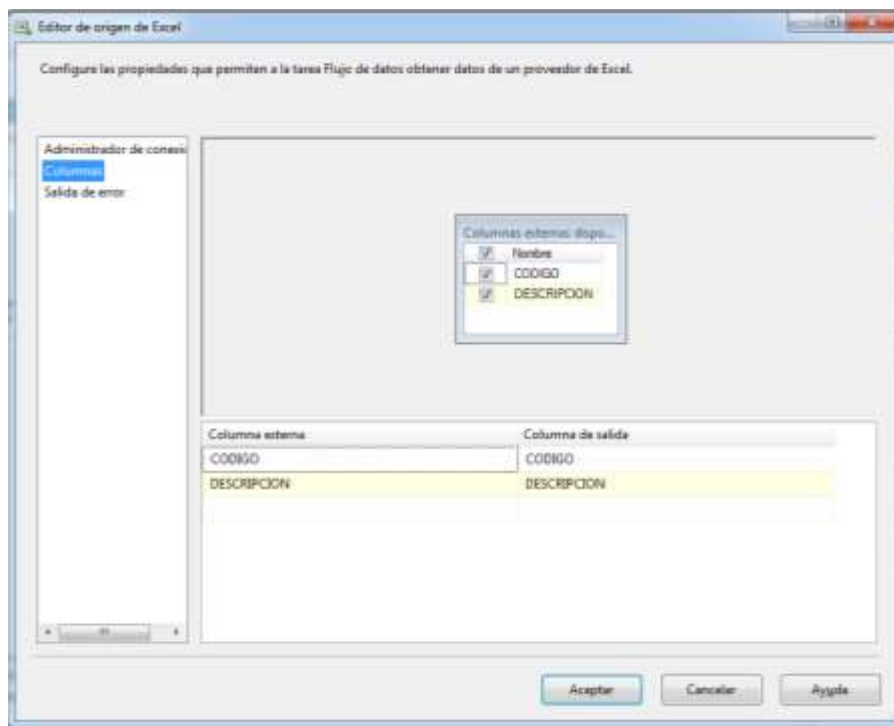


Ilustración 83: Selección de campos del origen Excel
Fuente: Propia

Finalmente antes de depositar todos los datos obtenidos en el destino, para el caso de Excel se debe realizar una transformación del tipo de datos a cadena, de la siguiente manera:

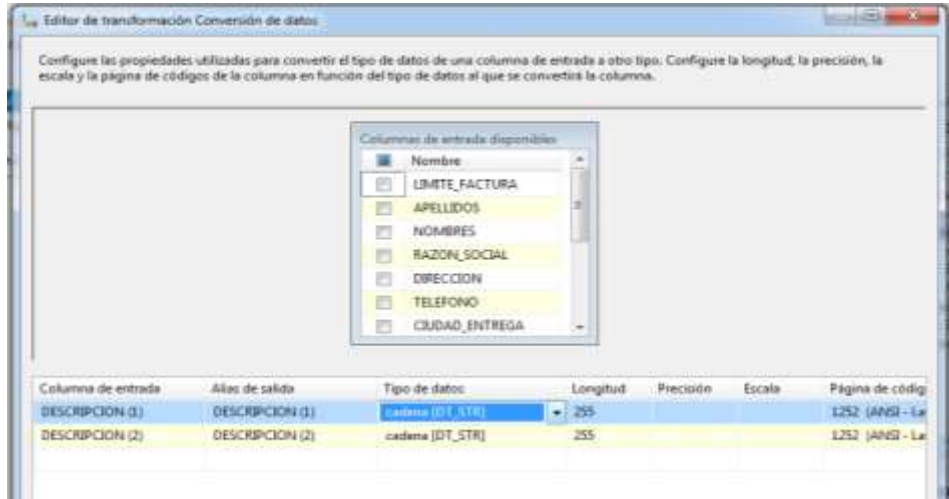


Ilustración 84: Transformación de datos para Excel

Fuente: Propia

Para depositar los datos obtenidos correspondientes a la información de CLIENTES, se utiliza el componente DESTINO OLE DB, en donde se selecciona la base de datos destino y la tabla correspondiente.

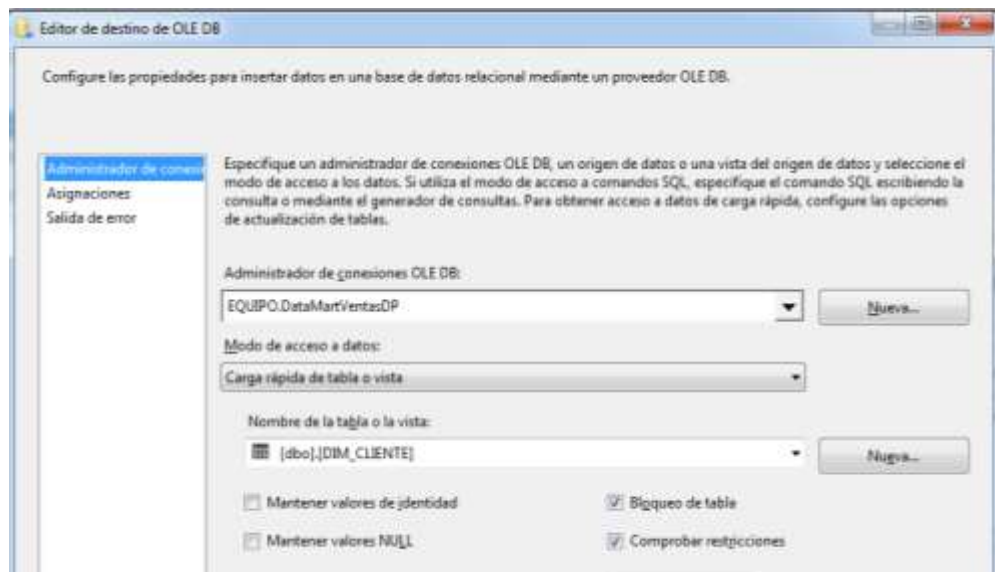


Ilustración 85: Selección de la tabla destino

Fuente: Propia

En la parte de Asignaciones se realiza el mapeo entre los datos de origen y su correspondiente con los campos de destino, de la siguiente manera:

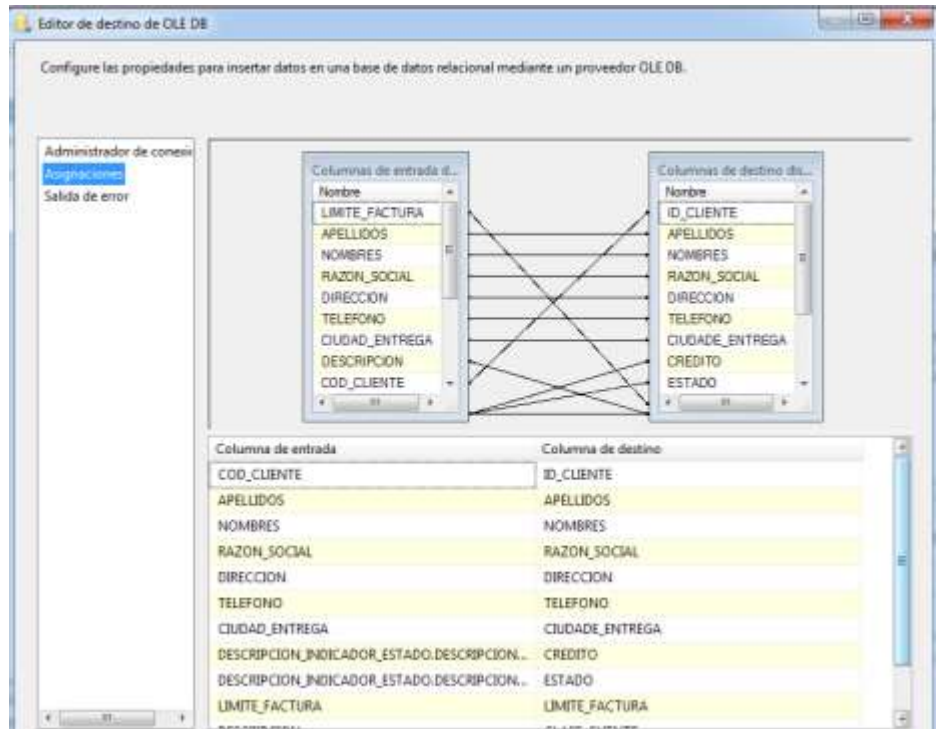


Ilustración 86: Mapeo de datos

Fuente: Propia

Finalmente el proceso para cargar la tabla DIM_CLIENTE, se procesa y se puede verificar ya que los datos están cargados a la base de datos multidimensional:

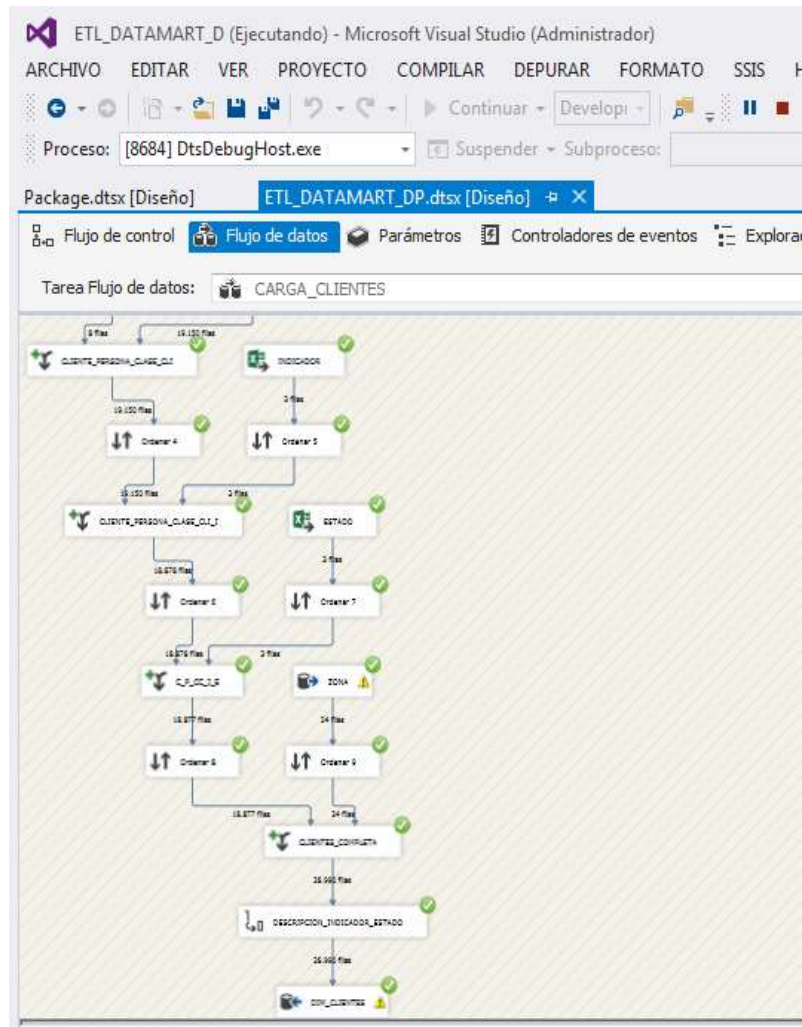


Ilustración 87: Procesamiento de tareas
Fuente: Propia

La información cargada es la siguiente:

ID_CLIENTE	APELLIDOS	NOMBRES	RAZON SOCIAL	DIRECCION	TELEFONO	CIUDADE_ENT	CREDITO	ESTADO	LIMITE_FACTU	CLASE_CLIENTE	ZONA
99	VENTAS DE OHL	NULL	VENTAS DE OHL	IBARRA	NULL	IBARRA	CERRADO	ACTIVO	9999.00	TIENDAS	VARIOS
100	MARCALLA ANL	MARCELO RA...	MARCALLA ANL	CALLE TENA Y...	29058902	IBARRA ALPAC...	ABERTO	ACTIVO	270.00	TIENDAS	IBARRA
101	FLORES PERUG...	MARIA CARMEN	FLORES PERUG...	SIAN MONTAL...	2914356	QUIROGA	ABERTO	ACTIVO	100.00	OTROS	COTACACHI
102	GUERRA ANDR...	TATIANA ALEX...	GUERRA ANDR...	ESMERALDAS Z...	2918107	COTACACHIVIL	CERRADO	ACTIVO	1.00	TIENDAS	COTACACHI
103	ACHINA JEREZ	SILVIA VERON...	ACHINA JEREZ...	SECTOR LA LO...	2918600	CUICOCHA	CERRADO	ACTIVO	140.00	TIENDAS	COTACACHI
104	CALDERON AR...	NELSON ERIVEL	CALDERON AR...	24 DE MAYO E...	2918909	COTACACHI	ABERTO	ACTIVO	100.00	OTROS	COTACACHI
105	MORILLO QUE...	EDWIN FERNA...	MORILLO QUE...	VENEZUELA Y S...	2983305	TULCAN	CERRADO	ACTIVO	100.00	MAYORISTAS	TULCAN
107	SUAREZ SANTA...	EVA YANET	SUAREZ SANTA...	CORY CORY Y...	NULL	IBARRA	ABERTO	ACTIVO	100.00	TIENDAS	IBARRA
108	TULCANAZ NA...	CARLOS PETER	TULCANAZ NA...	TOMAS CAJEL...	NULL	IBARRA	ABERTO	ACTIVO	100.00	TIENDAS	IBARRA
109	GALIANO GAL...	CESAR GONZA...	GALIANO GAL...	BOLIVAR 947 Y V...	NULL	IBARRA	ABERTO	ACTIVO	100.00	TIENDAS	IBARRA
110	TUTILLO CEVA...	ZOLA VICTORIA	TUTILLO CEVA...	GRAN COLOM...	2933578	SAN ANTONIO	ABERTO	ACTIVO	100.00	OTROS	IBARRA
111	MARTINEZ VAR...	OLGA DE IESUS	MARTINEZ VAR...	PANAMA VIA A...	NULL	ATUNTAQUE O...	ABERTO	ACTIVO	200.00	OTROS	ATUNTAQUE
112	ESPAA CORRAL	KARINA VICTO...	ESPAA CORRAL	CALLE BIVBAEL...	2907223	ANDRADE MAR...	CERRADO	ACTIVO	150.00	OTROS	ATUNTAQUE
113	CHAUCA TAMB...	OSCAR PATREC...	CHAUCA TAMB...	SANCHEZ Y OF...	999423094	ANDRADE MAR...	ABERTO	ACTIVO	100.00	OTROS	ATUNTAQUE
114	CALDERON ESP...	MONICA MARE...	CALDERON ESP...	GENERAL ENEL...	2906525	ANDRADE MAR...	ABERTO	ACTIVO	100.00	OTROS	ATUNTAQUE
118	BENAVIDES RO...	EDSON ARVIY	BENAVIDES RO...	TULCAN COM...	99489960	TULCAN	ABERTO	ACTIVO	2000.00	OTROS	TULCAN
120	HURTADO RIVE...	JAIIME WLADIMIR	HURTADO RIVE...	CALDERON RD...	NULL	TULCAN	ABERTO	ACTIVO	100.00	OTROS	TULCAN
121	ROSERO	KARINA	ROSERO KARINA	SUCRE Y URUG...	NULL	TULCAN	ABERTO	ACTIVO	200.00	OTROS	TULCAN
122	LIGHT & FRESH	NULL	LIGHT & FRESH	AV CALDERON...	NULL	TULCAN	ABERTO	ACTIVO	100.00	OTROS	TULCAN
123	REVELO ALZ	ANGEL EFRAIN	REVELO ALZ A...	MERCADO QUL...	996256418	TULCAN	ABERTO	ACTIVO	100.00	TIENDAS	TULCAN
124	ORTIGA FUEL	ROSA RIES	ORTIGA FUEL...	PANAMA TRAS...	NULL	TULCAN	ABERTO	ACTIVO	150.00	OTROS	TULCAN
125	FUEL BLUES	MIREYA ANDREA	FUEL BLUES ME...	BOLIVIA 3023 Y...	990570704	TULCAN	ABERTO	ACTIVO	200.00	TIENDAS	TULCAN
130	MELO CASTRO	BETTY MARICE...	MELO CASTRO...	B DE DOC Y J.M...	2973232	HUACA	ABERTO	ACTIVO	100.00	OTROS	JULIO ANDRADE...
132	CORAL PEREIRA	AMPARO DEL R...	CORAL PEREIR...	AV ANDRES BE...	NULL	TULCAN	ABERTO	ACTIVO	100.00	OTROS	TULCAN

Ilustración 88: Información cargada en tabla destino

Fuente: Propia

Para cargar los datos a la tabla DIM_ARTICULOS y DIM_VENDEDOR se utiliza el mismo procedimiento que DIM_CLIENTE, con los datos correspondientes a cada tabla.

En cuanto a la carga de datos para la tabla DIM_FECHA se crea una vista en el motor de base de datos ORACLE para extraer la fecha del detalle de factura, y se crea un ID para la fecha, se divide la fecha en Mes, Día, Año y mes en letras, usando el siguiente código:

```

SELECT      to_number(TO_CHAR(FECHA_FACTURA, 'ddmmyyyy')) AS id,
            to_date(TO_CHAR(FECHA_FACTURA, 'dd/mm/yyyy'), 'dd/mm/yyyy') AS fecha,
            TO_CHAR(FECHA_FACTURA, 'yyyy') AS ANIO, TO_CHAR(FECHA_FACTURA, 'MM') AS MES,
            TO_CHAR(FECHA_FACTURA, 'DD') AS DIA, TO_CHAR(FECHA_FACTURA, 'MONTH') AS MES_LETRAS
FROM FAC_FACTURA_C

GROUP BY to_number(TO_CHAR(FECHA_FACTURA, 'ddmmyyyy')),
         to_date(TO_CHAR(FECHA_FACTURA, 'dd/mm/yyyy'), 'dd/mm/yyyy'),
         TO_CHAR(FECHA_FACTURA, 'yyyy'),
         TO_CHAR(FECHA_FACTURA, 'MM'),
         TO_CHAR(FECHA_FACTURA, 'DD'),
         TO_CHAR(FECHA_FACTURA, 'MONTH')

```

Ilustración 89: Código de Vista para Fecha
Fuente: Propia

Posteriormente se procede a referenciar esta vista desde el SSIS usando el componente ORIGEN DE OLE DB y depositar esta información mediante el componente DESTINO OLE DB en la tabla de destino DIM_FECHA, en donde se realiza el mapeo de la información:

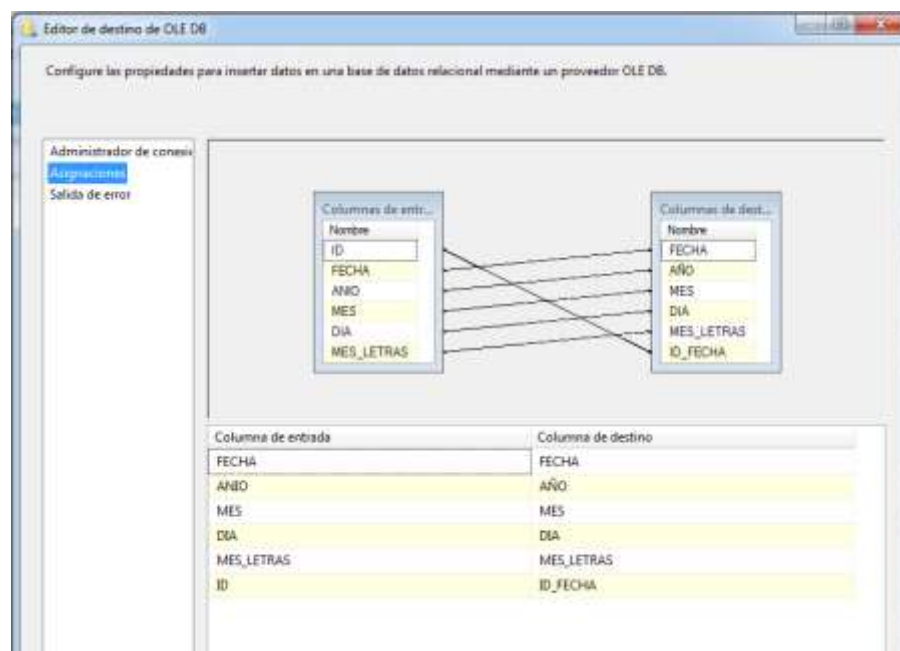
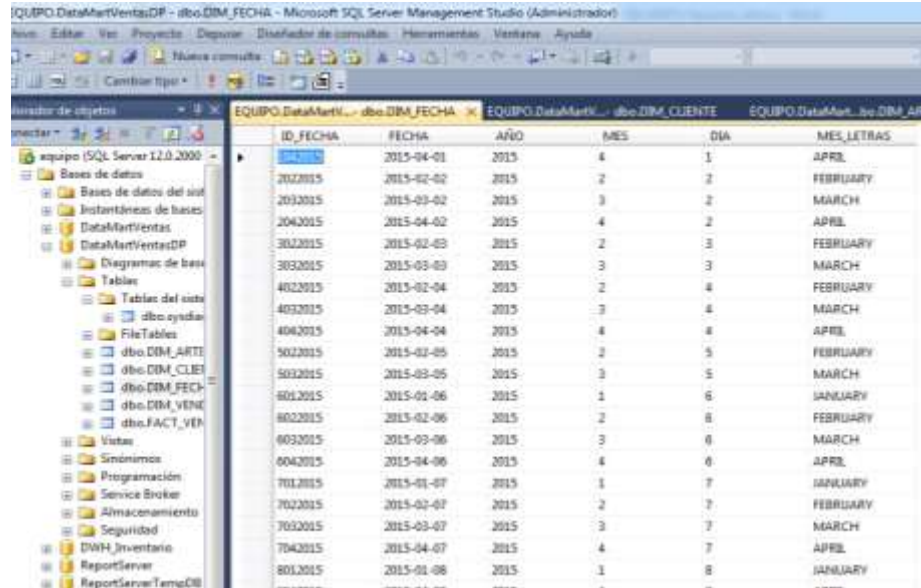


Ilustración 90: Mapeo de datos para Fecha
Fuente: Propia

La información que se carga a la tabla DIM_FECHA es la siguiente:



ID_FECHA	FECHA	AÑO	MES	DIA	MES_LETRAS
2012015	2015-04-01	2015	4	1	APRIL
2022015	2015-02-02	2015	2	2	FEBRUARY
2032015	2015-03-02	2015	3	2	MARCH
2042015	2015-04-02	2015	4	2	APRIL
3022015	2015-02-03	2015	2	3	FEBRUARY
3032015	2015-03-03	2015	3	3	MARCH
4022015	2015-02-04	2015	2	4	FEBRUARY
4032015	2015-03-04	2015	3	4	MARCH
4042015	2015-04-04	2015	4	4	APRIL
5022015	2015-02-05	2015	2	5	FEBRUARY
5032015	2015-03-05	2015	3	5	MARCH
6012015	2015-01-06	2015	1	6	JANUARY
6022015	2015-02-06	2015	2	6	FEBRUARY
6032015	2015-03-06	2015	3	6	MARCH
6042015	2015-04-06	2015	4	6	APRIL
7012015	2015-01-07	2015	1	7	JANUARY
7022015	2015-02-07	2015	2	7	FEBRUARY
7032015	2015-03-07	2015	3	7	MARCH
7042015	2015-04-07	2015	4	7	APRIL
8012015	2015-01-08	2015	1	8	JANUARY

Ilustración 91: Información cargada en tabla DIM_FECHA

Fuente: Propia

Para el ingreso de información en la tabla FACT_VENTAS se utiliza mediante el componente de ORIGEN OLE DEB en la opción modo de acceso a los datos “Comando SQL”, en donde mediante código de SQL se referencia a los datos que forman parte de esta tabla, como son los ID de cada dimensión, y los valores numéricos disponibles para analizar, el código es el siguiente:


```

SELECT F.COD_CLIENTE AS ID_CLIENTE,
D.COD_ARTICULO AS ID_ARTICULO,
F.COD_VENDEDOR AS ID_VENDEDOR,
to_number(TO_CHAR(F.FECHA_FACTURA, 'ddmmyyyy')) AS ID_FECHA,
F.NUM_FACTURA AS NUMERO_FACTURA,
ROUND(D.TOTAL_REG,2) AS TOTAL,
D.AUX_CANTIDAD AS CANTIDAD,
ROUND(ROUND(D.TOTAL_REG,2)/D.AUX_CANTIDAD,2) AS PRECIO_UNITARIO,

ROUND( FUN_GET_COSTO_PROMEDIO_MES
( '01'
, F.FECHA_FACTURA
, D.COD_ARTICULO
, D.COD_BODEGA
, D.COD_INVENTARIO
),2) AS COSTO_PROM_U,

ROUND( FUN_GET_COSTO_PROMEDIO_MES
( '01'
, F.FECHA_FACTURA
, D.COD_ARTICULO
, D.COD_BODEGA
, D.COD_INVENTARIO
) * D.AUX_CANTIDAD , 2)
AS COSTO_TOTAL

```

Ilustración 92: Código para extraer y transformar datos para la tabla de Hechos
Fuente: Propia

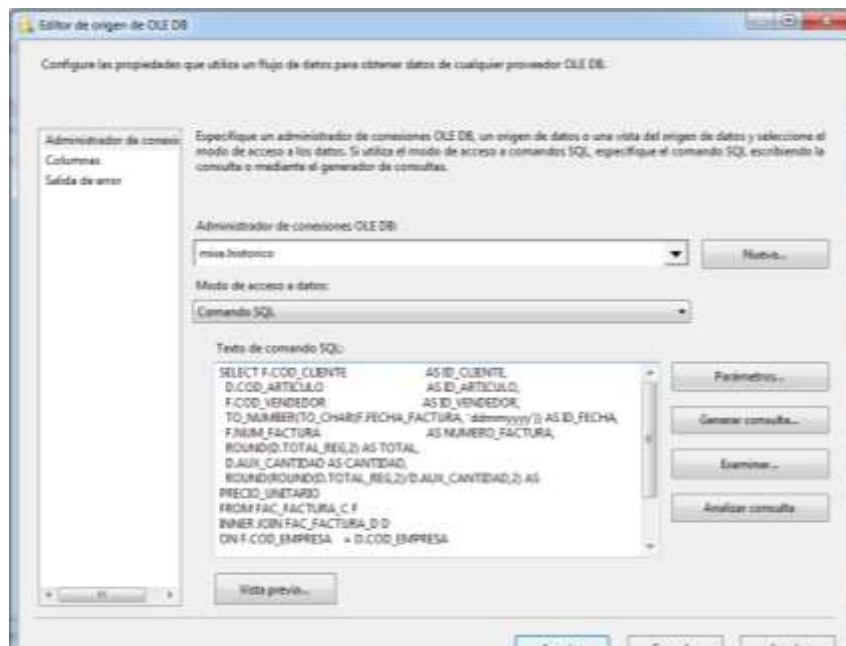


Ilustración 93: Selección origen de datos para tabla de hechos
Fuente: Propia

Y la información que se muestra es la siguiente:

ID_CLIENTE	ID_ARTICULO	ID_VENDEDOR	ID_FECHA	NUMERO_FACTURA	PRECIO_UNIT	TOTAL	CANTIDAD
6177	6177	27	30062015	2713	1.53	9.39	6
5955	6038	27	30062015	2713	0.00	0.00	2
5955	7383	27	30062015	2713	1.04	8.39	8
5955	7180	27	30062015	2713	0.62	3.74	6
5955	7180	27	30062015	2713	0.62	3.74	6
5955	301	27	30062015	2713	0.99	5.95	6
1314	5883	27	30062015	2714	2.89	4.78	2
1314	7187	27	30062015	2714	0.88	7.02	13
1314	912	27	30062015	2714	2.12	8.37	3
1314	7180	27	30062015	2714	0.62	12.48	20
1314	7748	27	30062015	2714	2.17	2.17	1
1314	8138	27	30062015	2714	2.17	2.17	1
1314	8114	27	30062015	2714	1.41	1.41	1
1314	8115	27	30062015	2714	1.41	1.41	1
1314	8127	27	30062015	2714	0.38	2.28	6
1387	3723	27	30062015	2715	2.19	17.51	8
1387	3396	27	30062015	2715	0.89	0.89	1
61286	4454	27	30062015	2716	0.38	196.43	24
6636	6052	27	30062015	2717	4.54	4.54	1
6636	6118	27	30062015	2717	1.40	11.18	8
6636	6118	27	30062015	2717	0.00	0.00	1
6636	3496	27	30062015	2717	1.27	3.80	3
6636	7180	27	30062015	2717	0.62	3.74	6
6636	7527	27	30062015	2717	1.43	1.43	1

Ilustración 94: Información cargada en Tabla de hechos

Fuente: Propia

3.3.6 CONSTRUCCIÓN DEL CUBO CON MODELO MULTIDIMENSIONAL

Para la construcción del Cubo se utiliza SSAS (SQL SERVER Analysis Services), en donde se referencia los datos ya transformados que residen en la base de datos de SQL, que ha sido poblados en el paso anterior, el procedimiento a seguir es el siguiente:

Crear un nuevo proyecto de Anlisy Services “Proyecto Multidimensional y de Minería de datos” el cual lleva el nombre de CUBOS_VENTAS_DN. El primer paso es seleccionar un origen de datos para lo cual se da clic derecho en la pestaña Orígenes de datos y se utiliza el asistente para seleccionar el origen que será la base de datos multidimensional DATAMART_VENTAS_DP de la siguiente manera:

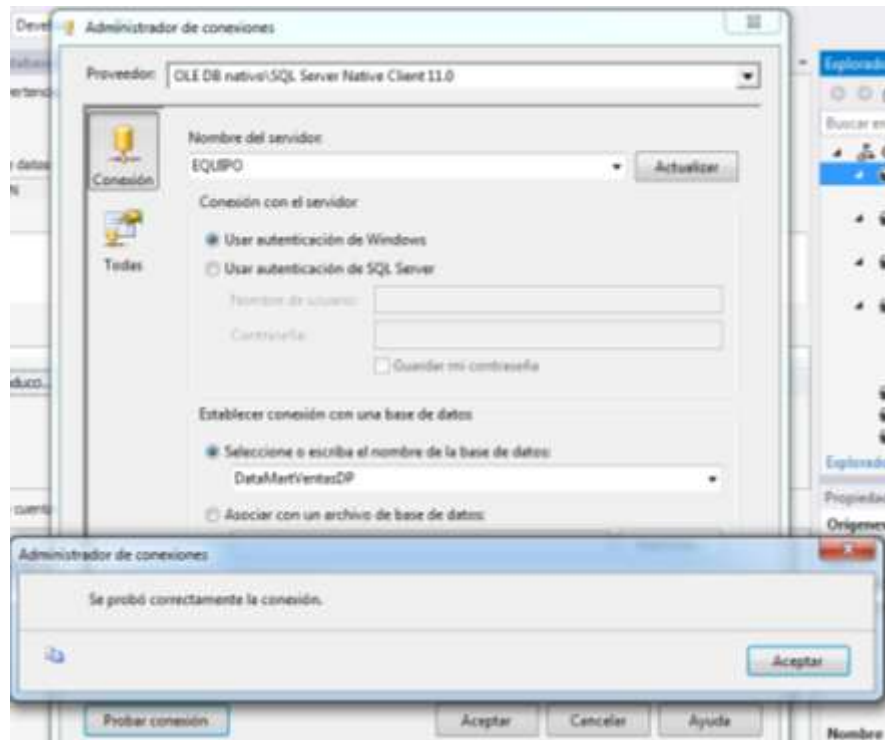


Ilustración 95: Selección del Origen de datos

Fuente: Propia

En este proceso de origen de datos es importante en la información de suplantación especificar la manera como se conecta al origen de datos, especificando el Usuario y contraseña de Windows:

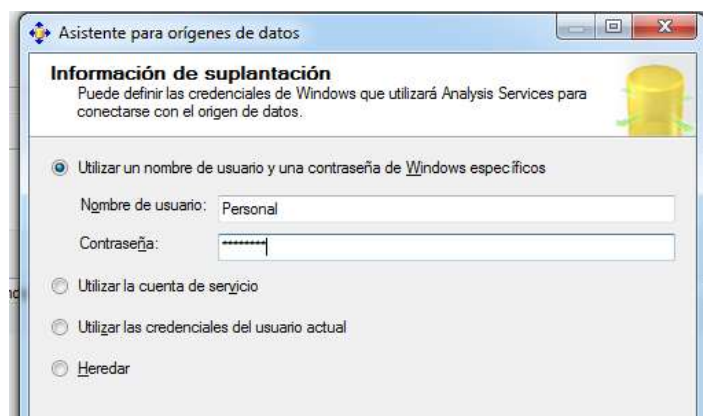


Ilustración 96: Información de Suplantación

Fuente: Propia

Finalmente especificar el nombre del origen de datos y finalizar.

Como siguiente paso corresponde configurar las Vistas de origen de Datos, en esta parte se debe especificar las tablas que se usan del origen seleccionado anteriormente. Las tablas disponibles que se seleccionan son todas las creadas en la base de datos poblada en el motor de base de datos de SQL:

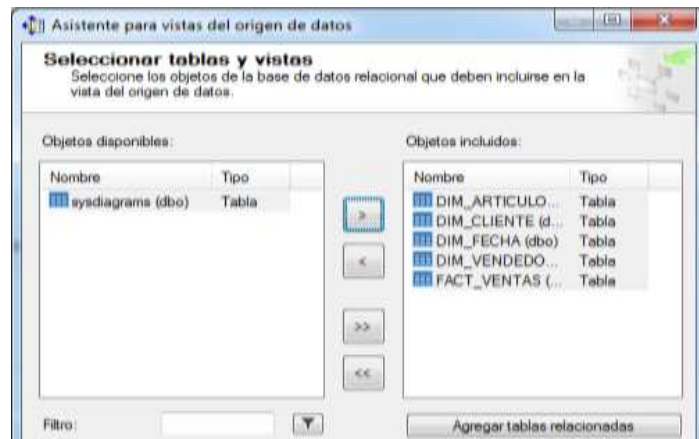


Ilustración 97: Selección de Vistas de origen de datos

Fuente: Propia

Finalmente se especifica el nombre de las vistas de origen y se puede ya visualizar la información de las vistas disponibles:

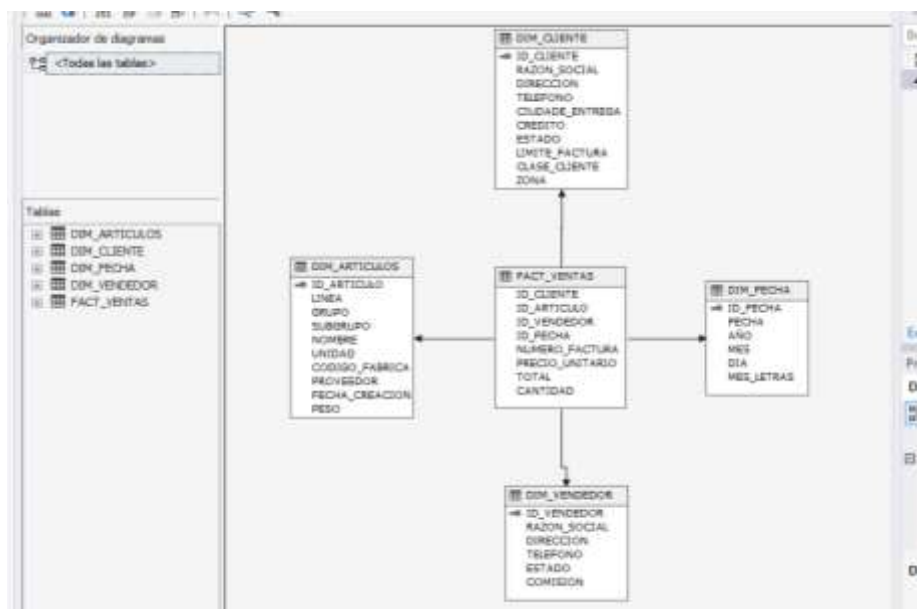


Ilustración 98: Vistas de Origen de datos

Fuente: Propia

Una vez definido los orígenes, se debe crear las dimensiones que formaran parte del cubo, para lo cual se da clic derecho en dimensiones, crear nueva dimensión y mediante el asistente se crea cada dimensión especificando las tablas y los campos a usar para cada dimensión, el proceso a seguir para cada dimensión es el siguiente:

Primero especificar la tabla de origen:

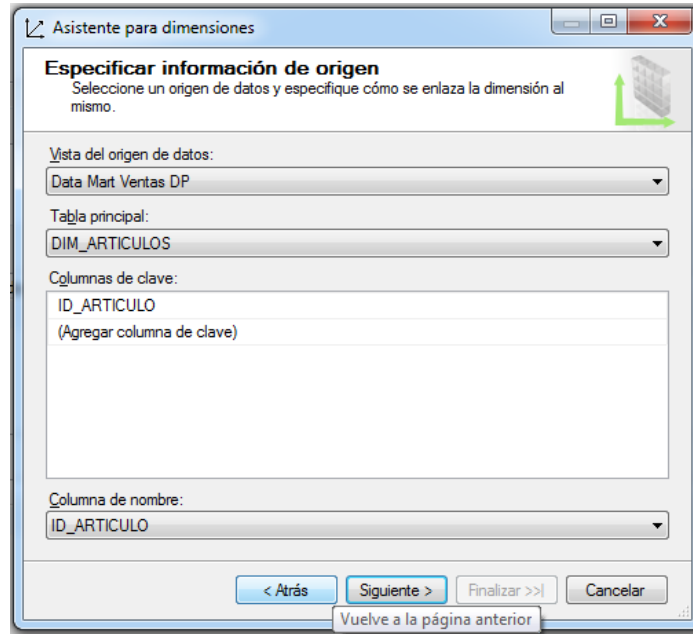


Ilustración 99: Especificación de tabla origen

Fuente: Propia

Luego se selecciona los campos a usar

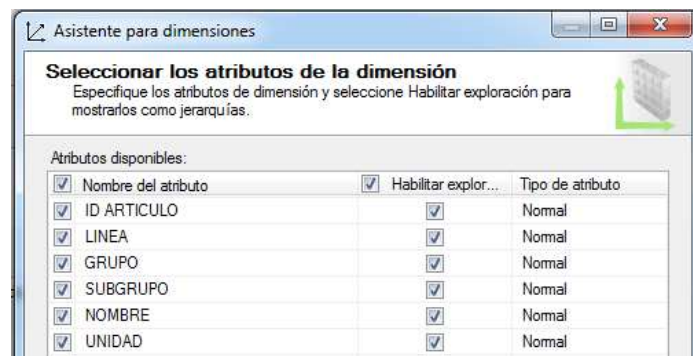


Ilustración 100: Especificación de atributos

Fuente: Propia

Finalmente se especifica el nombre de la dimensión y se puede visualizar la información asignada, en la pestaña estructura de dimensión en donde se puede trabajar con cada dimensión.

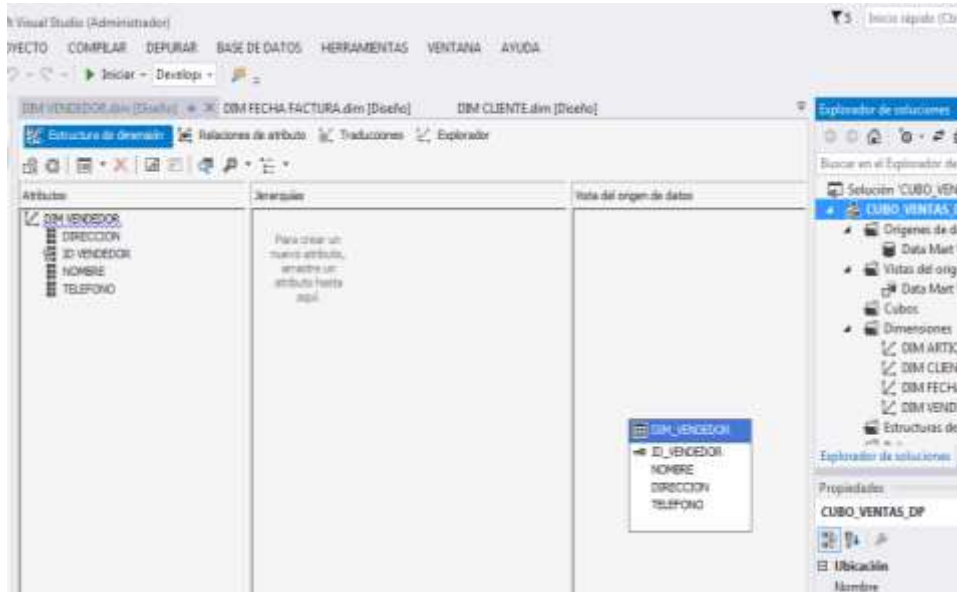


Ilustración 101: Estructura de la dimensión

Fuente: Propia

Una vez definido cada una de las dimensiones se procede a crear el cubo en donde se especifica que tabla actúa como tabla de hechos y se definen las medidas, para lo cual se sigue el siguiente proceso: Clic derecho en Cubos, nuevo cubo y especificar, usar tablas existentes:

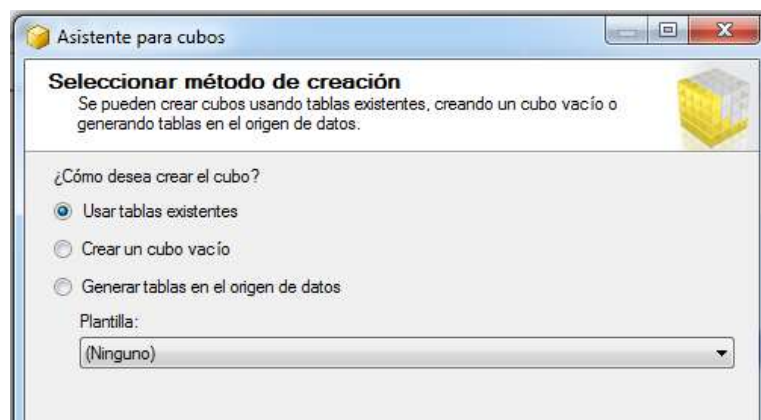


Ilustración 102: Selección método de creación Cubo

Fuente: Propia

Especificar la tabla de hechos:

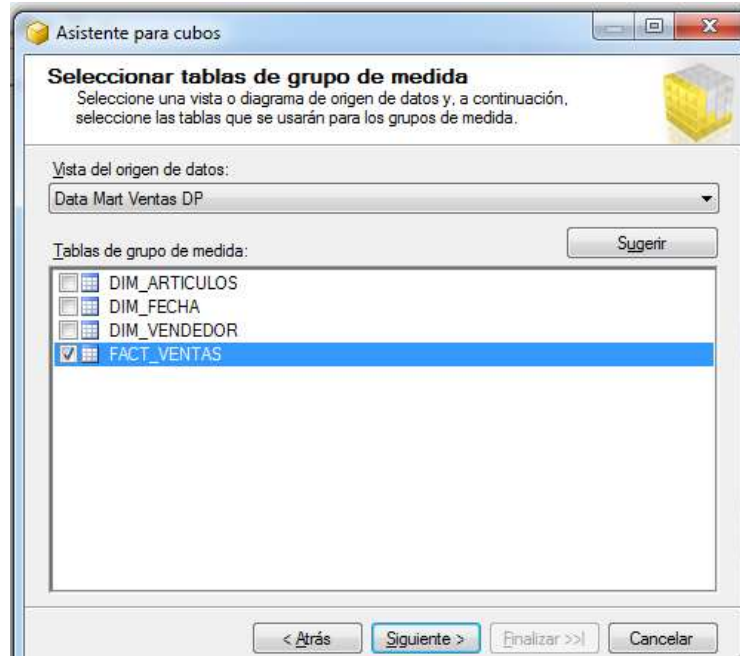


Ilustración 103: Selección de tabla de hechos

Fuente: Propia

Especificar el grupo de medidas:

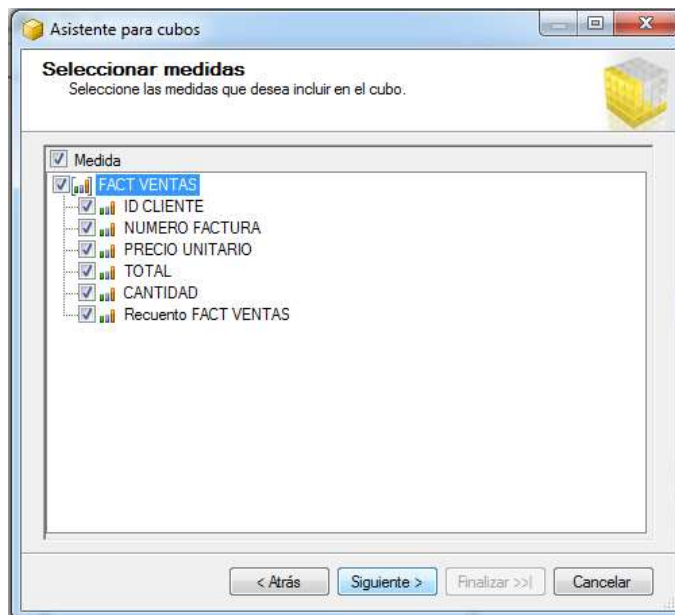


Ilustración 104: Selección de medidas

Fuente: Propia

Especificar las dimensiones:

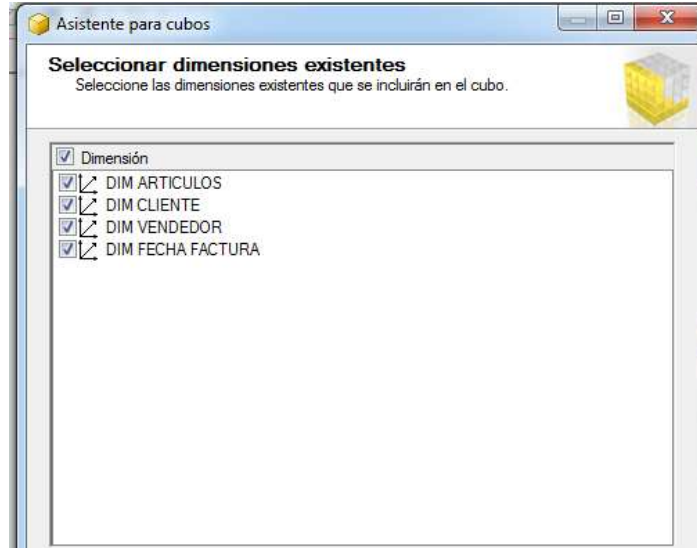


Ilustración 105: Selección de Dimensiones

Fuente: Propia

Y finalmente se puede observar el cubo:

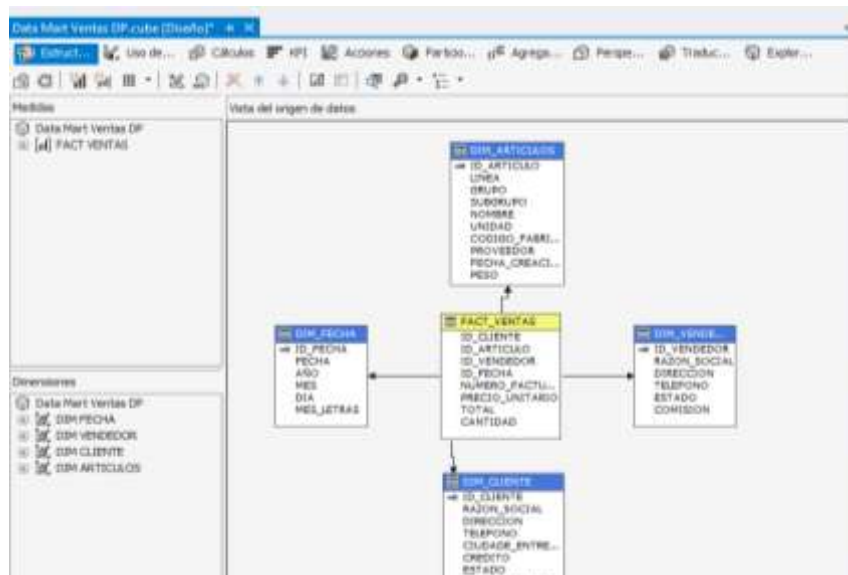


Ilustración 106: Visualización del Cubo

Fuente: Propia

Es importante para que el cubo funcione correctamente verificar las relaciones de las dimensiones con el grupo de medidas:

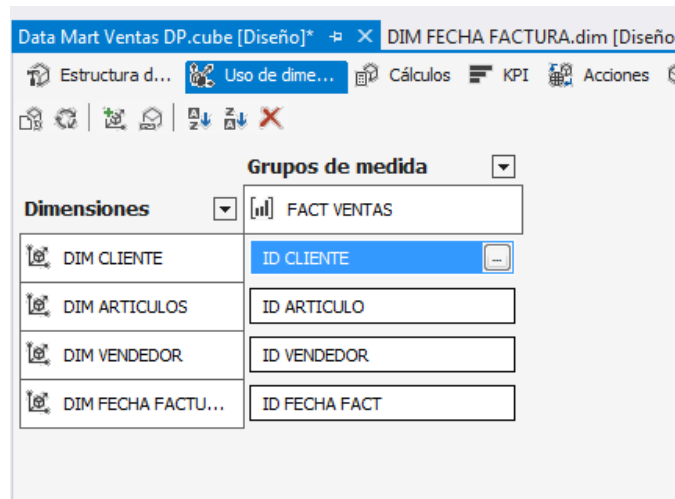


Ilustración 107: Correspondencia de dimensiones con grupo de medidas
Fuente: Propia

Finalmente se implementa y ejecuta el cubo para que la información se pueda visualizar.

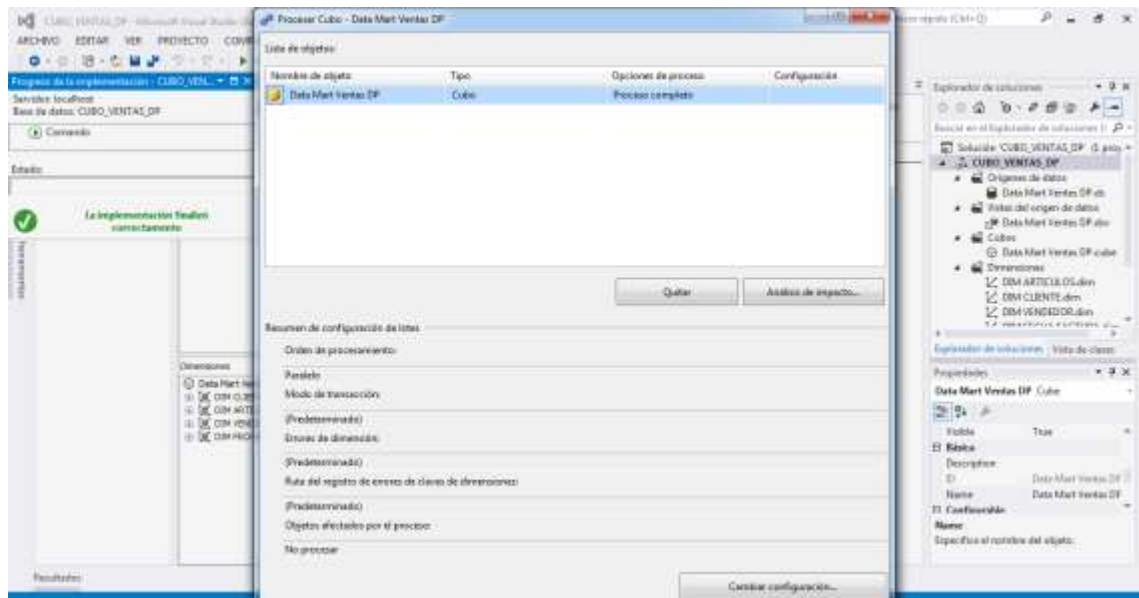


Ilustración 108: Implementación del Cubo
Fuente: Propia

Ahora ya se puede referenciar la información directamente dese el explorador de Analysis Services mediante una pequeña pestaña ubicada en la parte

superior del explorador, que automáticamente exporta los datos a Excel, de la siguiente manera:

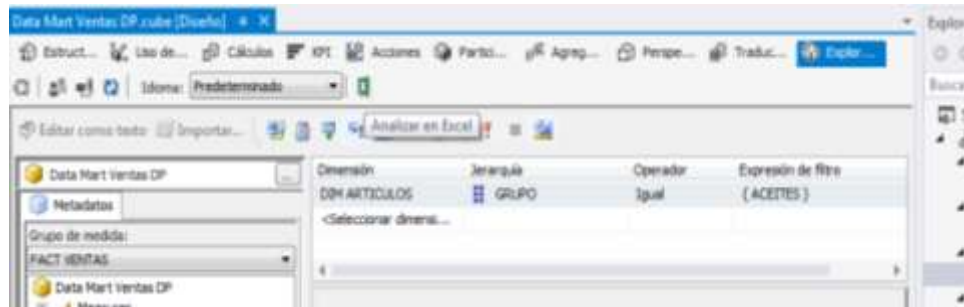


Ilustración 109: Exportación datos a Excel
Fuente: Propia

A continuación se observa al lado derecho los campos de las dimensiones con los que se puede trabajar y al lado izquierdo el formato de la tabla dinámica:

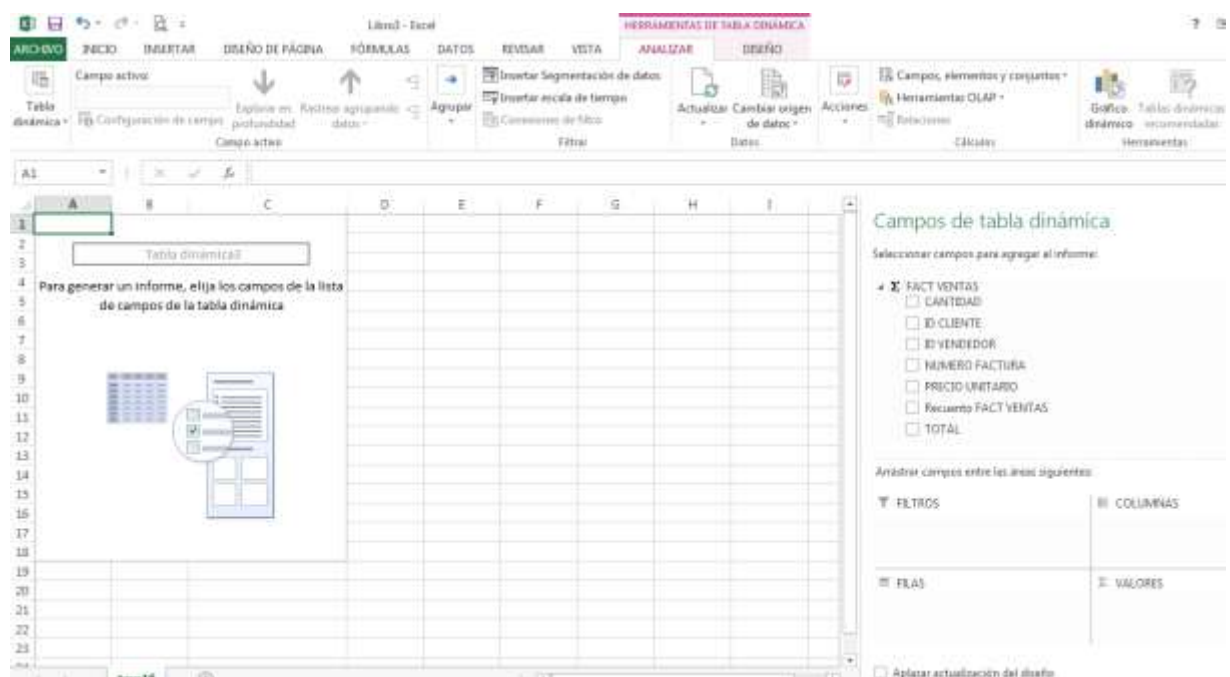


Ilustración 110: Visualización de datos en Excel
Fuente: Propia

Cabe indicar también que en este tipo de modelos multidimensionales una vez construido el cubo se puede interactuar con la información, creando jerarquías, cálculos, KPI minería de datos, entre otras posibilidades que hacen de este

modelo un componente poderoso de análisis de datos recomendado especialmente para nivel corporativo, como se mencionó anteriormente.

3.3.7 CONSTRUCCIÓN DEL MODELO TABULAR

Como se ha propuesto en el presente proyecto los informes o reportes empresariales se muestran en Excel usando el componente Power Pivot, el mismo que admite el acceso al modelo de datos desde Excel, para ser analizados y visualizados por el usuario final, a través del uso de tablas dinámicas que muestran la información y permiten interactuar con ella.

Ahora bien para empezar desde Excel abrir la pestaña POWER PIVOT y luego administrar:

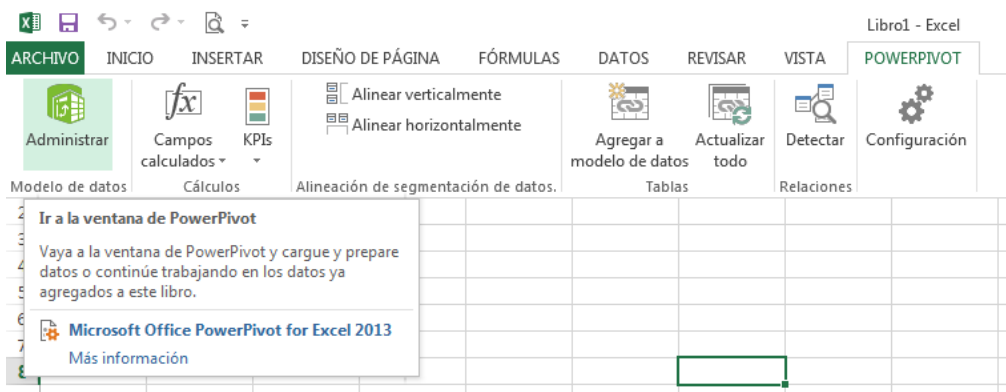


Ilustración 111: Menú de POWER PIVOT

Fuente: Propia

Dentro del entorno de Power Pivot, se procede a referenciar el origen de datos, que en este caso al igual que en el modelo multidimensional desarrollado en el tema anterior, el origen es la base de datos que fue cargada mediante los procesos de Extracción Transformación y Carga, de la siguiente manera:

Dentro de las opciones de “Obtener datos externos” abrir la opción “De otros orígenes”, en esta lista como primera opción se encuentra Microsoft SQL Server, que es la opción que se escoge, sin embargo se puede observar que Power Pivot permite crear conexión a múltiples orígenes de datos como Oracle, modelos multidimensionales, archivos de Excel, etc.

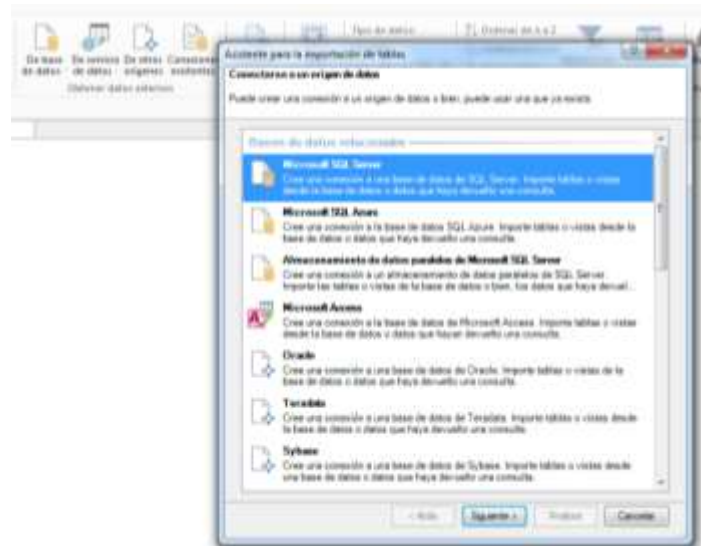


Ilustración 112: Selección de origen de datos
Fuente Propia

Una vez seleccionado el origen se ingresa la información de la base de datos y se prueba conexión de la siguiente manera:

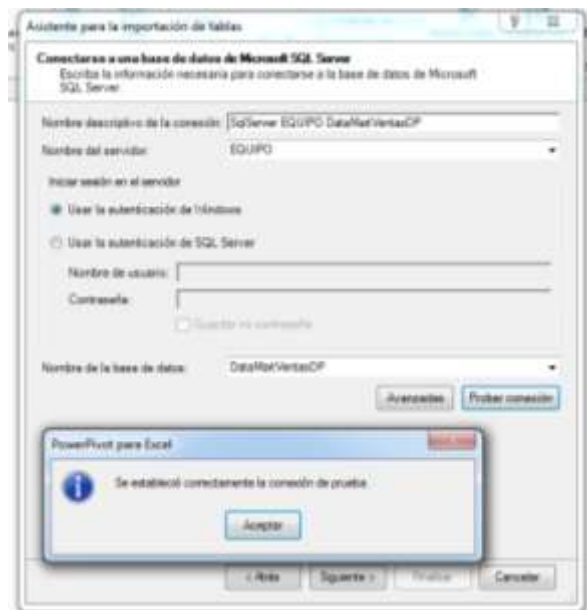


Ilustración 113: Conexión al origen de datos
Fuente: Propia

Una vez seleccionado el origen en el administrador para importación de tablas, seleccionar la opción “Seleccionar en una lista de tablas y vistas para elegir los datos a importar”:



Ilustración 114: Administrador para la importación de tablas
Fuente: Propia

Dentro de esta opción se muestra una lista con las tablas disponibles en el origen de datos, se puede en esta parte cambiar el nombre con el que se desea que aparezca cada dimensión en una libro diferente de Excel, como también seleccionar filtros y ver hacer una vista previa de la información, de las tablas de la lista se escoge las siguientes:

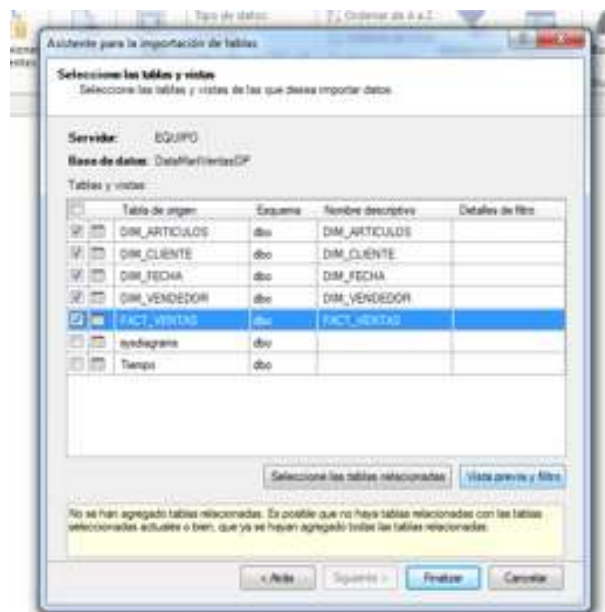


Ilustración 115: Selección de tablas de origen
Fuente: Propia

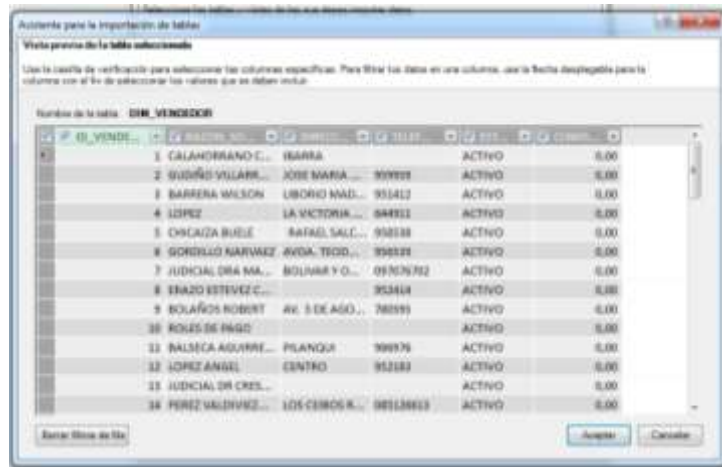


Ilustración 116: Vista previa de tablas seleccionadas
Fuente: Propia

Una vez seleccionado y configurado las tablas a usar empieza el proceso de exportación de la información hacia Excel, como se puede ver en la siguiente imagen de cada dimensión se exporta con la cantidad de registros existentes, y en la tabla de hechos por ejemplo se puede ver que se exportaron más de 900.000 mil registros, información que en el Excel con Power Pivot es muy rápido y fácil de navegar ya que es guardada y comprimida en memoria, mientras que en el modelo multidimensional cada vez que se haga una petición de información se conecta la base de datos origen, tardando un poco más en los tiempos de respuesta de la información:

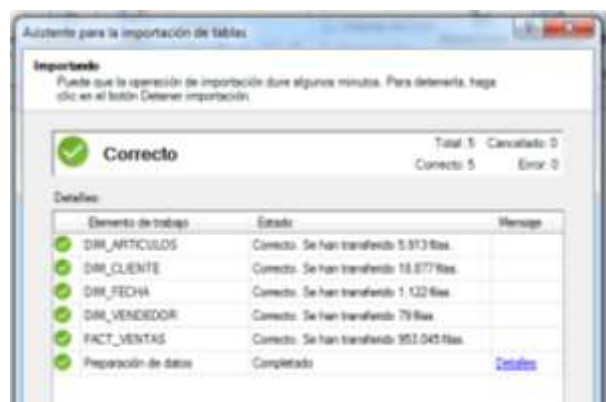


Ilustración 117: Importación de tablas
Fuente: Propia

Una vez exportado los datos a Excel se puede ver la información en un entorno bastante familiar, por ejemplo cada dimensión aparece como si fuera un libro de Excel, cada columna tiene la opción de filtrar, en la parte inferior existe un espacio para cálculos y agregaciones, y en el menú hay varias opciones para trabajar con la información existente, realizando clic derecho en cada columna igual hay varias opciones que permiten editar dicha columna, como por ejemplo permitir o no mostrar en el entorno del usuario final, entre otras.

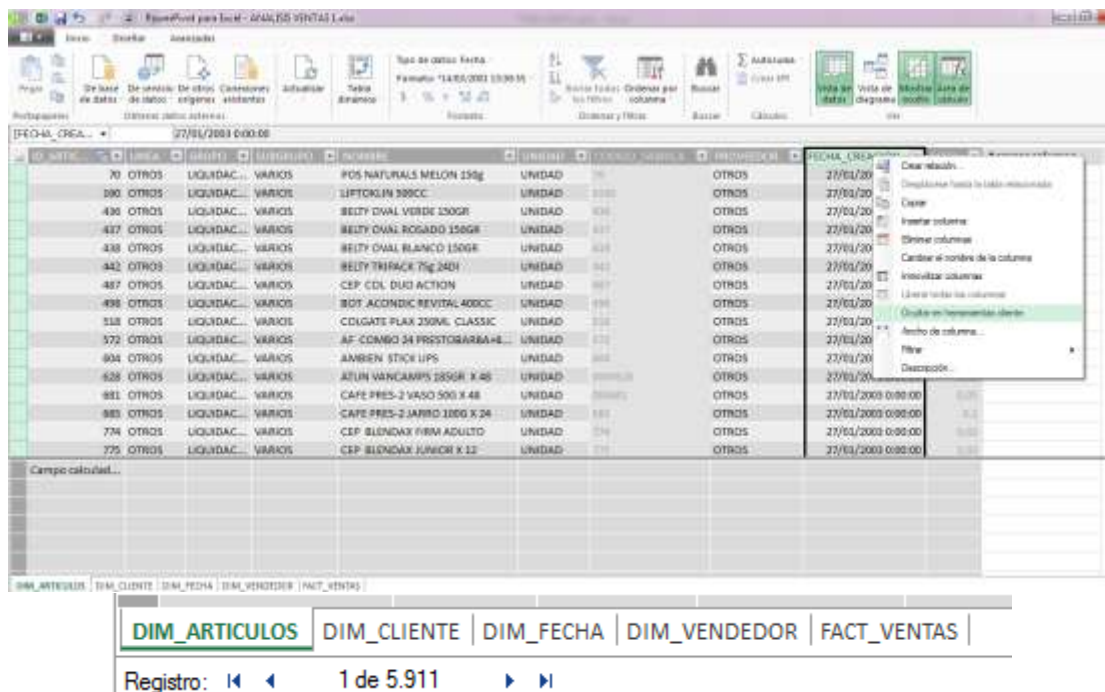


Ilustración 118: Visualización de la información

Fuente: Propia

Para crear una nueva medida como lo es la utilidad, en este mismo entorno se agrega una columna a la que se le aplica la fórmula: $=[\text{TOTAL}] - [\text{COSTO_TOTAL}]$ y de igual manera por ejemplo para crear una columna con el estado de la utilidad con la siguiente formula: $=\text{IF}([\text{UTILIDAD}] > 1, \text{"GANANCIA"}, \text{"PERDIDA"})$.

ID_VENDEDOR	ID_CLIENTE	NUMERO_FACTURA	PRECIO_UNITARIO	CANTIDAD	TOTAL	COSTO	COSTO_TOTAL	UTILIDAD	Estado
22	34	14022012	252083	0	1	0	0	0	50,00 PERDIDA
22	43	28322012	254715	0	1	0	0	0	50,00 PERDIDA
22	14	28022012	254729	0	1	0	0	0	50,00 PERDIDA
22	48	30252012	200000	0	1	0	0	0	50,00 PERDIDA
22	43	17042012	259308	0	1	0	0	0	50,00 PERDIDA
22	43	30042012	200009	0	1	0	0	0	50,00 PERDIDA
22	14	3052012	200094	0	1	0	0	0	50,00 PERDIDA
22	32	28042012	209606	0	1	0	0	0	50,00 PERDIDA
22	9	24042012	258610	0	1	0	0	0	50,00 PERDIDA
22	54	28042012	208610	0	1	0	0	0	50,00 PERDIDA
22	43	9052012	261245	0	1	0	0	0	50,00 PERDIDA
22	52	10052012	200420	0	1	0	0	0	50,00 PERDIDA
22	68	28042012	200762	0	1	0	0	0	50,00 PERDIDA
22	05	4052012	261099	0	1	0	0	0	50,00 PERDIDA
22	09	4052012	261087	0	1	0	0	0	50,00 PERDIDA
22	08	13052012	261525	0	1	0	0	0	50,00 PERDIDA
					TOT 1: 29052012				

Ilustración 119: Visualización de cálculos

Fuente: Propia

Para crear jerarquías, mirar las relaciones o modificar el modelo de datos hay que dirigirse a Vista de diagramas, en donde se ha creado las jerarquías para productos, clientes y tiempo de la siguiente manera:

En vista de diagramas se muestra el modelo con cada dimensión, en la parte superior derecha de cada dimensión existe la opción para crear jerarquías, se puede poner un nombre a la jerarquía y arrastrar los campos que queremos que se desplieguen:

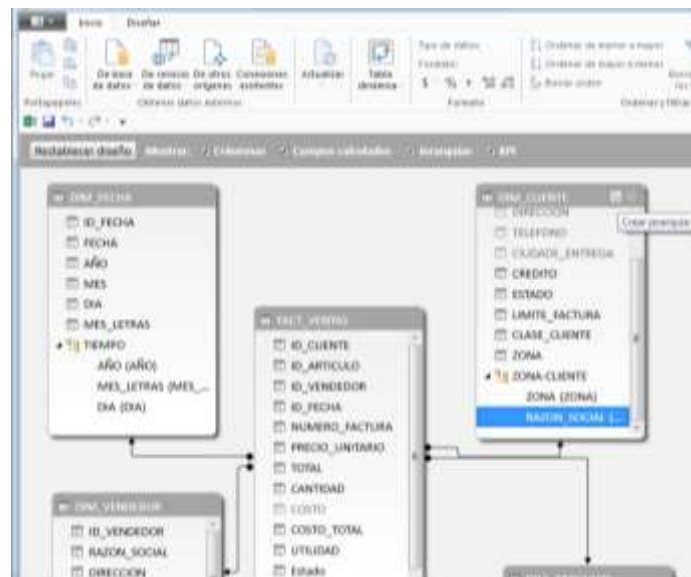


Ilustración 120: Vista de diagramas

Fuente: Propia

Para el presente análisis se ha creado las siguientes jerarquías que permitirán navegar por la información indicada:

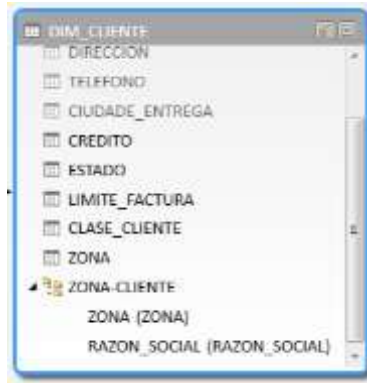


Ilustración 121: Jerarquía Cliente
Fuente Propia

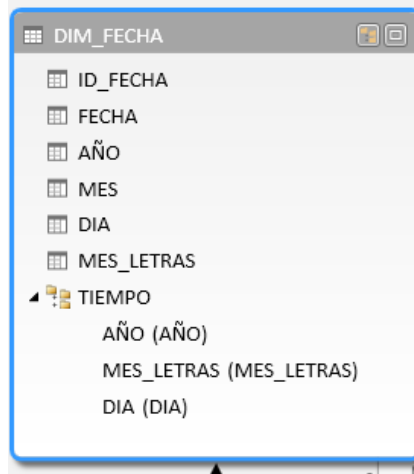


Ilustración 122: Jerarquía Tiempo
Fuente Propia

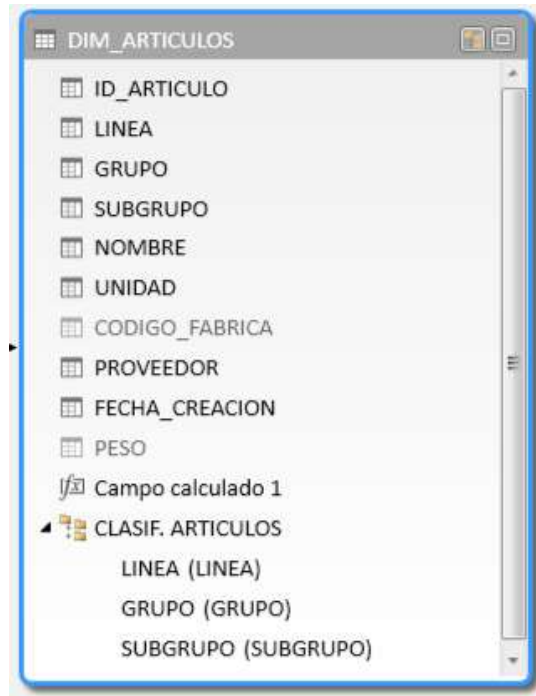


Ilustración 123: Jerarquía Clasificación Artículos
Fuente: Propia

Para visualizar los datos en el entorno del usuario final, hay que exportarlos a una ventana normal de Excel a través de la opción “Tabla dinámica” del menú principal, en donde se puede escoger varias opciones para la presentación de la información:

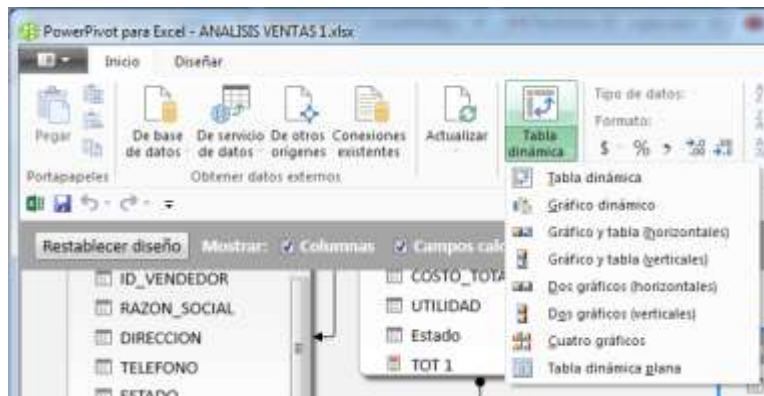


Ilustración 124: Opciones de Tabla dinámica
Fuente: Propia

Cabe destacar que de aquí en adelante se trabaja con dos ventanas de Excel, la una es del modelo tabular y la otra es del entorno del usuario final.

Una vez exportado el modelo a entorno de usuario final se puede interactuar con la información y ver que necesidades requiere de acuerdo a los reportes, por ejemplo en el presente proyecto al momento de consultar por fecha, el mes se ordena en forma alfabética automáticamente, lo cual no es correcto, por lo tanto hay que regresar al modelo tabular y ordenar la columna mes de la dimensión Fecha, primero seleccionar la columna MES_LETRAS luego hacer clic en la opción Ordenar por Columna del menú de la parte superior, y escoger ordenar por columna mes, ya que esta columna contiene los valores en números de los meses.

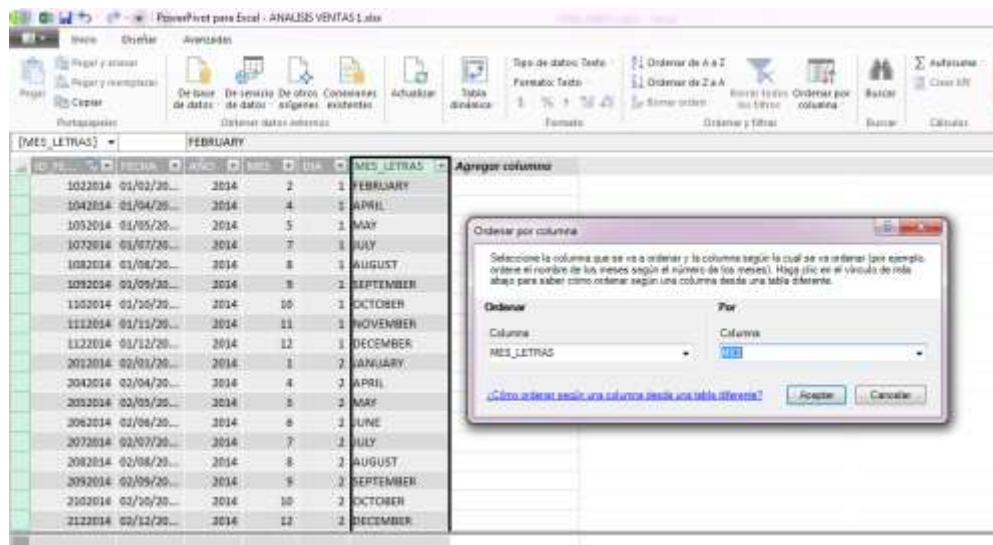


Ilustración 125: Ordenar fecha

Fuente: Propia

De esta manera se interactúa con el modelo tabular y la información final la cual se actualiza automáticamente, finalmente se procede a guardar el archivo como un archivo normal de Excel, en donde se puede evidenciar una propiedad muy importante de Power Pivot que es el alto nivel de compresión en la información, en donde el archivo de análisis es apenas de 18.29 KB.


Nombre	Tamaño
 ANALISIS VENTAS 1	18.298 KB

Ilustración 126: Tamaño en disco del archivo de análisis

Fuente: Propia

3.4 PRUEBAS

El objetivo en esta fase es probar el modelo y la información que contiene, por medio de los reportes definidos en el alcance del proyecto, para lo cual se interactúa con la información exportada al Excel, a continuación se muestran varios de los reportes que se obtienen como prueba del modelo.

Diseño de informes en Power Pivot

Dentro del ambiente de Excel una vez insertado la tabla dinámica desde Power Pivot, se puede observar a la derecha el menú para interactuar con la información de las tablas dinámicas, en la parte superior hay un menú de las herramientas de tabla dinámica con varias opciones para analizar la tabla o configurar su diseño, entre otras opciones.



Ilustración 127: Ambiente de Excel para tablas dinámicas con Power Pivot

Fuente: Propia

Para ir elaborando los reportes se hace uso de todas estas herramientas de acuerdo a los requerimientos, varios de los reportes elaborados son:

- Reporte de clase de clientes por épocas del año, en donde se puede observar la información tanto en una tabla dinámica como en gráfico e

interactuar con la información haciendo usos de los diferentes filtros que se a colocado, para colocar los filtros de la derecha hay que dirigirse al menú en la opción “insertar segmentación de datos” y elegir el campo por el que se desea filtrar la información. En este reporte se interactúa con tres dimensiones como son Cliente, Tiempo y Artículos a través de la medida “TOTAL” que registra el total de venta.

Interpretando este reporte se puede observar que la clase de cliente Tiendas en el mes de noviembre es cuando más compra, de ahí se puede seguir navegando por la información y consultar que línea de producto compra más, en qué año compro más, cual es la peor clase de clientes y muchos informes más.

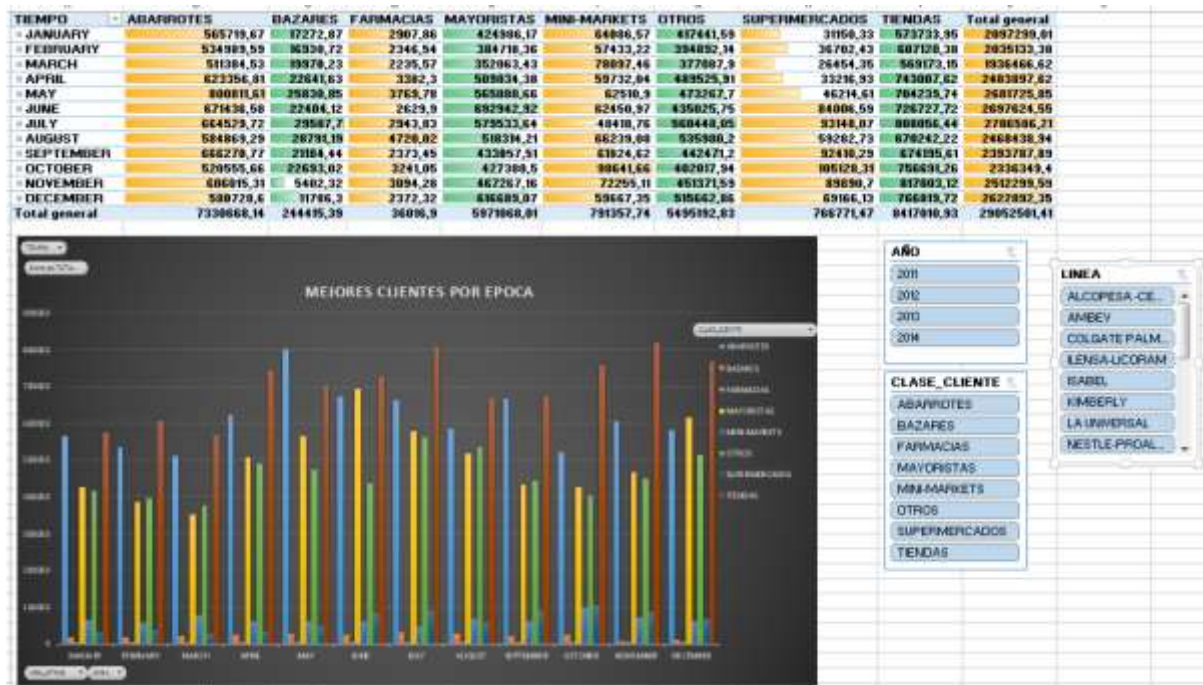


Ilustración 128: Reporte Compras Clientes por época

Fuente: Propia

- Para el siguiente reporte, se analiza a través de las medidas, las Ventas y Utilidad de los artículos, en este reporte se puede navegar de acuerdo a la granularidad de la dimensión artículo, es decir se puede observar las ventas y utilidad por línea, grupo y subgrupo; además se puede agregar también la zona en la que se comercializa dicho producto, y todo esto en

base al tiempo y su respectiva granularidad; dando así un reporte bastante interesante en cuanto a los productos.

En este reporte se pudo observar una desventaja en cuanto a la clasificación de los productos, ya que existe muchos productos que no están clasificados por línea, grupo o subgrupo, por lo cual se almacenan en la variable Otros, provocando que la información resultante no sea tan específica, pero sin embargo de los productos clasificados se puede también observar que la Línea de productos Isabel es la las vendida y la que genera mayor utilidad en general. En este reporte igual se puede obtener la información de acuerdo a todos los criterios disponibles en el panel de filtros.

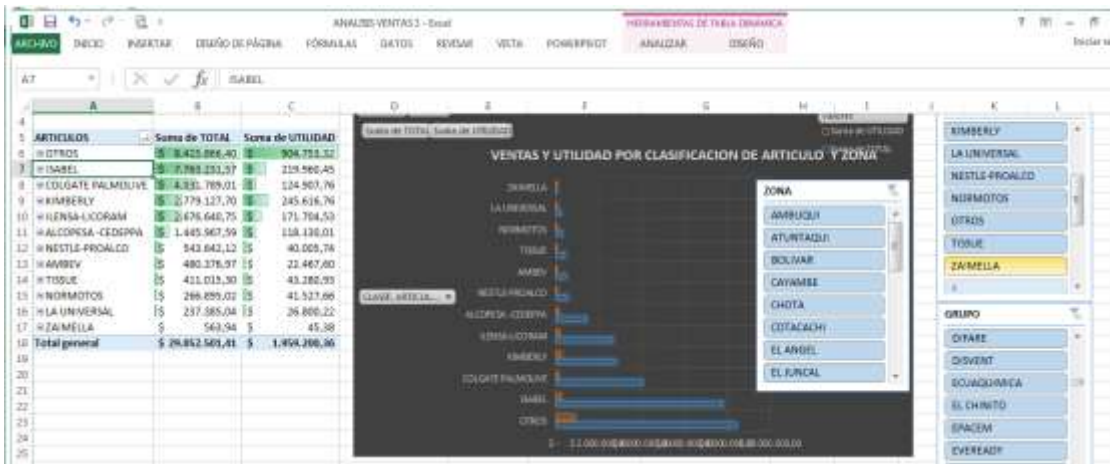


Ilustración 129: Reporte de Ventas y Utilidad por Artículo
Fuente: Propia



Ilustración 130: Ventas y Utilidad por año
Fuente: Propia

En el siguiente reporte dinámico, se analiza dos puntos de vista diferentes, como lo es las ventas por artículo y las ventas por clase de cliente, usando el mismo filtro de navegación que en este caso es el año:

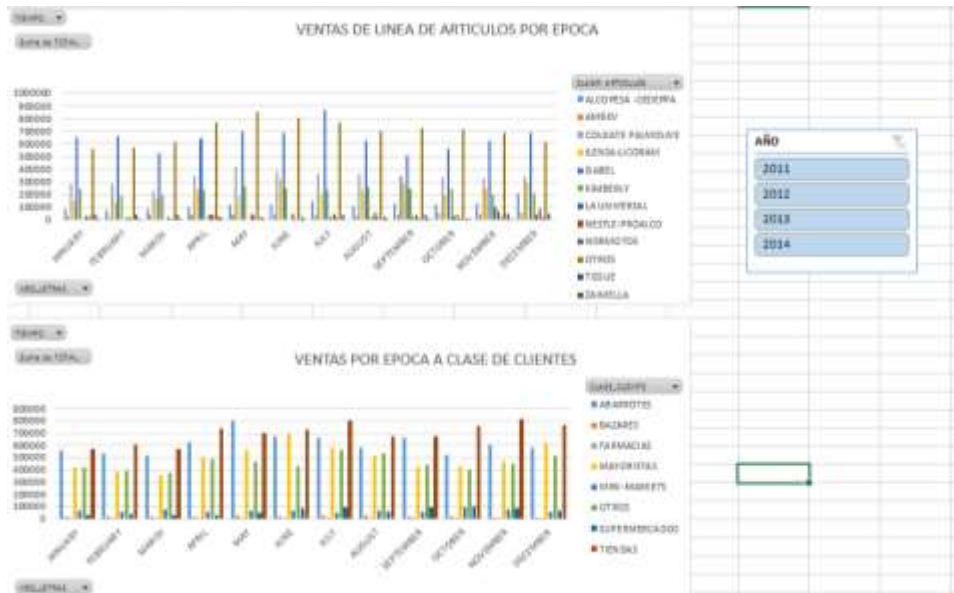


Ilustración 131: Reporte con el mismo filtro de dos puntos de vista
Fuente: Propia

En el siguiente reporte dinámico se puede analizar, las ventas por clase de cliente al año, usando la segmentación línea.



Ilustración 132: Ventas por Clase de Cliente al año
Fuente: Propia

Diseño de informes con Power View

Dentro del mismo ambiente de Excel se puede interactuar con la información usando el componente Power View, para lo cual como se mencionó antes, es necesario que esté instalado el componente Silverlight, para insertar una reporte con Power View, abrir Insertar y escoger la pestaña Power View, se abre un nuevo libro de Power View donde se puede encontrar varias herramientas disponibles para elaborar un reporte:



Ilustración 133: Ambiente de Excel para Power View
Fuente: Propia

En la parte derecha se encuentra el menú con la lista de campos, en donde aparecen las dimensiones con sus respectivos campos los cuales se los puede seleccionar, una vez seleccionado un campo se despliega en la parte superior del menú la pestaña **Diseñar** con opciones para trabajar con estos datos:



Ilustración 134: Menú de opciones de Power View

Fuente: Propia

Una gran ventaja en Power View es que se puede trabajar con secciones para elaborar reportes dinámicos, permitiendo que la información se visualice de manera muy dinámica y clara, para lo cual solo se debe dar clic en la parte vacía de la hoja y seguir arrastrando los campos para la nueva sección, a cada sección se le puede dar un formato independiente de las otras secciones haciendo uso del menú cambiar visualización.

Se ha elaborado el siguiente reporte para demostración, en el que se puede analizar la información, en tres secciones diferentes, por **ZONA** y por **CLIENTE** que se encuentran en formato de barras y por **LÍNEA** que se encuentra en formato de tabla tipo tarjeta, y también se puede analizar la información con el filtro de tiempo.

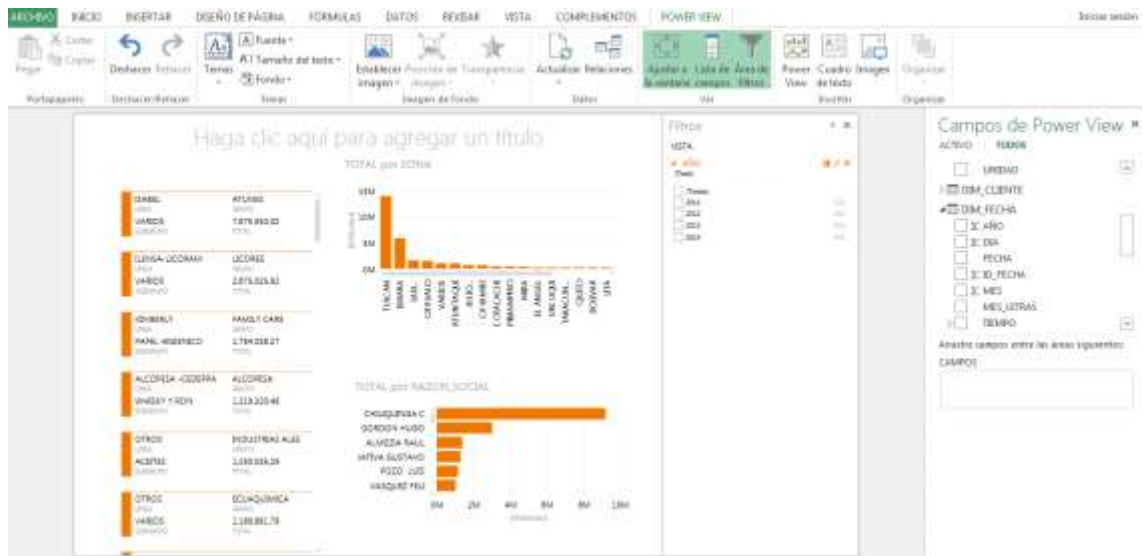


Ilustración 135: Informe con secciones en Power View
Fuente: Propia

En este reporte se puede también navegar por la información de manera específica, es decir si se da clic sobre una barra que la información se actualice, **por ejemplo** si se selecciona la barra de la zona Tulcán el reporte se actualiza con los datos correspondientes a esa Zona y se puede observar por ejemplo que: en Tulcán el mejor Cliente es el Sr. Chilingua y que el producto que más compra es de la línea ISABEL seguido de ILENSA.

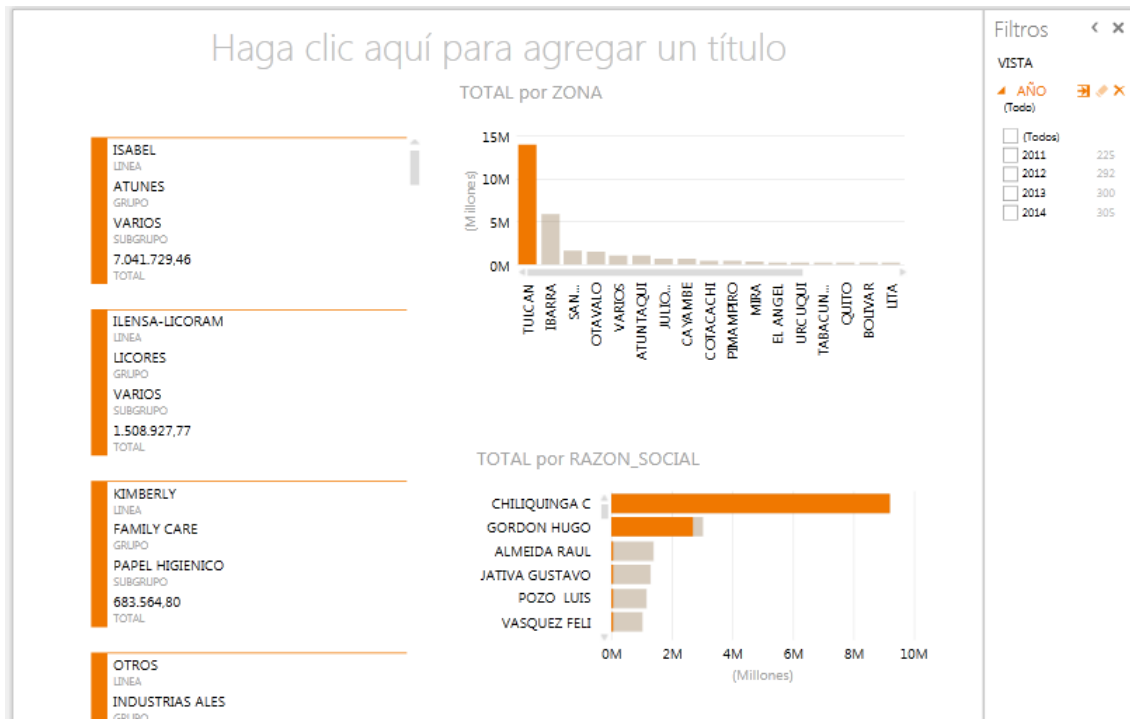


Ilustración 136: Informe interactivo de Ventas
Fuente Propia

En el siguiente reporte se hace uso del formato tipo grafico de dispersión para analizar el comportamiento de las ventas y la utilidad en una línea de tiempo, para lo cual se construye una nueva sección u hoja de Excel y se agrega los campos por los que se quiere evaluar la información en el menú de campos de la parte derecha, por ejemplo de la siguiente manera:

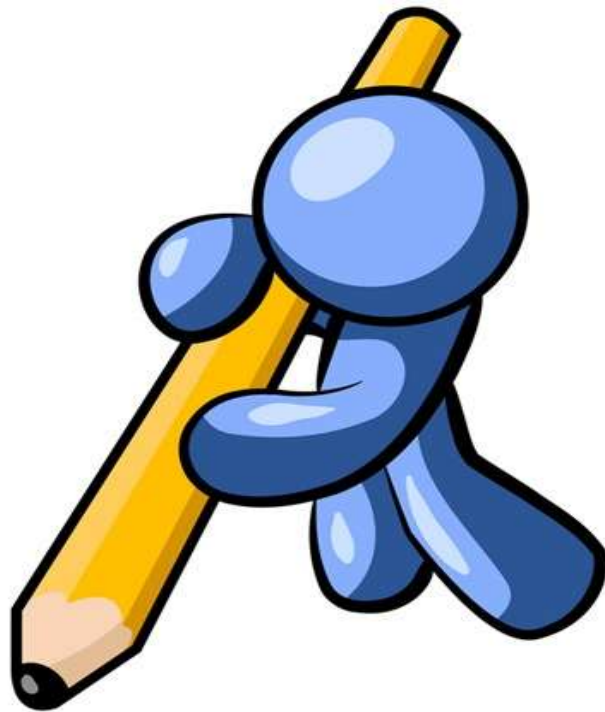


Ilustración 137: Informe interactivo con línea de tiempo
Fuente: Propia

De esta manera se puede analizar varios reportes más, en lo que respecta al departamento de ventas, cumpliendo así con los objetivos planteados en el presente proyecto.

CAPITULO 4

CONCLUSIONES Y RECOMENDACIONES



- ✓ CONCLUSIONES
- ✓ RECOMENDACIONES
- ✓ ANÁLISIS DE IMPACTOS

CONCLUSIONES

De la experiencia adquirida como resultado de la construcción del presente proyecto se puede concluir varios aspectos:

- La herramienta utilizada para el desarrollo del presente proyecto como se propuso en el objetivo, fue SQL Server 2014, la cual permitió la construcción adecuada del DataMart del área de Ventas, constatando así que es una herramienta eficiente, y que tiene muchos recursos para seguir explotando en el tema de Inteligencia de Negocios, como por ejemplo la minería de datos, entre otros.
- Las soluciones de Inteligencia de Negocios aplicadas a toda clase de empresas u organizaciones proveen mayor competitividad y rendimiento ya que impulsan de manera adecuada a las decisiones estratégicas, permitiendo al usuario profundizar en los datos y manejarlos desde diferentes perspectivas.
- El usar metodologías para el desarrollo y construcción de un modelo de datos para Inteligencia de negocios, es muy beneficioso, ya que aparte de ser una guía, permite que las actividades a realizar estén organizadas y planificadas ahorrando así recursos y tiempo.
- Para la implementación de un proyecto de Inteligencia de Negocios es muy importante seleccionar una herramienta eficiente, que disponga de la mayor cantidad de componentes, que permitan desarrollar una solución completa y reduzca al máximo el trabajo de los usuarios finales a la hora de consultar los reportes.
- En la actualidad existe en el mercado una gran variedad de herramientas para Inteligencia de negocios tanto en Software libre como con licencias, el escoger que herramienta usar dependerá de los requerimientos de la solución y el presupuesto disponible, en el presente proyecto se usó las herramientas que pone a disposición Microsoft, tanto para el diseño como para el análisis y presentación de datos.

RECOMENDACIONES

- En la actualidad es recomendable para las empresas que manejan gran cantidad de información implementar una solución de Inteligencia de Negocios, que le ayude a la gestión de la información y por ende al correcto funcionamiento de sus procesos, a través del conocimiento de información privilegiada que le guíen a la toma de decisiones adecuadas.
- Se recomienda que para realizar un proyecto de inteligencia de Negocios, se estudie y comprenda el funcionamiento y la estructura de la información fuente, así como también los requerimientos de la entidad, ya que de esta correcta comprensión depende el diseño de un modelo óptimo y eficaz, además de facilitar el desarrollo del proyecto y ahorrar tiempo en su construcción.
- Para la creación de un proyecto de inteligencia de Negocios se recomienda tomar en cuenta como requisito fundamental antes de iniciar con el proceso de construcción entender perfectamente y con claridad el funcionamiento del negocio, y sus procesos, ya que sin esta información clara la solución final no será totalmente eficiente.
- Se recomienda también tomar en cuenta a la hora de construir una solución de Inteligencia de negocios, que usando cualquier tipo de software para su desarrollo el costo siempre es un poco alto, mas sin embargo los resultados de usar estas soluciones generan muchos beneficios, justificando así la inversión.
- Se sugiere para otros estudiantes que como tema de tesis podrían profundizar en las herramientas de Microsoft incluidas en Excel, para análisis de datos y presentación de informes, partiendo de un cubo ya diseñado, explorando por ejemplo una aplicación Web de Power BI o Share Point ,que incluya aplicaciones de Power Pivot, Power View, Power Query, Power Map entre otros.

ANÁLISIS DE IMPACTOS

Para analizar los impactos que genera el uso de soluciones de Inteligencia de negocios de manera general en cualquier ámbito al que se lo aplique, se parte desde sus principales beneficios como lo son: el ahorro de tiempo y dinero, calidad en la información, mejora en los procesos, toma de decisiones estratégicas y eficientes, entre otras, que se resumen en el gran impacto que tiene la aplicación de dichos proyectos en el estado de resultados.

IMPACTO ECONÓMICO

En los últimos tiempos las empresas que manejan información han visto la necesidad de considerar inversiones tecnológicas como una necesidad para mantener su crecimiento económico y productividad. Es así que la inversión en soluciones de Inteligencia de Negocios se ha convertido en una de las inversiones más rentables, por todos los beneficios que brinda su aplicación.

Dentro de los beneficios más importantes de aplicar soluciones de inteligencia de negocios sin duda se encuentra el aspecto económico, ya que como resultado de su aplicación se logra optimizar los procesos empresariales para tener eficiencia y reducir costos.

En el presente proyecto se pudo observar un impacto por ejemplo a la hora de consultar reportes que antes necesitaban la intervención del proveedor del sistema transaccional, ahorrando tiempo y dinero, además se tuvo la posibilidad de explorar decisiones estratégicas que implican rentabilidad económica.

IMPACTO SOCIAL

La tecnología en la actualidad es aplicada en la mayoría de ambientes empresariales, y es muy importante en la sociedad por los beneficios que provee su aplicación puesto que permiten realizar varias actividades de manera más sencilla.

Es así que las soluciones de Inteligencia de Negocios se han convertido con el paso del tiempo en aplicaciones tecnológicas usadas por muchas empresas

como apoyo a los ejecutivos de alto nivel en la solución de problemas a los que se enfrentan y principalmente en la toma de decisiones; y se ha ido extendiendo poco a poco como apoyo a todas las personas que manejan información dentro de las empresas, eliminando los análisis extensos y las demoras en la obtención de información, aumentando la confianza en la toma de decisiones y la calidad de información.

En conclusión se puede decir que el impacto radica en el conocimiento y la información privilegiada que se da a conocer a las personas responsables del manejo de la información. En el presente proyecto por ejemplo se pudo conocer información histórica que no se había conocido antes, así como también se pudo realizar análisis de datos desde varios puntos de vista.

IMPACTO AMBIENTAL

En la actualidad es de suma importancia tomar conciencia acerca del cuidado que se debe dar al medio ambiente, por todas las consecuencias que conlleva su destrucción y que ya se están evidenciando y afectando a todo nuestro planeta, por lo cual entre una de los aportes al cuidado del medioambiente esta la reducción del consumo de papel por parte de todos, desde las personas como individuo hasta las grandes empresas que lo utilizan.

Al ser las empresas las mayores consumidoras de papel, tienen la responsabilidad y la obligación ética y moral de preservar el medio ambiente, realizando todos los procesos que les permita disminuir el daño ambiental. Con esta cooperación ambiental las empresas se benefician mejorando su imagen institucional.

Es por esta razón que el impacto ambiental que genera las soluciones de Inteligencia de Negocios en las empresas, es directamente en la reducción del uso del papel, puesto que a la hora de presentar los informes se reemplaza los informes impresos en papel, por informes digitales e interactivos mucho más óptimos. En el presente proyecto se presentan todos los análisis de datos en

varios reportes digitales que permiten al usuario final interactuar con la información.

BIBLIOGRAFÍA

Barbieri, C. (2011). *Business Intelligence*. Rio de Janeiro Brasil: Elsevier Editorial Ltda.

Cesar, P. L. (2014). *Técnicas de minería de datos e inteligencia de negocios : IBM SPSS Modeler*. Garceta Grupo Ed. 2014.

Diaz, L. E. (2014). *Inteligencia de Negocios: ¿la base para la solución de sus problemas? T21*.

IBM. (25 de 04 de 2015). *Aplique las nuevas herramientas analíticas para revelar nuevas oportunidades*. Obtenido de http://www.ibm.com/smarterplanet/ec/es/business_analytics/article/it_business_intelligence.html

MICROSOFT. (24 de 04 de 2015). *Libros en pantalla de SQL Server 2014*. Obtenido de <https://msdn.microsoft.com/es-es/library/ms130214.aspx>

Mistry, R., & Misner, S. (2012). *Introduccion Microsoft SQL Server 2012*. Madrid.

Muñiz, L. (2013). *Tablas dinamicas con excel aplicadas a la gestion empresarial para utilizar con excel 2007, 2010, 2013*. Barcelona: Profit Editorial.

UTN, D. d. (Marzo de 2011). *Guia para el diseño y presentacion de los trabajos de investigacion* .

Villarreal, R. X. (2013). "Estudio de metodologías de Data Warehouse para la implementación de repositorios de información para la toma de decisiones gerenciales.". *Tesis de Grado*. Universidad Tecnica del Norte, Ibarra Ecuador.

Roebuck, K. (2012). DATA WAREHOUSING. Emereo Publishing.

Cientificos,L. (2015). MODELADO PREDICTIVO PARA LA INTELIGENCIA DE NEGOCIOS. Createspace.

Kimball, R. (2013) THE DATA WAREHOUSE TOOLKIT (3RA Ed.). CANADA: Jhon Wiley & Sons INC.

GLOSARIO

A

Analisis Services

SQL Server Analysis Services (SSAS) ofrece funciones de procesamiento analítico en línea (OLAP) y minería de datos para aplicaciones de Business Intelligence. 53

B

Bottom-Up

Es una estrategia de procesamiento de información " de abajo arriba" 40

C

CRM

Customer Relationship Management, modelo de gestión enfocado principalmente en los clientes. 34

D

Dashboard y Scorecard

También llamados Cuadros de Mando Integral, son resúmenes visuales de información del negocio, que muestran de una mirada la comprensión del global de las condiciones del negocio mediante métricas e Indicadores Clave de Desempeño (KPIs) 32

DATAMART

Data mart es una versión especial de almacén de datos (data

warehouse). Constituye un subconjuntos de datos, con el propósito de ayudar a que un área específica dentro del negocio pueda tomar mejores decisiones. 20

Datamining

Minería de datos, es el conjunto de técnicas y tecnologías que permiten explorar grandes bases de datos, de manera automática o semiautomática, con el objetivo de encontrar patrones repetitivos, tendencias o reglas que expliquen el comportamiento de los datos en un determinado contexto. 32

DataWarehouse

Almacén de datos o colección de datos orientado a un determinado ámbito, integrado, no volátil y variable en el tiempo, que ayuda a la toma de decisiones 20

DAX

Data Analysis Expression, o Expresiones de Analisis de Datos es el lenguaje de fórmulas empleado para crear cálculos personalizados en PowerPivot para libros de Microsoft Excel y proyectos de modelos tabulares de Analysis Services. Las fórmulas DAX incluyen funciones, operadores y valores para realizar cálculos avanzados sobre datos de tablas y columnas. 58

E

ERP

son sistemas de información gerenciales que integran y manejan muchos de los negocios asociados con las operaciones de producción y de los aspectos de distribución de una compañía en la producción de bienes o servicios. 30

ETL

Extract, transform and load, o Extraer, transformar y cargar, es el proceso que permite mover datos desde múltiples fuentes, reformatearlos, limpiarlos y cargarlos en otra base de datos, data mart, o data warehouse para ser analizados. 21

H

HOLAP: Hybrid OLAP

Hybrid Online Analytical Process, procesamiento analítico en línea híbrido, es una combinación de ROLAP y MOLAP 45

I

Integration Services

Es una plataforma para la creación de soluciones empresariales, de transformaciones de datos e integración de datos. 53

M

Microsoft Visual Studio

Es un entorno de desarrollo integrado, para sistemas operativos Windows. Soporta múltiples

lenguajes de programación tales como C++, C#, Visual Basic .NET, F#, Java, Python, Ruby, PHP 52

MOLAP: Multidimensional OLAP

Multidimensional Online Analytical Processing, es decir, 'procesamiento analítico multidimensional en línea', su forma de acceso es sobre una base de datos multidimensional. 45

O

ODS

Es un almacén operacional de datos, que se alimenta de los sistemas transaccionales y al mismo tiempo les da soporte. 35

OLAP

On-Line Analytical Processing procesamiento analítico en línea, es una solución utilizada en el campo de Inteligencia de negocios con el objetivo de agilizar la consulta de grandes cantidades de datos a través del uso de estructuras multidimensionales. 21

P

Power Map

Es un complemento de Excel que permite fácilmente visualizar información georreferenciada (datos que incluyan latitud y longitud) en mapas 3D. Permite crear "paseos interactivos" de los datos permitiendo un mejor entendimiento de los mismos. 52

Power Pivot

Es una tecnología que permite procesar y analizar información compleja en poco tiempo, con poco esfuerzo y con poco hardware, su velocidad radica en que es tecnología in-memory. 23

Power View

es un componente para Microsoft Excel 2013 que brinda una experiencia interactiva de exploración, visualización y presentación de datos que fomenta la elaboración intuitiva de informes ad hoc. 52

R

Reporting Services

Proporciona una gama completa de herramientas y servicios listos para usar que ayudan a crear, implementar y administrar informes para la organización. Incluye características de programación que le permitirán ampliar y personalizar la funcionalidad de informes. 53

ROLAP: Relational OLAP

Significa Procesamiento Analítico OnLine Relacional, se trata de sistemas y herramientas OLAP (Procesamiento Analítico OnLine) construidos sobre una base de datos relacional. 45

S

SCM

Supply Chain Management, modelo de gestión o administración de la cadena de suministro. 34

Share Point

Es un entorno de colaboración que pueden usar organizaciones de todos los tamaños para incrementar la eficacia de los procesos empresariales, estos sitios proporcionan entornos seguros que los administradores pueden configurar para proporcionar un acceso personalizado a la información. 59

Silverlight

Es un complemento de Microsoft que permite desarrollar aplicaciones enriquecidas para la web. 59

SQL Server 2014

Es un sistema de manejo de bases de datos del modelo relacional, desarrollado por la empresa Microsoft 20

SQL Server Management Studio

(SSMS) es un entorno integrado para acceder, configurar, gestionar, administrar, y el desarrollo de todos los componentes de SQL Server. 23

Staging

Es un área temporal que se encuentra en el flujo de datos entre las fuentes y los repositorios para facilitar la extracción de los datos. 35

X

XVelocity

Tecnología de administración de
datos optimizada para memoria.58