

UNIVERSIDAD NACIONAL DE CAJAMARCA



FACULTAD DE INGENIERÍA

ESCUELA ACADÉMICO PROFESIONAL
DE INGENIERÍA DE SISTEMAS

“IMPACTO DE LA IMPLEMENTACIÓN DE UN DATA WAREHOUSE PARA OPTIMIZAR LA TOMA DE DECISIONES EN EL ÁREA DE COMERCIALIZACIÓN Y PRODUCCIÓN DE LA INDUSTRIA ALIMENTARIA HUACARIZ”

TESIS PARA OPTAR POR EL TÍTULO DE INGENIERO DE SISTEMAS

BACHILLER: TUCTO BASTERES ANTHONY DEIVY

ASESOR: ING. JAIME AMADOR MEZA HUAMÁN

CAJAMARCA - PERÚ

2017

AGRADECIMIENTO

En primer lugar a Dios por brindarme salud, por dejar que me rodee de buenas personas y permitirme llegar a este punto de mi vida.

Gracias mamá por acompañarme en mis alegrías y tristezas, por darme esas palabras de aliento que solo tú sabes decir, siempre aconsejándome para hacer lo correcto, por ser mi amiga y cómplice incondicional con la que siempre contaré.

Gracias papá por tu compañía y apoyo incondicional, por enseñarme que las cosas que empiezan bien, terminan bien, por ayudarme a forjar un carácter fuerte ante los problemas, por enseñarme a no rendirme y demostrar que sí se puede.

A mi asesor de tesis, el Ingeniero Jaime Meza, por brindarme su apoyo, su confianza y su tiempo desinteresado para recorrer este camino que ha sido la tesis.

A la industria alimentaria Huacariz por abrir sus puertas y brindarme todas las facilidades para hacer esta tesis.

DEDICATORIA

A Dios por acompañar siempre mis pasos.

A mis padres por brindarme su confianza y apoyo en mis decisiones desde
niño.

A mi hermana menor Shakira a quien quiero como una hija, y quien se
preocupa como una mamá por mí.

A mis hermanos mayores, quienes siempre me enseñan con sus experiencias.

A mis sobrinitos Nicolás y Sofía, para quienes espero ser un buen ejemplo.

A mis amigos Liliana y Renzo por acompañarme en este camino que ha sido la
tesis, brindándome su apoyo y buenos deseos todo este tiempo para no
detenerme y culminar con éxito.

A la señora Lidia por su apoyo incondicional todo este tiempo.

A mis tíos por sus consejos y buenos deseos para culminar esta tesis.

Para ellos es esta dedicatoria de tesis, por su apoyo incondicional.

CONTENIDO

AGRADECIMIENTO	ii
DEDICATORIA	iii
CONTENIDO	iv
ÍNDICE DE FIGURAS	vii
ÍNDICE DE TABLAS	x
RESUMEN	xii
ABSTRACT	xiii
CAPÍTULO I INTRODUCCIÓN	1
CAPÍTULO II MARCO TEÓRICO	5
2.1. Antecedentes teóricos de la investigación	5
2.2. Bases teóricas	6
2.2.1. Base de datos	6
2.2.2. Sistema de información	7
2.2.3. Data warehouse	9
2.2.3.1. Características de un data warehouse	9
2.2.3.2. Beneficios de un data warehouse	12
2.2.3.3. Diferencias entre un data warehouse y un data mart	12
2.2.4. Business intelligence (BI)	13
2.2.4.1. Arquitectura de una solución BI	14
2.2.5. Modelo dimensional	15
2.2.5.1. Procesos del modelado dimensional	15
2.2.5.2. Beneficios del modelado dimensional	16
2.2.6. Metodologías de diseño y construcción de un data warehouse	17
2.2.6.1. Metodología de Ralph Kimball	17
2.2.6.2. Metodología de Bill Inmon	23
2.2.6.3. Diferencias entre Ralph Kimball e Inmon	25
2.2.7. Toma de decisiones	26
2.2.8. Optimización	27
2.2.9. Información	27
2.2.10.1. Principales características de la información	27
CAPÍTULO III MATERIALES Y MÉTODOS	29
3.1. Procedimiento	29
3.1.1. Planeación del proyecto	29
3.1.1.1. Descripción de la empresa	29
3.1.1.2. Objetivos	30
3.1.1.3. Alcance	30
3.1.1.4. Riesgos	30
3.1.1.5. Recurso humano	31

3.1.1.6. Presupuesto y retorno de la inversión	31
3.1.2. Definición de requerimientos del negocio	34
3.1.2.1. Requerimientos del negocio	34
3.1.2.2. Entrevistas	37
3.1.2.3. Base de datos transaccional en SQL Server del área de comercialización	38
3.1.2.4. Diccionario de tablas del área de comercialización	43
3.1.2.5. Diccionario de datos del área de comercialización	44
3.1.2.6. Base de datos transaccional en SQL Server del área de producción	55
3.1.2.7. Diccionario de tablas del área de producción	58
3.1.2.8. Diccionario de datos del área de producción	58
3.1.3. Modelo dimensional	64
3.1.3.1. Elección del esquema a utilizar	64
3.1.3.2. Elección de las dimensiones para el área de comercialización	66
3.1.3.3. Elección de las dimensiones para el área de producción	67
3.1.3.4. Total de dimensiones encontradas	68
3.1.3.5. Medidas encontradas para el área de comercialización	68
3.1.3.6. Medidas encontradas para el área de producción	69
3.1.3.7. Diseño del modelo dimensional para el área de producción	69
3.1.3.8. Diseño del modelo dimensional para el área de comercialización	70
3.1.3.9. Modelo general del data warehouse	71
3.1.4. Diseño físico	72
3.1.4.1. Dimensión Dim_Producto	72
3.1.4.2. Dimensión Dim_Formula	72
3.1.4.3. Dimensión Dim_Venta	73
3.1.4.4. Dimensión Dim_Proveedor	73
3.1.4.5. Dimensión Dim_Sucursal	74
3.1.4.6. Dimensión Dim_Cliente	74
3.1.4.7. Dimensión Dim_Calendario	75
3.1.4.8. Dimensión Dim_Usuario	76
3.1.4.9. Dimensión Dim_Insumo	76
3.1.4.10. Dimensión Dim_Documento	77
3.1.4.11. Dimensión Dim_Orden	77
3.1.4.12. Tabla hechos Fact_Producción	77
3.1.4.13. Tabla hechos Fact_Comercialización	78
3.1.5. Microsoft SQL Server Integration Services (SSIS)	78
3.1.6. Diseño e implementación del subsistema de ETL en SSIS	79
3.1.5.1. Capa stage	79
3.1.5.2. Capa ODS (Operational Data Stores)	81
3.1.5.3. Capa DW (Data Warehouse)	107
3.1.5.4. Tipo de carga	112

3.1.7. QlikView	116
3.1.8. Desarrollo de la aplicación BI en QlikView	118
Aplicación BI para el área de comercialización.....	118
Aplicación BI para el área de Producción.....	126
3.2. Tratamiento, análisis de datos y presentación de resultados	129
3.2.1 .Tiempo de trabajo por cada reporte generado	129
3.2.2 .Grado de satisfacción en la presentación de reportes	130
CAPÍTULO IV ANÁLISIS Y DISCUSIÓN DE RESULTADOS	131
4.1. Análisis de resultados.....	131
4.1.1. Indicador razón: Tiempo de trabajo por cada reporte generado.....	131
a) Contrastación de la hipótesis.....	131
b) Formulación de hipótesis estadística	132
c) Nivel de significancia.....	133
d) Valor estadístico del procedimiento.....	133
e) Establecer región crítica.....	133
f) Toma de decisión	134
4.1.2. Indicador ordinal: Grado de satisfacción de los usuarios en cuanto al Data Warehouse.....	134
4.2. Discusión de resultados.....	136
CAPÍTULO V CONCLUSIONES Y RECOMENDACIONES.....	139
5.1. Conclusiones.....	139
5.2. Recomendaciones.....	140
REFERENCIAS BIBLIOGRÁFICAS	141
ANEXOS	144
ANEXO 1: MODELO DE ENCUESTA – GRADO DE SATISFACCIÓN	144

ÍNDICE DE FIGURAS

FIGURA 1 BASE DE DATOS [10].....	7
FIGURA 2 SISTEMA DE INFORMACIÓN [12]	8
FIGURA 3 DATA WAREHOUSE [15].....	9
FIGURA 4 DATA WAREHOUSE – INTEGRADO [16].....	10
FIGURA 5 DATA WAREHOUSE – TEMÁTICO [16].....	10
FIGURA 6 DATA WAREHOUSE – HISTÓRICO [16]	11
FIGURA 7 DATA WAREHOUSE - NO VOLÁTIL [16].....	11
FIGURA 8 DATOS - INFORMACIÓN - CONOCIMIENTO [18]	13
FIGURA 9 ARQUITECTURA DE UNA SOLUCIÓN BI [19].....	14
FIGURA 10 DIAGRAMA DE LA METODOLOGÍA DE RALPH KIMBALL [21]	18
FIGURA 11 ARQUITECTURA DE DATA WAREHOUSE - BILL INMON [3].....	23
FIGURA 12 LOGO - INDUSTRIAS ALIMENTARIAS HUACARIZ.....	29
FIGURA 13 BASE DE DATOS - COMERCIALIZACIÓN	38
FIGURA 14 TABLA - TB_PRODUCTOSTERMINADOS	39
FIGURA 15 TABLA - TB_DETALLEVENTA.....	39
FIGURA 16 TABLA - TB_TIPOCLIENTE.....	39
FIGURA 17 TABLA - TB_SUCURSALLUGAR.....	39
FIGURA 18 TABLA - TB_USUARIO.....	40
FIGURA 19 TABLA - TB_TIPODOCUMENTO	40
FIGURA 20 TABLA - TB_CLIENTE	40
FIGURA 21 TABLA - TB_VENTAS	41
FIGURA 22 TABLA - TB_PROVEEDORES	41
FIGURA 23 TABLA - TB_CALEAR	42
FIGURA 24 TABLA - TB_CATEGORIA	42
FIGURA 25 TABLA - UBIDEPARTAMENTO	42
FIGURA 26 TABLA - UBIPROVINCIA.....	43
FIGURA 27 TABLA - UBIDISTRITO.....	43
FIGURA 28 BASE DE DATOS - PRODUCCIÓN.....	55
FIGURA 29 TABLA - TB_FORMULAS.....	55
FIGURA 30 TABLA - TB_ORDENPRODUCCION	55
FIGURA 31 TABLA - TB_USUARIO.....	56
FIGURA 32 TABLA - TB_DETALLEORDENPRODUCCION.....	56
FIGURA 33 TABLA - TB_INSUMOS	56
FIGURA 34 TABLA - TB_CATEGORIA	57
FIGURA 35 TABLA - TB_CALEAR	57
FIGURA 36 ESQUEMA ESTRELLA [27]	64
FIGURA 37 ESQUEMA DE COPO DE NIEVE [30].....	65
FIGURA 38 MODELO DIMENSIONAL PARA EL ÁREA DE PRODUCCIÓN	69
FIGURA 39 MODELO DIMENSIONAL PARA EL ÁREA DE COMERCIALIZACIÓN	70
FIGURA 40 MODELO GENERAL DEL DATA WAREHOUSE	71
FIGURA 41 MICROSOFT SQL SERVER INTEGRATION SERVICES (SSIS) [33].....	79
FIGURA 42 CAPA STAGE.....	80
FIGURA 43 FLUJO DE DATOS STG_TB_USUARIO	81
FIGURA 44 CAPA ODS (OPERATIONAL DATA STORE)	82
FIGURA 45 ODS_DIM_VENTA - ORIGEN STG_TB_VENTAS.....	82
FIGURA 46 ODS_DIM_VENTA - FLUJO DE DATOS	83
FIGURA 47 ODS_DIM_VENTA - DESTINO ODS_DIM_VENTA.....	83
FIGURA 48 ODS_DIM_CALEAR - ORIGEN STG_TB_CALEAR.....	84
FIGURA 49 ODS_DIM_CALEAR - FLUJO DE DATOS.....	84
FIGURA 50 ODS_DIM_CALEAR - DESTINO ODS_DIM_CALEAR.....	85
FIGURA 51 ODS_DIM_CLIENTE - ORIGEN STG_TB_CLIENTE	85
FIGURA 52 ODS_DIM_CLIENTE - ORIGEN STG_TB_TIPOCLIENTE	86
FIGURA 53 ODS_DIM_CLIENTE - ORIGEN STG_UBIDEPARTAMENTO.....	86
FIGURA 54 ODS_DIM_CLIENTE - ORIGEN STG_UBIPROVINCIA	87

FIGURA 55 ODS_DIM_CLIENTE - ORIGEN STG_UBIDISTRITO	87
FIGURA 56 ODS_DIM_CLIENTE - FLUJO DE DATOS.....	88
FIGURA 57 ODS_DIM_CLIENTE - DESTINO ODS_DIM_CLIENTE.....	89
FIGURA 58 ODS_DIM_DOCUMENTO - ORIGEN STG_TB_TIPODOCUMENTO.....	89
FIGURA 59 ODS_DIM_DOCUMENTO - FLUJO DE DATOS	90
FIGURA 60 ODS_DIM_DOCUMENTO - DESTINO ODS_DIM_DOCUMENTO	90
FIGURA 61 ODS_DIM_PRODUCTO - ORIGEN STG_TB_PRODUCTOSTERMINADOS	91
FIGURA 62 ODS_DIM_PRODUCTO - ORIGEN STG_TB_CATEGORIA.....	91
FIGURA 63 ODS_DIM_PRODUCTO - FLUJO DE DATOS.....	92
FIGURA 64 ODS_DIM_PRODUCTO - DESTINO ODS_DIM_PRODUCTO	92
FIGURA 65 ODS_DIM_PROVEEDOR - ORIGEN STG_TB_PROVEEDORES.....	93
FIGURA 66 ODS_DIM_PROVEEDOR - FLUJO DE DATOS	93
FIGURA 67 ODS_DIM_PROVEEDOR - DESTINO ODS_DIM_PROVEEDOR.....	94
FIGURA 68 ODS_DIM_SUCURSAL - ORIGEN STG_TB_SUCURSALLUGAR	94
FIGURA 69 ODS_DIM_SUCURSAL - FLUJO DE DATOS.....	95
FIGURA 70 ODS_DIM_SUCURSAL - DESTINO ODS_DIM_SUCURSAL	95
FIGURA 71 ODS_DIM_USUARIO - ORIGEN STG_TB_USUARIO.....	96
FIGURA 72 ODS_DIM_USUARIO - ORIGEN STG_TB_SUCURSALLUGAR.....	96
FIGURA 73 ODS_DIM_USUARIO - FLUJO DE DATOS	97
FIGURA 74 ODS_DIM_USUARIO - DESTINO ODS_DIM_USUARIO	97
FIGURA 75 ODS_DIM_FORMULA - ORIGEN STG_TB_FORMULAS	98
FIGURA 76 ODS_DIM_FORMULA - FLUJO DE DATOS	98
FIGURA 77 ODS_DIM_FORMULA - DESTINO ODS_DIM_FORMULA	99
FIGURA 78 ODS_DIM_ORDEN - ORIGEN STG_TB_ORDENPRODUCCION	99
FIGURA 79 ODS_DIM_ORDEN - FLUJO DE DATOS	100
FIGURA 80 ODS_DIM_ORDEN - DESTINO ODS_DIM_ORDEN.....	100
FIGURA 81 ODS_DIM_INSUMO - ORIGEN STG_TB_INSUMOS.....	101
FIGURA 82 ODS_DIM_INSUMO - ORIGEN STG_TB_CATEGORIA.....	101
FIGURA 83 ODS_DIM_INSUMO - FLUJO DE DATOS	102
FIGURA 84 ODS_DIM_INSUMO - DESTINO ODS_DIM_INSUMO.....	102
FIGURA 85 ODS_FACT_COMERCIALIZACION - ORIGEN STG_TB_VENTAS	103
FIGURA 86 ODS_FACT_COMERCIALIZACION - ORIGEN STG_TB_DETALLEVENTA	103
FIGURA 87 ODS_FACT_COMERCIALIZACION - ORIGEN STG_TB_PRODUCTOSTERMINADOS.....	104
FIGURA 88 ODS_FACT_COMERCIALIZACION - FLUJO DE DATOS.....	104
FIGURA 89 ODS_FACT_COMERCIALIZACION - DESTINO ODS_FACT_COMERCIALIZACION	105
FIGURA 90 ODS_FACT_PRODUCCION - ORIGEN STG_TB_ORDENPRODUCCION	105
FIGURA 91 ODS_FACT_PRODUCCION - ORIGEN STG_TB_DETALLEORDENPRODUCCION	106
FIGURA 92 ODS_FACT_PRODUCCION - FLUJO DE DATOS	106
FIGURA 93 ODS_FACT_PRODUCCION - DESTINO ODS_FACT_PRODUCCION.....	107
FIGURA 94 CAPA DW (DATA WAREHOUSE)	107
FIGURA 95 FLUJO DE DATOS DW_DIM_CLIENTE	108
FIGURA 96 DW_FACT_COMERCIALIZACION - ORIGEN ODS_FACT_COMERCIALIZACION	108
FIGURA 97 DW_FACT_COMERCIALIZACION - FLUJO DE DATOS.....	109
FIGURA 98 DW_FACT_COMERCIALIZACION - DESTINO DW_FACT_COMERCIALIZACION	110
FIGURA 99 DW_FACT_COMERCIALIZACION - ORIGEN ODS_FACT_COMERCIALIZACION	110
FIGURA 100 DW_FACT_COMERCIALIZACION - FLUJO DE DATOS	111
FIGURA 101 DW_FACT_COMERCIALIZACION - DESTINO DW_FACT_COMERCIALIZACION	112
FIGURA 102 EJECUCIÓN PACKAGE MAESTRO	112
FIGURA 103 CREAR UNA CREDENCIAL	113
FIGURA 104 CREAR UN JOB EN EL AGENTE SQL	113

FIGURA 105 CREAR LOS PASOS DEL JOB.....	114
FIGURA 106 CONFIGURAR EL PASO PARA EL JOB.....	114
FIGURA 107 CREAR LA PROGRAMACIÓN DE EJECUCIÓN PARA EL JOB.....	115
FIGURA 108 CONFIGURACIÓN DE LA PROGRAMACIÓN DEL JOB.....	115
FIGURA 109 DESCUBRIR DATOS VS DESCUBRIR EL NEGOCIO [36]	117
FIGURA 110 DESCUBRIR DATOS VS DESCUBRIR EL NEGOCIO (2) [36]	118
FIGURA 111 MODELO DIMENSIONAL DE COMERCIALIZACIÓN EN QLIKVIEW	118
FIGURA 112 VENTAS Y META MENSUAL – VENTAS GENERALES	119
FIGURA 113 VENTAS Y META MENSUAL – POR SUCURSAL	119
FIGURA 114 VENTAS Y META MENSUAL – POR PRODUCTO	120
FIGURA 115 VENTAS POR DOCUMENTO – VENTAS GENERALES.....	120
FIGURA 116 VENTAS POR DOCUMENTO – POR PRODUCTO	121
FIGURA 117 VENTAS POR DOCUMENTO – POR CLIENTE	121
FIGURA 118 VENTAS POR PRECIO – VENTAS GENERALES	122
FIGURA 119 VENTAS POR PRECIO – POR PRODUCTO	122
FIGURA 120 VENTAS POR PRECIO – POR TIPO DE VENTA.....	123
FIGURA 121 PRODUCTOS CON MÁS DEMANDA - GENERAL.....	123
FIGURA 122 PRODUCTOS CON MÁS DEMANDA - POR PRODUCTO	124
FIGURA 123 PRODUCTOS CON MÁS DEMANDA - POR CLIENTE	124
FIGURA 124 CUENTAS POR COBRAR – POR AÑOS.....	125
FIGURA 125 CUENTAS POR COBRAR – EN MESES POR AÑO	125
FIGURA 126 CUENTAS POR COBRAR - DETALLE POR CLIENTE	126
FIGURA 127 MODELO DIMENSIONAL DE PRODUCCIÓN EN QLIKVIEW.....	126
FIGURA 128 CONSUMO Y PROYECCIÓN DE LECHE	127
FIGURA 129 CONSUMO Y PROYECCIÓN DE INSUMO	127
FIGURA 130 PRODUCTOS MÁS PRODUCIDOS	128
FIGURA 131 GRÁFICO DE MEDICIÓN DE TIEMPOS.....	129
FIGURA 132 DISTRIBUCIÓN T-STUDENT	134
FIGURA 133 GRÁFICO COMPARATIVO DEL GRADO DE SATISFACCIÓN	135
FIGURA 134 GRADO DE SATISFACCIÓN	137
FIGURA 135 GRÁFICO DE MEDICIÓN DE TIEMPOS.....	137

ÍNDICE DE TABLAS

TABLA 1 DIFERENCIAS ENTRE UN DATA WAREHOUSE Y UN DATA MART [17].....	12
TABLA 2 DIFERENCIAS ENTRE RALPH KIMBALL E INMON [3].....	25
TABLA 3 RECURSO HUMANO	31
TABLA 4 PRESUPUESTO – ACTIVOS TANGIBLES.....	31
TABLA 5 PRESUPUESTO - ACTIVOS INTANGIBLES	31
TABLA 6 PRESUPUESTO - OTROS GASTOS	31
TABLA 7 PRESUPUESTO - GASTOS PERSONAL	31
TABLA 8 PRESUPUESTO - COSTOS PROYECTADOS	32
TABLA 9 COMPARACIÓN DE GASTOS ANUALES POR HORAS TRABAJADAS.....	32
TABLA 10 BENEFICIO	32
TABLA 11 INGRESOS PROYECTADOS	32
TABLA 12 DIFERENCIA ENTRE INGRESOS Y COSTOS PROYECTADOS.....	33
TABLA 13 CÁLCULO DEL VALOR COK.....	33
TABLA 14 INDICADORES DE EVALUACIÓN.....	33
TABLA 15 CUADRO DE REQUERIMIENTO NO FUNCIONAL N°1	34
TABLA 16 CUADRO DE REQUERIMIENTO NO FUNCIONAL N°2	34
TABLA 17 CUADRO DE REQUERIMIENTO NO FUNCIONAL N°3	34
TABLA 18 CUADRO DE REQUERIMIENTO FUNCIONAL N°4	35
TABLA 19 CUADRO DE REQUERIMIENTO FUNCIONAL N°5	35
TABLA 20 CUADRO DE REQUERIMIENTO FUNCIONAL N°6	35
TABLA 21 CUADRO DE REQUERIMIENTO FUNCIONAL N°7	35
TABLA 22 CUADRO DE REQUERIMIENTO FUNCIONAL N°8	35
TABLA 23 CUADRO DE REQUERIMIENTO FUNCIONAL N°9	35
TABLA 24 CUADRO DE REQUERIMIENTO FUNCIONAL N°10	36
TABLA 25 CUADRO DE REQUERIMIENTO FUNCIONAL N°11	36
TABLA 26 CUADRO DE REQUERIMIENTO FUNCIONAL N°12	36
TABLA 27 CUADRO DE REQUERIMIENTO FUNCIONAL N°13	36
TABLA 28 CUADRO DE REQUERIMIENTO FUNCIONAL N°14	36
TABLA 29 TABLAS DEL ÁREA DE COMERCIALIZACIÓN	43
TABLA 30 DESCRIPCIÓN TB_PRODUCTOSTERMINADOS.....	44
TABLA 31 DESCRIPCIÓN TB_DETALLEVENTA.....	45
TABLA 32 DESCRIPCIÓN TB_TIPOCLIENTE.....	46
TABLA 33 DESCRIPCIÓN TB_SUCURSALLUGAR	46
TABLA 34 DESCRIPCIÓN UBIDEPARTAMENTO	46
TABLA 35 DESCRIPCIÓN TB_USUARIO.....	47
TABLA 36 DESCRIPCIÓN TB_CLIENTE	48
TABLA 37 DESCRIPCIÓN TB_VENTAS.....	49
TABLA 38 DESCRIPCIÓN TB_PROVEEDORES	51
TABLA 39 DESCRIPCIÓN TB_CALENDAR	52
TABLA 40 DESCRIPCIÓN TB_CATEGORIA	53
TABLA 41 DESCRIPCIÓN UBIPROVINCIA	53
TABLA 42 DESCRIPCIÓN UBIDISTRITO.....	54
TABLA 43 DESCRIPCIÓN TB_TIPODOCUMENTO	54
TABLA 44 TABLAS DEL ÁREA DE PRODUCCIÓN.....	58
TABLA 45 DESCRIPCIÓN TB_FORMULAS.....	58
TABLA 46 DESCRIPCIÓN TB_USUARIO.....	59
TABLA 47 DESCRIPCIÓN TB_CALENDAR	60
TABLA 48 DESCRIPCIÓN TB_CATEGORIA	61
TABLA 49 DESCRIPCIÓN TB_DETALLEORDENPRODUCCION.....	61
TABLA 50 DESCRIPCIÓN TB_INSUMOS	62
TABLA 51 DESCRIPCIÓN TB_ORDENPRODUCCION	63
TABLA 52 ELECCIÓN DE DIMENSIONES PARA EL ÁREA DE COMERCIALIZACIÓN.....	67
TABLA 53 ELECCIÓN DE DIMENSIONES PARA EL ÁREA DE PRODUCCIÓN	68
TABLA 54 DISEÑO FÍSICO DE DIM_PRODUCTO	72

TABLA 55 DISEÑO FÍSICO DE DIM_FORMULA	72
TABLA 56 DISEÑO FÍSICO DE DIM_VENTA.....	73
TABLA 57 DISEÑO FÍSICO DE DIM_PROVEEDOR	73
TABLA 58 DISEÑO FÍSICO DE DIM_SUCURSAL	74
TABLA 59 DISEÑO FÍSICO DE DIM_CLIENTE	74
TABLA 60 DISEÑO FÍSICO DE DIM_CALENDARIO	75
TABLA 61 DISEÑO FÍSICO DE DIM_USUARIO.....	76
TABLA 62 DISEÑO FÍSICO DE DIM_INSUMO.....	76
TABLA 63 DISEÑO FÍSICO DE DIM_DOCUMENTO	77
TABLA 64 DISEÑO FÍSICO DE DIM_ORDEN	77
TABLA 65 DISEÑO FÍSICO DE FACT_PRODUCION.....	77
TABLA 66 DISEÑO FÍSICO DE FACT_COMERCIALIZACION	78
TABLA 67 MEDICIÓN DE TIEMPO - PRE TEST Y POST TEST	129
TABLA 68 GRADO DE SATISFACCIÓN - PRE TEST	130
TABLA 69 RESUMEN DE MEDICIÓN DE TIEMPO - PRE TEST Y POST TEST	132
TABLA 70 GRADO DE SATISFACCIÓN - PRE TEST	135
TABLA 71 GRADO DE SATISFACCIÓN - POST TEST	135

RESUMEN

El presente proyecto de investigación se desarrolló con el objetivo de realizar el seguimiento a las ventas en el área de comercialización y de los productos producidos e insumos involucrados en el área de producción con la finalidad de tener un impacto positivo en la toma de decisiones en las áreas mencionadas de la industria alimentaria Huacariz con la implementación de un data warehouse bajo la metodología de Ralph Kimball. Se analizó el estado de las áreas de comercialización y producción al inicio del proyecto, logrando así, determinar que cuentan con un sistema informático proporcionado por el área de sistemas, el cual presenta ciertas limitaciones en cuanto a la información disponible, causando que los usuarios tengan que recurrir al área de sistemas para solicitar reportes adicionales con información histórica y más detallada, aspecto que les demanda invertir tiempo extra para el procesamiento manual y análisis de la información obtenida, teniendo impacto en la toma de decisiones.

Lo primero que se realizó fue la identificación y entendimiento de la base de datos de la industria alimentaria Huacariz, específicamente de las tablas involucradas en los procesos de las áreas de comercialización y producción. Se identificó los requerimientos de los usuarios de cada área siguiendo los pasos de la metodología de Ralph Kimball, para luego poder extraer la información necesaria para el proyecto, logrando así, diseñar los modelos dimensionales necesarios, crear los flujos de carga en SQL Server Integration Services y consumir la información por medio de los reportes creados en QlikView.

Finalmente se recopiló información a través de encuestas para obtener resultados del pre-test y del post-test una vez presentados los reportes y se validó la hipótesis con la prueba T-Student, obteniendo que la implementación de un data warehouse tuvo un impacto positivo en la toma de decisiones en el área de comercialización y producción de la industria alimentaria Huacariz disminuyendo los tiempos de respuesta y aumentando el grado de satisfacción en las áreas correspondientes.

Palabras clave: Data warehouse, inteligencia de negocios, toma de decisiones, industria alimentaria, comercialización y producción.

ABSTRACT

The present research project was developed with the objective of monitoring the sales in the area of marketing and the products produced and inputs involved in the production area in order to have a positive impact on the decision making in the mentioned areas of the Huacariz food industry with the implementation of a data warehouse under the methodology of Ralph Kimball. The status of the marketing and production areas was analyzed at the beginning of the project, thus determining that they have a computer system provided by the systems area, which presents certain limitations as to the available information, causing users to have they resort to the systems area to request additional reports with historical information and more detailed, which requires them to invest extra time for manual processing and analysis of the information obtained, having an impact on decision making.

The first thing that was done was the identification and understanding of the logic of the Huacariz food industry database, specifically the tables involved in the processes of the marketing and production areas. The requirements of the users of each area were identified following the steps of the Ralph Kimball methodology, in order to extract the necessary information for the project, thus, designing the necessary dimensional models, creating the load flows in SQL Server Integration Services and consume the information through the reports created in QlikView.

Finally, information was collected through surveys to obtain pre-test and post-test results once the reports were presented and the hypothesis was validated with the T-Student test, obtaining that the implementation of a data warehouse had a positive impact on the decision making in the marketing and production area of the Huacariz food industry, reducing response times and increasing the degree of satisfaction in the corresponding areas.

Keywords: Data warehouse, business intelligence, decision making, food industry, marketing and production.

CAPÍTULO I INTRODUCCIÓN

A nivel mundial ha crecido la capacidad para generar y almacenar información, la cual no puede ser analizada por los métodos tradicionales existentes, mientras mayor es la capacidad para almacenar datos, mayor es la incapacidad para extraer información realmente útil [1], en este contexto la información cobra una gran importancia estratégica en la competencia empresarial en donde todo está marcado por la información, el conocimiento del mercado y la toma de decisiones [2].

A nivel nacional son muchas las entidades públicas y privadas que han empezado a confiar en las soluciones de business intelligence para la toma de decisiones, las cuales son críticas para una mejora continua y una buena administración de sus recursos, la Contraloría General de la República es un ejemplo de ello, en donde se implementó un data mart para optimizar la toma de decisiones en el departamento de finanzas [3]; el ámbito local no es la excepción, y es que en la Municipalidad Distrital de los Baños del Inca se desarrolló un data mart en el área de administración y finanzas permitiéndoles una mejor administración y gestión de la información logrando mejorar las decisiones estratégicas basadas en información de calidad [4].

Cada vez que se habla de queso en el Perú se habla de Cajamarca, eso es un orgullo para la región, la cual brinda un producto de calidad con gran aceptación en el mercado nacional que crece cada día [5]. En la actualidad, en la mayoría de empresas productoras de lácteos de Cajamarca no se cuenta con sistemas de información que permitan analizar con exactitud la información de los procesos de producción y comercialización de productos. La demanda dentro de la región Cajamarca es cambiante debido a la temporada turística, lo que hace que existan algunos meses de mayor demanda de productos lácteos, sumado al envío de productos a todo el Perú. En estas empresas no se cuenta con una información adecuada para la toma de decisiones en la compra de materias primas e insumos que permitan una mejor distribución y producción de productos, lo que conlleva a que las empresas se queden sin stocks y en caso contrario los productos se malogren o se venzan ocasionando pérdidas económicas.

Actualmente, en las áreas de comercialización y producción de la industria alimentaria Huacariz los usuarios invierten mucho tiempo en la elaboración de reportes como tablas dinámicas para la toma de decisiones, si bien el sistema genera reportes en hojas de cálculo o archivos pdf, hay ocasiones en las que se necesita información adicional, la cual tiene que ser solicitada al área de sistemas, lo cual genera dependencia y un tiempo de espera hasta la creación de las consultas y su emisión en archivos excel, una vez que los usuarios tienen la información requerida del sistema aún tienen que depurar, ordenar, organizar y clasificar estos datos con el fin de obtener los reportes que quieren para realizar sus labores en la toma de decisiones. Este proceso de depurar, ordenar, organizar y clasificar los datos lo realiza de manera individual cada usuario por lo que los criterios que utilizan puede variar dando como resultado información no necesariamente confiable; adicional a ello, al hacerlo de forma manual se invierte mucho tiempo, y al tratarse de una gran cantidad de datos el usuario esta propenso a cometer errores. Al realizarse de manera individual los usuarios obtienen hojas de cálculo como resultado del procedimiento que desarrollan para obtener sus reportes, de modo que no es posible contar con información histórica durante los análisis. Todo lo mencionado genera como consecuencia que la gerencia de esta empresa no cuente con información exacta de cuánto invertir, qué comprar, ni qué cantidades producir; esto no permite una buena calidad y rentabilidad en el negocio.

Después de describir el contexto y el problema podemos cuestionarnos, ¿De qué manera la implementación de un data warehouse optimizaría la toma de decisiones en el área de comercialización y producción de la industria alimentaria Huacariz? La hipótesis que se plantea es que al implementar un data warehouse en el área de comercialización y producción de la industria alimentaria Huacariz se optimizará la toma de decisiones. Además, en las hipótesis específicas se considera que con la implementación del data warehouse, mediante la programación de tareas, se evitara los procesos de solicitar datos al área de sistemas ya que se dispondrá de manera inmediata con reportes con la información necesaria, permitiendo analizar información más veraz y oportuna lo que dará paso a elecciones más certeras al momento de tomar decisiones en las áreas de comercialización y producción.

Con lo expuesto anteriormente se sostiene que implementar un data warehouse en la industria alimentaria Huacariz, ayudaría a explotar de una mejor manera su información, con la posibilidad de poder manipularlos de una manera más sencilla y entender el porqué de su desempeño del área de comercialización y producción o, mejor aún, plantear escenarios a futuro, lo cual ayudaría a tomar mejores decisiones [2] en la compra de materias primas e insumos, logrando así, tener un stock que vaya acorde a sus ventas, esto permitiría reducir el número de productos vencidos en el almacén.

La reducción de tiempo en la elaboración de reportes para la toma de decisiones en las áreas de comercialización y producción sería notoria, pues ya no invertirían tiempo en depurar, ordenar y clasificar los datos, además de ello, contarían con información más confiable pues la manipulación de datos sería automatizada a través del data warehouse llegando a tener un mejor panorama sobre sus ventas y su producción lo que le permitirá aumentar su calidad y rentabilidad.

El alcance de esta solución abarca las áreas de comercialización y producción, en donde se hará uso de toda la data histórica existente en la base de datos actual para la creación de reportes dando como resultado modelos dimensionales que serán la fuente de datos para los reportes que muestren información para la toma de decisiones.

Las limitaciones que se encuentran en la investigación son el uso de computadoras existentes cuyo hardware podría no satisfacer los requerimientos del software pues la herramienta a utilizar para la creación de reportes trabaja a nivel de memoria ram, y el consumo de la misma es equivalente a la cantidad de datos a procesar, además de ello los usuarios funcionales podrían no tener tiempo disponible al inicio de la investigación para un análisis de requerimientos a corto tiempo.

El objetivo general es evaluar de qué manera la implementación de un data warehouse optimizaría la toma de decisiones en el área de comercialización y producción de la industria alimentaria Huacariz. Como objetivos específicos se tiene, identificar y analizar los requerimientos de información que intervendrán en los modelos dimensionales que se construirá e implementará para el área de comercialización y producción, lo que servirá en la creación de reportes para la toma de decisiones, permitiendo así, determinar el impacto analizando el nivel

de optimización en la toma de decisiones en el área de comercialización y producción cuando se implemente el data warehouse.

El informe de tesis está estructurado y desarrollado en 5 capítulos que a continuación se detallan:

Capítulo I, en este capítulo se presenta el problema, situación actual de la industria alimentaria Huacariz, la hipótesis, la justificación, los alcances y objetivo principal de la Investigación.

Capítulo II, contiene el marco teórico (antecedentes históricos, bases teóricas y definición de términos básicos) relacionado con los requerimientos y la parte funcional de la implementación de un data warehouse.

Capítulo III, corresponde a los materiales y métodos que se utilizaron para desarrollar el proyecto de investigación, donde se detalla y se explica la forma en que se obtuvieron los datos para luego transformarlos en información, además de la generación, análisis y presentación de reportes.

Capítulo IV, contiene el análisis de datos que se realizó y se explica los resultados obtenidos en la investigación.

Capítulo V, contiene las conclusiones de acuerdo a los objetivos planteados y también las recomendaciones necesarias para seguir ampliando los conocimientos sobre el tema de investigación que no se incluyeron en el presente informe de tesis.

CAPÍTULO II MARCO TEÓRICO

2.1. Antecedentes teóricos de la investigación

Fiorelly Shirley Guillén Rodríguez en su tesis desarrollada en el 2012 para optar por el título profesional de Ingeniero de Sistemas, desarrollo de un data mart para mejorar la toma de decisiones en el área de tesorería de la Municipalidad Provincial de Cajamarca; el trabajo consistió en desarrollar un data mart para hacer un uso adecuado de lo recaudado en las temporadas más críticas según cada área involucrada, para lo que se diseñó una base de datos dimensional y se creó un cubo de acuerdo a la tabla hechos; como conclusión se logró mejorar, crear reportes anuales y mensuales para la mejor administración de lo recaudado hacia las demás áreas de la Municipalidad Provincial de Cajamarca [6].

Melsi Ocas Terrones en su tesis desarrollada en el 2012 para optar el título profesional de Ingeniero de Sistemas, desarrollo de un data mart en el área de administración y finanzas de la Municipalidad Distrital de los Baños del Inca; el trabajo consistió en recolectar y analizar los requerimientos para el análisis de la información mediante fichas de observación, lo que permitió diseñar la base de datos dimensional, diseñar cubos y también los reportes necesarios para el mejor estudio de los datos; como conclusión la implementación del data mart en la institución contribuyó a la mejor administración y gestión de la información, disminuyendo los tiempos del desarrollo de reportes y tiempos en la toma de decisiones que son necesarios para decisiones estratégicas basadas en información de calidad [4].

José Eduardo Córdova Yupanqui en su tesis desarrollada en el 2013 para optar por el título de Ingeniero Informático, análisis diseño e implementación de una solución de inteligencia de negocios para el área de importaciones en una empresa comercializadora/importadora; el trabajo consistió en analizar y desarrollar los requerimientos funcionales y no funcionales planteados en el levantamiento de información con lo que se pudo definir la arquitectura para el proceso ETL para la correcta obtención de los datos; como conclusión se pudo automatizar los procesos en la obtención de

información permitiendo tener una respuesta más rápida y correcta, ayudando en la toma de decisiones [7].

Alejandro Rojas Zaldívar en su tesis desarrollada el 2014 para optar por el título profesional de Ingeniero de Computación y Sistemas, implementación de un data mart como solución de inteligencia de negocios, bajo la metodología de Ralph Kimball para optimizar la toma de decisiones en el departamento de finanzas de la Contraloría General de la República; el trabajo consistió en mejorar el proceso de análisis de información mediante la implementación de un data mart para poder generar información veraz y oportuna; como conclusión se logró analizar y seleccionar los datos indispensables para el desarrollo de la solución de inteligencia de negocios [3].

Claudia Estefanía Jaramillo Mejía en su trabajo de graduación, desarrollado en el año 2015, previo a la obtención del título de Ingeniero Informático, implementación de un sistema de inteligencia de negocios en el departamento de ventas de una empresa nacional del sector lácteo y desarrollo de reportes dinámicos utilizando herramientas de software libre; el trabajo consistió en analizar la información relacionada al departamento de ventas de las fuentes existentes, que cumpla con las especificaciones y pautas necesarias para la toma de decisiones, para ello se desarrolló un data mart en el que se cargó la información analizada; como conclusión se logró empaquetar debidamente la ejecución de las consultas, personalizándose según las necesidades de cada área [8].

2.2.Bases teóricas

2.2.1.Base de datos

Una base de datos o banco de datos es un conjunto de datos pertenecientes a un mismo contexto y almacenados sistemáticamente para su posterior uso. En este sentido; una biblioteca puede considerarse una base de datos compuesta en su mayoría por documentos y textos impresos en papel e indexados para su consulta. Actualmente, y debido al desarrollo tecnológico de campos como la informática y la electrónica,

la mayoría de las bases de datos están en formato digital, siendo este un componente electrónico, por tanto se ha desarrollado y se ofrece un amplio rango de soluciones al problema del almacenamiento de datos. Existen programas denominados sistemas gestores de bases de datos, abreviado SGBD (del inglés database management system o DBMS), que permiten almacenar y posteriormente acceder a los datos de forma rápida y estructurada [9].



Figura 1 Base de Datos [10]

2.2.2.Sistema de información

Un sistema de información (SI) es un conjunto de elementos orientados al tratamiento y administración de datos e información, organizados y listos para su uso posterior, generados para cubrir una necesidad o un objetivo. Habitualmente el término "sistema de información" se usa de manera errónea como sinónimo de sistema de información informático, en parte porque en la mayoría de los casos los recursos materiales de un sistema de información están constituidos casi en su totalidad por sistemas informáticos. Estrictamente hablando, un sistema de información no tiene por qué disponer de dichos recursos (aunque en la práctica esto no suele ocurrir). Se podría decir entonces que los sistemas de información informáticos son una subclase o un subconjunto de los sistemas de información en general [11].



Figura 2 Sistema de información [12]

Un sistema de información realiza cuatro actividades básicas [13]:

Entrada de información: Es el proceso mediante el cual el sistema de información toma los datos que requiere para procesar la información.

Almacenamiento de información: El almacenamiento es una de las actividades o capacidades más importantes, ya que a través de esta propiedad el sistema puede recordar la información guardada en la sección o proceso anterior.

Procesamiento de información: Es la capacidad del sistema de información para efectuar cálculos de acuerdo con una secuencia de operaciones preestablecida. Estos cálculos pueden efectuarse con datos introducidos recientemente en el sistema o bien con datos que están almacenados. Esta característica de los sistemas permite la transformación de datos fuente en información que puede ser utilizada para la toma de decisiones.

Salida de información: La salida es la capacidad de un sistema de información para sacar la información procesada o bien datos de entrada al exterior. Es importante aclarar que la salida de un sistema de información puede constituir la entrada a otro sistema de información o módulo.

2.2.3.Data warehouse

Es una base de datos que se caracteriza por integrar y depurar información de una o más fuentes distintas, para luego procesarla permitiendo su análisis desde infinidad de perspectivas y con grandes velocidades de respuesta. La creación de un data warehouse representa en la mayoría de las ocasiones el primer paso, desde el punto de vista técnico, para implantar una solución completa y fiable de business intelligence.

La ventaja principal de este tipo de bases de datos radica en las estructuras en las que se almacena la información (modelos de tablas en estrella, en copo de nieve, cubos relacionales, etc.). Este tipo de persistencia de la información es homogénea y fiable, y permite la consulta y el tratamiento jerarquizado de la misma (siempre en un entorno diferente a los sistemas operacionales) [14].

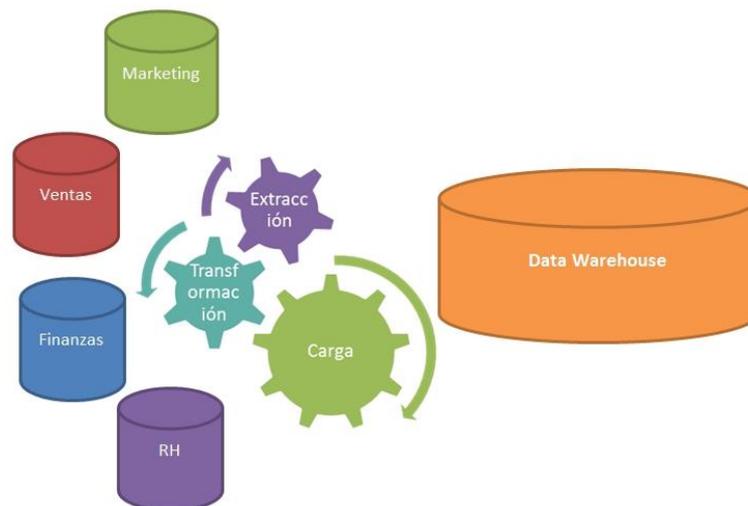


Figura 3 Data warehouse [15]

2.2.3.1.Características de un data warehouse

Los data warehouse cuentan con 4 características principales [16]:

Integrado, porque los datos almacenados en el data warehouse deben integrarse en una estructura consistente, por lo que las inconsistencias existentes entre los diversos sistemas operacionales deben ser eliminadas. La información suele estructurarse también en distintos niveles de detalle para adecuarse a las distintas necesidades de los usuarios.

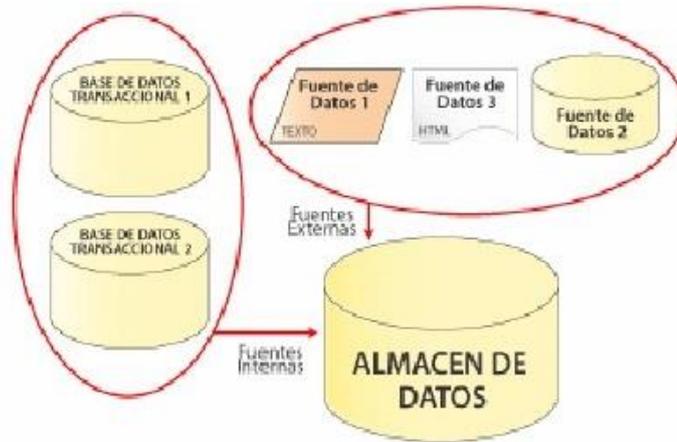
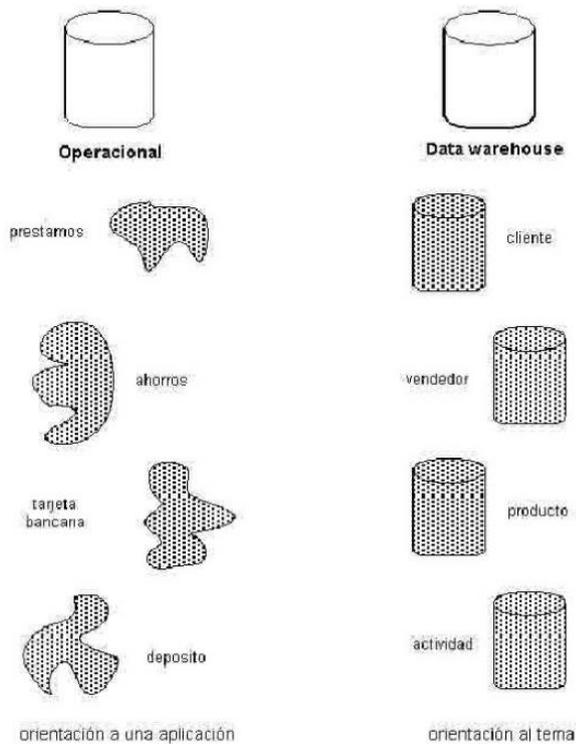


Figura 4 Data warehouse – integrado [16]

Temático, debido a que sólo los datos necesarios para el proceso de generación del conocimiento del negocio se integran desde el entorno operacional. Los datos se organizan por temas para facilitar su acceso y entendimiento por parte de los usuarios finales. Por ejemplo, todos los datos sobre clientes pueden ser consolidados en una única tabla del data warehouse. De esta forma, las peticiones de información sobre clientes serán más fáciles de responder dado que toda la información reside en el mismo lugar.



El data warehouse tiene una fuerte orientación al tema

Figura 5 Data warehouse – temático [16]

Histórico, pues el tiempo es parte implícita de la información contenida en un data warehouse. En los sistemas operacionales, los datos siempre reflejan el estado de la actividad del negocio en el momento presente. Por el contrario, la información almacenada en el data warehouse sirve, entre otras cosas, para realizar análisis de tendencias. Por lo tanto, el data warehouse se carga con los distintos valores que toma una variable en el tiempo para permitir comparaciones.

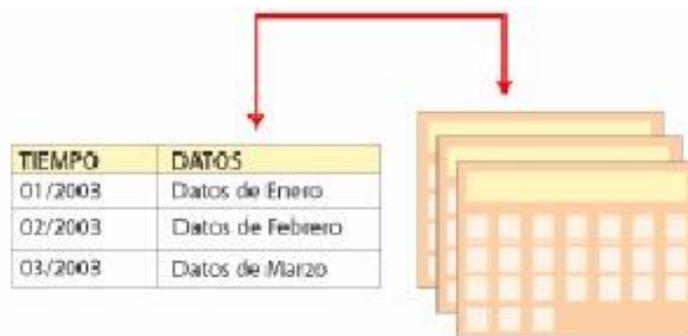


Figura 6 Data warehouse – histórico [16]

No volátil, ya que el almacén de información de un data warehouse existe para ser leído, pero no modificado. La información es por tanto permanente, significando la actualización del data warehouse la incorporación de los últimos valores que tomaron las distintas variables contenidas en él sin ningún tipo de acción sobre lo que ya existía.

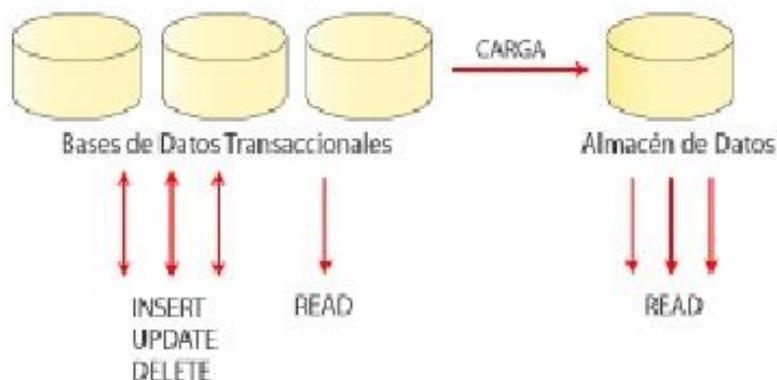


Figura 7 Data warehouse - no volátil [16]

2.2.3.2. Beneficios de un data warehouse

Entre los beneficios de un data warehouse tenemos [14]:

- ✓ Proporciona una herramienta para la toma de decisiones en cualquier área funcional, basándose en información integrada y global del negocio.
- ✓ Facilita la aplicación de técnicas estadísticas de análisis y modelización para encontrar relaciones ocultas.
- ✓ Proporciona la capacidad de aprender de los datos del pasado y de predecir situaciones futuras en diversos escenarios.
- ✓ Simplifica dentro de la empresa la implantación de sistemas de gestión integral de la relación con el cliente.
- ✓ Supone una optimización tecnológica y económica en entornos de centro de información, estadística o de generación de informes.

2.2.3.3. Diferencias entre un data warehouse y un data mart

Tabla 1 Diferencias entre un data warehouse y un data mart [17]

	Data warehouse	Data mart
Alcance	Satisface las necesidades de información de toda la organización.	Satisface las necesidades de información de un área específica.
Objetivo	Diseñado para optimizar la integración y la administración de los datos fuente.	Diseñado para optimizar la entrega de información de soporte a decisiones.
Características de los datos	Administra grandes cantidades de datos históricos a nivel atómico.	Se concentra en administrar resúmenes y/o datos totalizados.
Pertenencia	Pertenece a toda la organización.	Orientado a un área del negocio.
Administración	Es administrado por la unidad de sistemas de la organización.	Es administrado por el personal de sistemas de la unidad propietaria del data mart.

2.2.4. Business intelligence (BI)

Business intelligence es la habilidad para transformar los datos en información, y la información en conocimiento, de forma que se pueda optimizar el proceso de toma de decisiones en los negocios [18].



Figura 8 Datos - información - conocimiento [18]

Desde un punto de vista más pragmático, y asociándolo directamente con las tecnologías de la información, podemos definir business intelligence como el conjunto de metodologías, aplicaciones y tecnologías que permiten reunir, depurar y transformar datos de los sistemas transaccionales e información desestructurada (interna y externa a la compañía) en información estructurada, para su explotación directa (reporting, análisis OLTP / OLAP, alertas...) o para su análisis y conversión en conocimiento, dando así soporte a la toma de decisiones sobre el negocio.

La inteligencia de negocio actúa como un factor estratégico para una empresa u organización, generando una potencial ventaja competitiva, que no es otra que proporcionar información privilegiada para responder a los problemas de negocio: entrada a nuevos mercados, promociones u ofertas de productos, eliminación de islas de información, control financiero, optimización de costes, planificación de la producción, análisis de perfiles de clientes, rentabilidad de un producto concreto, etc.

En definitiva una solución BI completa permite [18]:

- ✓ Observar ¿Qué está ocurriendo?
- ✓ Comprender ¿Por qué ocurre?
- ✓ Predecir ¿Qué ocurriría?
- ✓ Colaborar ¿Qué debería hacer el equipo?
- ✓ Decidir ¿Qué camino se debe seguir?

2.2.4.1. Arquitectura de una solución BI

Una solución de business intelligence parte de los sistemas de origen de una organización (bases de datos, ERPs, ficheros de texto...), sobre los que suele ser necesario aplicar una transformación estructural para optimizar su proceso analítico.

Para ello se realiza una fase de extracción, transformación y carga (ETL) de datos. Esta etapa suele apoyarse en un almacén intermedio, llamado ODS, que actúa como pasarela entre los sistemas fuente y los sistemas destino (generalmente un data warehouse), y cuyo principal objetivo consiste en evitar la saturación de los servidores funcionales de la organización [19].

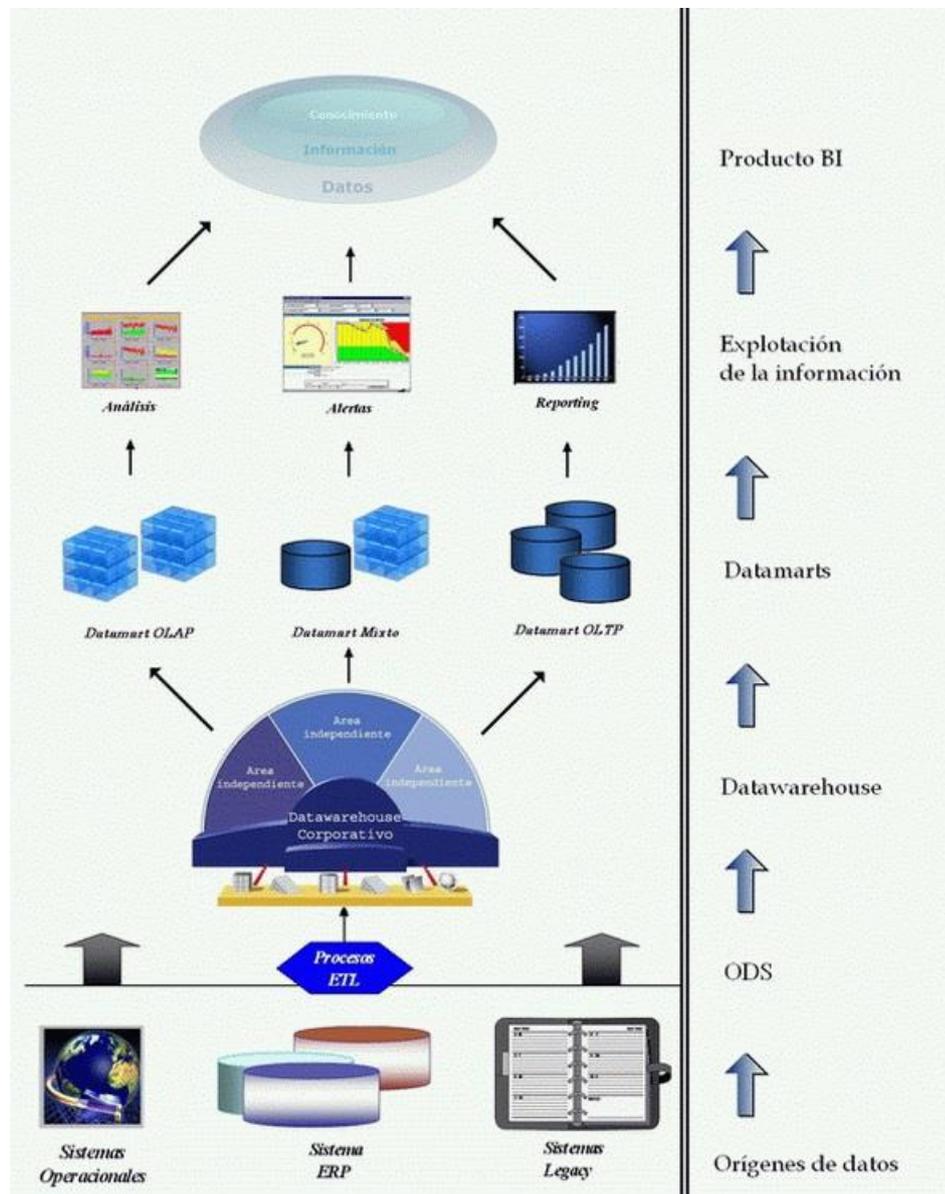


Figura 9 Arquitectura de una solución BI [19]

Los datos albergados en el data warehouse o en cada datamart se explotan utilizando herramientas comerciales de análisis, reporting, alertas... etc. En estas herramientas se basa también la construcción de productos BI más completos, como los sistemas de soporte a la decisión (DSS), los sistemas de información ejecutiva (EIS) y los cuadros de mando (CMI) o Balanced Scorecard (BSC) [19].

2.2.5. Modelo dimensional

El modelado dimensional siempre utiliza los conceptos de hechos (medidas) y dimensiones (contexto). Los hechos son normalmente (pero no siempre) los valores numéricos que se pueden agregar, y las dimensiones son grupos de jerarquías y descriptores que definen los hechos. Por ejemplo, la cantidad de ventas es un hecho; marca de tiempo, producto, número de registro, número de tienda, etc., son elementos de dimensiones. Los modelos dimensionales son construidos por el área de proceso de negocio, por ejemplo, las ventas en tiendas, inventarios, reclamaciones, etc. Debido a que las diferentes áreas de proceso de negocio comparten algunas pero no todas las dimensiones, la eficiencia en el diseño, la operación y la coherencia, se logra usando tablas de dimensión, es decir, utilizando una copia de la dimensión compartida. El término "tablas de dimensión" se originó por Ralph Kimball [20].

2.2.5.1. Procesos del modelado dimensional

El modelo dimensional se construye sobre un esquema de estrella, en donde los pasos a seguir son [20]:

- ✓ **Escoger el proceso de negocio**, describir el proceso de negocio en el que se basa el modelo. Esto podría ser por ejemplo una situación de ventas en una tienda al por menor. Para describir el proceso de negocio, se puede optar por hacer esto en texto plano o utilizar Notación de Modelado de Procesos de Negocio (BPMN en inglés) u otras guías de diseño, como el Lenguaje Unificado de Modelado (UML en inglés).
- ✓ **Declarar el "grain"**, es la descripción exacta en lo que el modelo dimensional debería concentrarse. Para aclarar lo que significa el

"grain", usted debe escoger el proceso central y describirlo con una sola oración. Además el "grain" (oración) es a lo que se le va a construir sus dimensiones y tabla de hechos. Puede que le resulte necesario volver a este paso para alterar el "grain" debido a nueva información obtenida en lo que su modelo supone entregar.

- ✓ **Identificar las dimensiones**, las dimensiones deben ser definidas dentro del "grain". Las dimensiones son la base de la tabla de hechos, y es donde se recogen los datos de la tabla de hechos. Normalmente las dimensiones son sustantivos, como fecha, tienda, inventario, etc. Estas dimensiones son donde se almacenan todos los datos, por ejemplo, la dimensión fecha podría contener datos tales como año, mes y día de la semana.
- ✓ **Identificar los hechos**, consiste en crear las llaves de la tabla de hechos. Este paso es identificar los hechos numéricos que poblarán cada fila de la tabla de hechos. Este paso está estrechamente relacionado con los usuarios de negocio del sistema, ya que es donde consiguen el acceso a los datos almacenados en el almacén de datos. Por lo tanto la mayor parte de las filas de la tabla de hecho son cifras numéricas, aditivos tales como cantidad o costo por unidad, etc.

2.2.5.2. Beneficios del modelado dimensional

Los beneficios del modelo dimensional son los siguientes [20]:

- ✓ **Comprensibilidad** - En comparación con el modelo normalizado, el modelo dimensional es más fácil de entender y más intuitivo. En los modelos dimensionales, la información se agrupa en dimensiones coherentes, por lo que es más fácil de leer e interpretar. La simplicidad también permite al software navegar las bases de datos de manera eficiente. En los modelos normalizados, los datos se divide en muchas entidades discretas e incluso un proceso de negocio simple podría resultar en docenas de tablas unidas entre sí de una manera compleja.
- ✓ **El rendimiento de consultas** - Los modelos dimensionales están más desnormalizados y optimizados para las consultas de datos, mientras que los modelos normalizados buscan eliminar redundancias de datos

y están optimizados para la carga de transacciones y actualización. El marco predecible de un modelo dimensional permite a la base de datos hacer fuertes supuestos sobre los datos, por lo que puede tener un impacto positivo en el rendimiento. Cada dimensión es el equivalente a un punto de entrada en la tabla de hechos, y esta estructura simétrica permite un manejo eficaz de consultas complejas. La optimización de consulta se hace simple, predecible y controlable.

- ✓ Extensibilidad - Los modelos dimensionales son escalables y fácilmente acomodados a nuevos datos inesperados. Las tablas existentes pueden ser cambiadas, ya sea por la simple adición de nuevas filas de datos en la tabla o ejecutar en SQL algún comando de tipo "Alter table". Las consultas o aplicaciones montadas sobre el almacén de datos no necesitan ser reprogramadas para acomodarse a los nuevos cambios. Las consultas y aplicaciones antiguas continúan funcionando sin producir resultados diferentes. Pero en los modelos normalizados cada modificación se debe considerar cuidadosamente, debido a las complejas dependencias entre las tablas de bases de datos.

2.2.6. Metodologías de diseño y construcción de un data warehouse

2.2.6.1. Metodología de Ralph Kimball

La Metodología Kimball, es una metodología empleada para la construcción de un almacén de datos (data warehouse, DW) que no es más que, una colección de datos orientada a un determinado ámbito (empresa, organización, etc.), integrado, no volátil y variable en el tiempo, que ayuda a la toma de decisiones en la entidad en la que se utiliza.

La metodología se basa en lo que Kimball denomina Ciclo de Vida Dimensional del Negocio (Business Dimensional Life Cycle). Este ciclo de vida del proyecto de DW, está basado en cuatro principios básicos:

- ✓ Centrarse en el negocio
- ✓ Construir una infraestructura de información adecuada

- ✓ Realizar entregas en incrementos significativos (este principio consiste en crear el almacén de datos (DW) en incrementos entregables en plazos de 6 a 12 meses, en este punto, la metodología se parece a las metodologías ágiles de construcción de software)
- ✓ Ofrecer la solución completa (en este punto se proporcionan todos los elementos necesarios para entregar valor a los usuarios de negocios, para esto ya se debe tener un almacén de datos bien diseñado, se deberán entregar herramientas de consulta ad hoc, aplicaciones para informes y análisis avanzado, capacitación, soporte, sitio web y documentación).

La construcción de una solución de DW/BI (Data Warehouse/Business Intelligence) es sumamente compleja, y Kimball nos propone una metodología que nos ayuda a simplificar esa complejidad. Las tareas de esta metodología (ciclo de vida) se describen a continuación [21]:

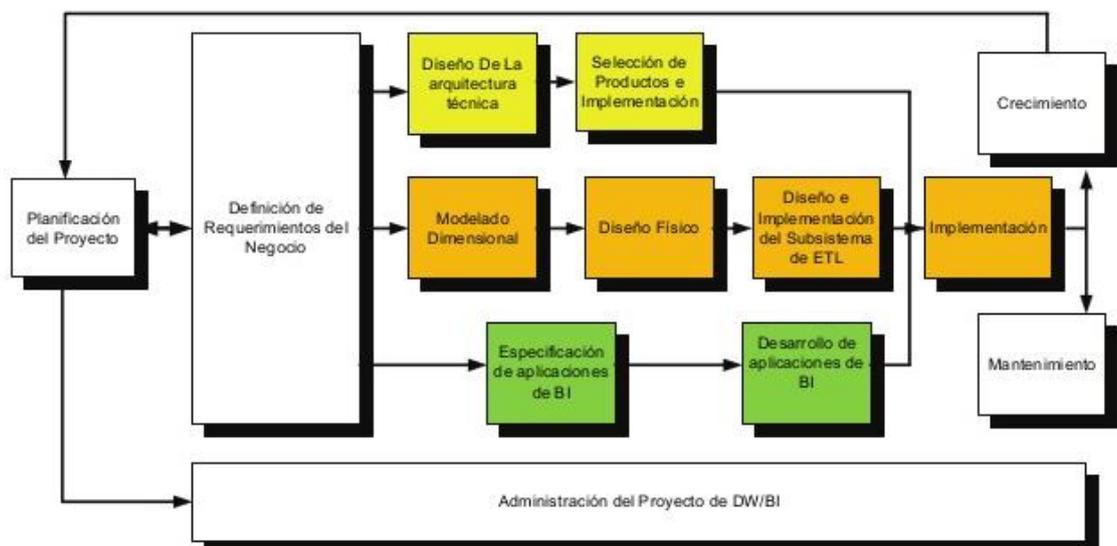


Figura 10 Diagrama de la metodología de Ralph Kimball [21]

❖ **Planificación del proyecto:** En este proceso se determina el propósito del proyecto de DW/BI, sus objetivos específicos y el alcance del mismo, los principales riesgos y una aproximación inicial a las necesidades de información.

Esta tarea incluye las siguientes acciones típicas de un plan de proyecto:

- ✓ Definir el alcance (entender los requerimientos del negocio).
- ✓ Identificar las tareas

- ✓ Programar las tareas
- ✓ Planificar el uso de los recursos.
- ✓ Asignar la carga de trabajo a los recursos
- ✓ Elaboración de un documento final que representa un plan del proyecto.

Además en esta parte definimos cómo realizar la administración o gestión de esta sub-fase que es todo un proyecto en sí mismo, con las siguientes actividades:

- ✓ Monitoreo del estado de los procesos y actividades.
- ✓ Rastreo de problemas
- ✓ Desarrollo de un plan de comunicación comprensiva que dirija la empresa y las áreas de TI

❖ **Definición de requerimientos del negocio:** La definición de requerimientos, es un proceso de entrevistar al personal de negocio y técnico, aunque siempre conviene, tener un poco de preparación previa. En esta tarea, se debe aprender sobre el negocio, los competidores, la industria y los clientes del mismo. Se debe dar una revisión a todos los informes posibles de la organización; rastrear los documentos de estrategia interna; entrevistar a los empleados, analizar lo que se dice en la prensa acerca de la organización, la competencia y la industria y se deben conocer los términos y la terminología del negocio.

Se sugiere entrevistar al personal que se encuentra en los cuatro grupos que se mencionan a continuación:

- ✓ El directivo responsable de tomar las decisiones estratégicas.
- ✓ Los administradores intermedios y de negocio responsables de explorar alternativas estratégicas y aplicar decisiones
- ✓ El personal de sistemas, si existe (estas son las personas que realmente saben qué tipos de problemas informáticos y de datos existen en la organización)
- ✓ El personal que se entrevista por razones políticas.

Entre las tareas antes descritas, existe una flecha bidireccional, esto indica que los requerimientos del negocio son el soporte inicial de las tareas subsiguientes, también tiene influencia en el plan de proyecto.

Si avanzamos por el camino central del diagrama, encontramos las tareas asociadas al área de datos, en esta, diseñaremos e implementaremos el modelo dimensional, y desarrollaremos el subsistema de extracción, transformación y carga (Extract, Transformation, and Load - ETL) para cargar el DW. Las tareas pertenecientes al área, se describen a continuación:

❖ **Modelado dimensional:** Es un proceso dinámico y altamente iterativo. Comienza con un modelo dimensional de alto nivel obtenido a partir de los procesos priorizados y descritos en la tarea anterior, y el proceso iterativo consiste en cuatro pasos:

- a) Elegir el proceso de negocio.
- b) Establecer el nivel de granularidad.
- c) Elegir las dimensiones.
- d) Identificar medidas y las tablas de hechos.

❖ **Diseño físico:** En esta tarea, se contestan las siguientes preguntas:

- ✓ ¿Cómo puede determinar cuán grande será el sistema de DW/BI?
- ✓ ¿Cuáles son los factores de uso que llevarán a una configuración más grande y más compleja?
- ✓ ¿Cómo se debe configurar el sistema?
- ✓ ¿Cuánta memoria y servidores se necesitan? ¿Qué tipo de almacenamiento y procesadores?
- ✓ ¿Cómo instalar el software en los servidores de desarrollo, prueba y producción?
- ✓ ¿Qué necesitan instalar los diferentes miembros del equipo de DW/BI en sus estaciones de trabajo?
- ✓ ¿Cómo convertir el modelo de datos lógico en un modelo de datos físicos en la base de datos relacional?
- ✓ ¿Cómo conseguir un plan de indexación inicial?
- ✓ ¿Debe usarse la partición en las tablas relacionales?

❖ **Diseño e implementación del subsistema de Extracción, Transformación y Carga (ETL):** El subsistema de Extracción, Transformación y Carga (ETL) es la base sobre la cual se alimenta el data warehouse. Si se diseña adecuadamente, puede extraer los datos de los sistemas de origen de datos, aplicar diferentes reglas para aumentar la calidad y consistencia de los mismos, consolidar la información proveniente de distintos sistemas, y finalmente cargar (grabar) la información en el DW en un formato acorde para la utilización por parte de las herramientas de análisis.

❖ **Implementación:** La implementación representa la convergencia de la tecnología, los datos y las aplicaciones de usuarios finales accesible desde el escritorio del usuario del negocio. Existen varios factores extras que aseguran el correcto funcionamiento de todas estas piezas, entre ellos se encuentran la capacitación, el soporte técnico, la comunicación y las estrategias de feedback.

❖ **Mantenimiento y crecimiento del data warehouse:** Para administrar el entorno del data warehouse existente es importante enfocarse en los usuarios de negocio, los cuales son el motivo de su existencia, además de gestionar adecuadamente las operaciones del data warehouse, medir y proyectar su éxito y comunicarse constantemente con los usuarios para establecer un flujo de retroalimentación, en esto consiste el mantenimiento. Finalmente, es importante sentar las bases para el crecimiento y evolución del data warehouse en donde el aspecto clave es manejar el crecimiento y evolución de forma iterativa utilizando el ciclo de vida propuesto, y establecer las oportunidades de crecimiento y evolución en orden por nivel prioridad.

Si avanzamos por el camino inferior del diagrama, encontramos las tareas asociadas al área aplicaciones de inteligencia de negocios, en esta ruta se encuentran tareas en las que diseñamos y desarrollamos las aplicaciones de negocios para los usuarios finales. Las tareas pertenecientes al área, se describen a continuación:

❖ **Especificación de aplicaciones de BI:** En esta tarea se proporciona, a una gran comunidad de usuarios una forma más estructurada y por lo tanto, más fácil, de acceder al almacén de datos. Se proporciona este acceso estructurado a través de lo que llamamos, aplicaciones de inteligencia de negocios (Business Intelligence Applications). Las aplicaciones de BI son la cara visible de la inteligencia de negocios: los informes y aplicaciones de análisis proporcionan información útil a los usuarios. Las aplicaciones de BI incluyen un amplio espectro de tipos de informes y herramientas de análisis, que van desde informes simples de formato fijo, a sofisticadas aplicaciones analíticas que usan complejos algoritmos e información del dominio. Kimball divide a estas aplicaciones en dos categorías basadas en el nivel de sofisticación, y les llama:

- a) Informes estándar: son informes relativamente simples, de formato predefinido, y parámetros de consulta fijos, proporcionan a los usuarios un conjunto básico de información acerca de lo que está sucediendo en un área determinada de la empresa y se utilizan día a día.
- b) Aplicaciones analíticas: Son más complejas que los informes estándar. Estas aplicaciones pueden incluir algoritmos y modelos de minería de datos, que ayudan a identificar oportunidades o cuestiones subyacentes en los datos, y el usuario puede pedir cambios en los sistemas transaccionales basándose en los conocimientos obtenidos del uso de la aplicación de BI.

Por último, en el camino superior, encontramos las tareas asociadas al área de tecnología en esta ruta, se encuentran las tareas relacionadas con software específico, por ejemplo, Microsoft SQL Analysis Services, etc. Las tareas pertenecientes al área, se describen a continuación:

❖ **Diseño de la arquitectura técnica:** El área de arquitectura técnica cubre los procesos y herramientas que se aplican a los datos. En el área técnica existen dos conjuntos que tienen distintos requerimientos, brindan sus propios servicios y componentes de almacenaje de datos, por lo que se consideran cada uno aparte: El back room (habitación trasera) y el front room (habitación frontal). El back room es el

responsable de la obtención y preparación de los datos, por lo que también se conoce como adquisición de datos y el front room es responsable de entregar los datos a la comunidad de usuario y también se le conoce como acceso de datos.

❖ **Conclusiones:** La metodología de Kimball proporciona una base empírica y metodológica adecuada para las implementaciones de almacenes de datos pequeños y medianos, dada su gran versatilidad y su enfoque ascendente, que permite construir los almacenes en forma escalonada. Además presenta una serie de herramientas, tales como planillas, gráficos y documentos, que proporcionan una gran ayuda para iniciarse en el ámbito de la construcción de un data warehouse [22].

2.2.6.2. Metodología de Bill Inmon

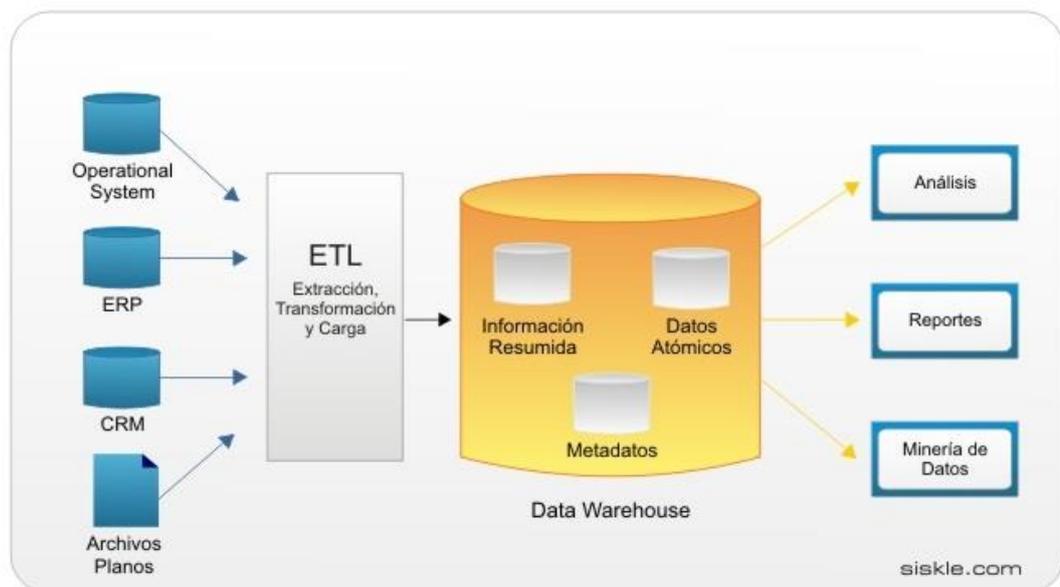


Figura 11 Arquitectura de Data Warehouse - Bill Inmon [3]

Bill Inmon ve la necesidad de transferir la información de los diferentes OLTP (sistemas transaccionales) de las organizaciones a un lugar centralizado donde los datos puedan ser utilizados para el análisis (sería el CIF o Corporate Information Factory). Insiste, además, en que ha de tener las siguientes características [3]:

- ✓ **Orientado a temas:** Los datos sobre 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 estos 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.

La información ha de estar a los máximos niveles de detalle. Los data warehouse departamentales o data marts son tratados como subconjuntos de este data warehouse corporativo, que son construidos para cubrir las necesidades individuales de análisis de cada departamento, y siempre a partir de este data warehouse central (del que también se pueden construir los ODS (Operational Data Stores) o similares). El enfoque Inmon también se referencia normalmente como top down. Los datos son extraídos de los sistemas operacionales por los procesos ETL y cargados en las áreas de stage, donde son validados y consolidados en el DW corporativo, y además existen los llamados metadatos que documentan de una forma clara y precisa el contenido del DW. Una vez realizado estas etapas, los procesos, de refresco de los data mart departamentales obtienen la información de él, y con las consiguientes transformaciones, organizan los datos en las estructuras particulares requeridas por cada uno de ellos, refrescando su contenido. Al tener este enfoque global es más difícil de desarrollar en un proyecto sencillo (pues estamos intentando abordar el “todo”, a partir del cual, luego, iremos al “detalle”) [3].

2.2.6.3. Diferencias entre Ralph Kimball e Inmon

Tabla 2 Diferencias entre Ralph Kimball e Inmon [3]

	Kimball	Inmon
Objetivo	Todas las empresas necesitan almacenar, analizar e interpretar los datos que van generando y acumulando, para luego tomar decisiones críticas que les permitan maximizar la prosperidad. Para ello, se necesita un sistema que les ayude a entender los datos y logren cumplir sus objetivos.	
Diseño	Utiliza el enfoque "Bottom - Up"	Utiliza el enfoque "Top - Down"
Enfoque	Tiene un enfoque por procesos que son manejados por las diferentes áreas del proceso. Trata de responder necesidades específicas según el tema.	Tiene un enfoque global de toda la empresa. No está basado en requerimientos específicos.
Costos	Implementar cada data mart permite que la solución no presente un alto costo.	Los costos aumentan, debido a que se replican grandes cantidades de datos
Tiempo	Debido a que en primer lugar debemos implementar los data marts, el tiempo de implementación es rápido.	Debido a que se implementa por completo el DWH se demanda mucho más tiempo.
Modelado	Plantea usar el modelamiento dimensional: esquema estrella, identificación de dimensiones y hechos.	Propone tres niveles en el modelo de datos del data warehouse: alto nivel (Entity Relationship Diagram), nivel medio (Data Item Set), nivel bajo (Physical Model). Sin embargo menciona que para implementar los data mart debe hacerse con modelamiento dimensional

Al establecer una comparación entre las dos metodologías más importantes que son la metodología de Ralph Kimball (y su enfoque dimensional), y la metodología de Bill Inmon (y su enfoque empresarial warehouse).

La metodología de Inmon es más apropiada para sistemas complejos, donde se quiere asegurar la perdurabilidad y consistencia de la información aunque cambien los procesos de negocio de la organización. Para proyectos pequeños donde se quiere asegurar la usabilidad de los usuarios que permita un desarrollo rápido e incremental de la solución donde no se tiene claro el panorama global, el enfoque de Kimball es el más apropiado [3].

2.2.7.Toma de decisiones

La toma de decisiones a nivel individual se caracteriza por el hecho de que una persona haga uso de su razonamiento y pensamiento para elegir una solución a un problema que se le presente en la vida; es decir, si una persona tiene un problema, deberá ser capaz de resolverlo individualmente tomando decisiones con ese específico motivo. También, la toma de decisiones es considerada como una de las etapas de la dirección.

En la toma de decisiones importa la elección de un camino a seguir, por lo que en un estado anterior deben evaluarse alternativas de acción. Si estas últimas no están presentes, no existirá decisión. Para tomar una decisión, cualquiera que sea su naturaleza, es necesario conocer, comprender, analizar un problema, para así poder darle solución.

En algunos casos, por ser tan simples y cotidianos, este proceso se realiza de forma implícita y se soluciona muy rápidamente, pero existen otros casos en los cuales las consecuencias de una mala o buena elección pueden tener repercusiones en la vida y si es en un contexto laboral en el éxito o fracaso de la organización, para los cuales es necesario realizar un proceso más estructurado que puede dar más seguridad e información para resolver el problema [23].

2.2.8.Optimización

Según Marta B. Ferrero y Omar J. A. Chiotti, se puede definir como optimización al proceso de seleccionar, a partir de un conjunto de alternativas posibles, aquella que mejor satisfaga el o los objetivos propuestos [24].

2.2.9.Información

La información está constituida por un grupo de datos ya supervisados y ordenados, que sirven para construir un mensaje basado en un cierto fenómeno o ente. La información permite resolver problemas y tomar decisiones, ya que su aprovechamiento racional es la base del conocimiento [25].

2.2.10.1.Principales características de la información

En general la información tiene una estructura interna y puede ser calificada según varias características [26]:

- ✓ Significado (semántica): Del significado extraído de una información, cada individuo evalúa las consecuencias posibles y adecúa sus actitudes y acciones de manera acorde a las consecuencias previsibles que se deducen del significado de la información. Esto se refiere a qué reglas debe seguir el individuo o el sistema experto para modificar sus expectativas futuras sobre cada posible alternativa.
- ✓ Importancia (relativa al receptor): Es decir, si trata sobre alguna cuestión importante. La importancia de la información para un receptor se referirá a en qué grado cambia la actitud o la conducta de los individuos. En las modernas sociedades, los individuos obtienen de los medios de comunicación masiva gran cantidad de información, una gran parte de la misma es poco importante para ellos, porque altera de manera muy poco significativa la conducta de los mismos. Esto se refiere a en qué grado cuantitativo deben alterarse las expectativas futuras. A veces se sabe que un hecho hace menos probables algunas cosas y más otras, la importancia tiene que ver con cuanto menos probables serán unas alternativas respecto a las otras.

- ✓ Vigencia (en la dimensión espacio-tiempo): Se refiere a si está actualizada o desfasada. En la práctica la vigencia de una información es difícil de evaluar, ya que en general acceder a una información no permite conocer de inmediato si dicha información tiene o no vigencia.
- ✓ Validez (relativa al emisor): Se evalúa si el emisor es fiable o puede proporcionar información no válida (falsa). Tiene que ver si los indicios deben ser considerados en la revaluación de expectativas o deben ser ignorados por no ser indicios fiables.
- ✓ Valor (activo intangible volátil): La utilidad que tiene dicha información para el destinatario.

CAPÍTULO III MATERIALES Y MÉTODOS

3.1.Procedimiento

Este proyecto de tesis tendrá como base fundamental la metodología de Ralph Kimball, quien presenta un esquema llamado BDL (Business Dimensional Life Cycle o Ciclo de Vida Dimensional del Negocio) en donde se muestra las diferentes etapas por la que debe pasar todo proyecto de data warehouse.

3.1.1.Planeación del proyecto

3.1.1.1.Descripción de la empresa



Figura 12 Logo - Industrias alimentarias Huacariz

La industria alimentaria Huacariz, es una empresa cajamarquina dedicada a la elaboración y comercialización de derivados lácteos. Inició operaciones en el año 1992, desde entonces y con arduo esfuerzo han logrado ser reconocidos como una de las empresas líderes del sector lácteo en el norte del país, ofreciendo una amplia variedad de productos en las líneas de quesos, yogurt, dulces de leche y cremas de leche.

Son una empresa dedicada a la elaboración de productos lácteos de alta calidad, orientada a la plena satisfacción de sus clientes, compitiendo cada vez en mercados más exigentes, logrando contar en la actualidad con 8 locales ubicados estratégicamente en la ciudad de Cajamarca, es por ello que utilizan la tecnología adecuada y un equipo calificado tanto para la producción de sus producto como también para el manejo de su información, contando así con un sistema de información para el registro

de sus ventas, productos en stock y también las ordenes de producción, el sistema mencionado se encuentra en ejecución desde el año 2015, brindándonos hasta el momento 2 años y medio de data histórica.

3.1.1.2.Objetivos

- ✓ Identificar y analizar los requerimientos de información que intervendrán en los modelos dimensionales (Data Warehouse)
- ✓ Diseñar e implementar el Data Warehouse para el área de comercialización y producción.
- ✓ Crear reportes con información veraz y oportuna, dirigida al usuario final en forma consolidada y representativa para el proceso de toma de decisiones.
- ✓ Determinar el impacto analizando el nivel de optimización en la toma de decisiones en el área de comercialización y producción.

3.1.1.3.Alcance

- ✓ El data warehouse abarca el área de comercialización y producción.
- ✓ Se hará uso de toda la data histórica existente en la base de datos actual para la creación de reportes.
- ✓ Como resultado se obtendrán los modelos dimensionales que serán la fuente de los reportes que muestren la información para la toma de decisiones.

3.1.1.4.Riesgos

- ✓ Se utilizarán los recursos existentes, cuyo hardware podría no satisfacer los requerimientos del software pues la herramienta a utilizar para la creación de reportes trabaja a nivel de memoria ram, y el consumo de la misma es equivalente a la cantidad de datos a procesar.
- ✓ Los usuarios funcionales podrían no tener tiempo disponible al inicio de la investigación para un análisis de requerimientos a corto tiempo.

3.1.1.5.Recurso humano

Tabla 3 Recurso humano

	Nombres y apellidos	Ocupación
Project manager	Tucto Basteres Anthony Deivy	Tesista
Project manager	Ing. Jaime Meza Huamán	Asesor

3.1.1.6.Presupuesto y retorno de la inversión

Tabla 4 Presupuesto – Activos tangibles

INVERSIÓN DE ACTIVOS TANGIBLES

ITEM	CANTIDAD INICIAL	MEDIDA	PRECIO UNITARIO	TOTAL INVERSIÓN
UTILES DE ESCRITORIO				
USB	1	Unidad	S/ 15.00	S/. 15.00
Papel A4 (millar)	2	millar	S/ 11.00	S/. 22.00
Tintas	4	Unidad	S/ 10.00	S/. 40.00
CD's	4	Unidad	S/ 1.00	S/. 4.00
Lapiceros	2	Unidad	S/ 1.00	S/. 2.00
Plumon indeleble	1	Unidad	S/ 2.50	S/. 2.50
Perforador	1	Unidad	S/ 12.00	S/. 12.00
Engrampador	1	Unidad	S/ 10.00	S/. 10.00
EQUIPOS DE OFICINA				
Laptop	1	Unidad	S/ 1,400.00	S/. 1,400.00
Impresora	1	Unidad	S/ 435.00	S/. 435.00
Escritorio	1	Unidad	S/ 200.00	S/. 200.00
Silla de oficina	1	Unidad	S/ 90.00	S/. 90.00
TOTAL INVERSION				2,232.50

Tabla 5 Presupuesto - Activos intangibles

INVERSIÓN DE ACTIVOS INTANGIBLES

ITEM	CANTIDAD INICIAL	MEDIDA	PRECIO UNITARIO	TOTAL INVERSIÓN
IMPLEMENTACIÓN				
SQL server 2008	5	Mes	S/ 148.00	S/. 740.00
Office 2010	5	Mes	S/ 32.00	S/. 160.00
CA Erwin Data Modeler	5	Mes	S/ 225.00	S/. 1,125.00
QlikView (Versión Gratuita)	5	Mes	S/ -	S/. -
TOTAL INVERSION				2,025.00

Tabla 6 Presupuesto - Otros gastos

OTROS GASTOS

ITEM	CANTIDAD	MEDIDA	PRECIO UNITARIO	TOTAL INVERSIÓN
Luz	5	Mes	S/ 10.00	S/. 50.00
Internet	5	Mes	S/ 100.00	S/. 500.00
TOTAL INVERSION				550.00

Tabla 7 Presupuesto - Gastos personal

GASTOS DE PERSONAL

ITEM	CANTIDAD	MEDIDA	PRECIO UNITARIO	TOTAL INVERSIÓN
Desarrollo	5	Mes	S/ 1,000.00	S/. 5,000.00

Tabla 8 Presupuesto - Costos proyectados

COSTOS PROYECTADOS - DATA WAREHOUSE

ITEMS	AÑO: 0	AÑO: 1	AÑO: 2	AÑO: 3	AÑO: 4	AÑO: 5
UTILES DE ESCRITORIO						
USB	S/. 15.00					
Papel A4 (millar)	S/. 22.00					
Tintas	S/. 40.00					
CD's	S/. 4.00					
Lapiceros	S/. 2.00					
Plumon indeleble	S/. 2.50					
Perforador	S/. 12.00					
Engrampador	S/. 10.00					
EQUIPOS DE OFICINA						
Laptop	S/. 1,400.00					
Impresora	S/. 435.00					
Escritorio	S/. 200.00					
Silla de oficina	S/. 90.00					
IMPLEMENTACIÓN						
SQL server 2008	S/. 740.00	S/. 1,776.00				
Office 2010	S/. 160.00					
CA Erwin Data Modeler	S/. 1,125.00					
QlikView	S/.					
OTROS GASTOS						
Luz	S/. 50.00	S/. 120.00				
Internet	S/. 500.00					
GASTOS DE PERSONAL						
Desarrollo	S/. 5,000.00					
TOTAL DE GASTOS	S/. 9,807.50	S/. 1,896.00				

Tabla 9 Comparación de gastos anuales por horas trabajadas

Reporte	Pre-Test (Mes)	Post-test (Mes)
RE1	960	0.19
RE2	35	0.17
RE3	1440	0.18
RE4	360	0.18
RE5	120	0.17
RE6	140	0.18
RE7	720	0.17
RE8	360	0.16
Minutos totales	4135.00	1.40
Horas totales	68.92	0.02
Sueldo por hora	S/ 18.75	S/ 18.75
Gasto Mensual Total	S/ 1,292.19	S/ 0.44
Gasto Anual Total	S/ 15,506.25	S/ 5.25

Tabla 10 Beneficio

ANALISIS DEL INDICADOR

INDICADOR	ANTES	BENEFICIO	DESPUES
Horas trabajadas	S/ 15,506.25	S/ 15,501.00	S/ 5.25

Tabla 11 Ingresos Proyectados

INGRESOS PROYECTADOS	AÑO 1	AÑO 2	AÑO 3	AÑO 4	AÑO 5
	S/ 15,501.00				

Tabla 12 Diferencia entre ingresos y costos proyectados

FLUJO DE CAJA NETO PROYECTO

AÑO 0	AÑO 1	AÑO 2	AÑO 3	AÑO 4	AÑO 5
-9,807.50	13,605.00	13,605.00	13,605.00	13,605.00	13,605.00

Tabla 13 Cálculo del valor COK

$CPPC = WACC = \frac{D}{D+C} \times Kd \times (1 - T) + \frac{C}{D+C} \times Ke$		LEYENDA D= Deuda K= Capital Kd= Costo Deuda 14.46% T= Impuesto a la Renta 30% Ke= Rentabilidad Accionista ROE	
DEUDA	239,071	33%	
CAPITAL	495,369	67%	
TOTAL	734,440	100%	
CPPC = Costo Prom Ponderado de Capital			
UTILIDAD ANTES DEL IMPUESTO		149,290	
IMP. A LA RENTA		44,787	
UTILIDAD NETA (DESPUES DE LA RENTA)		104,503	
$Ke = Roe = \frac{UTILIDAD NETA}{TOTAL PATRIMONIO}$		$\frac{104,503.00}{495,369.00}$	21%
CPPC=	22.85%		

Tabla 14 Indicadores de evaluación

Indicadores de evaluación

COK	22.85%
VA	S/. 38,258.21
VAN	28,450.71
TIR	137%
IR	3.90

Mejor alternativa de inversión en bonos

- VAN > 0 **acepta el proyecto**
- TIR > COK **se acepta el proyecto**
- IR > 1 **Indice de rentabilidad > 1 Acepta el proyecto**
- Por cada sol de inversión retorna S/2.90 de rentabilidad**

Si bien el retorno de la inversión en este proyecto no será recuperado como un ingreso de dinero directo para la empresa, se podrán ver los beneficios en la reducción de costos y un mejor análisis del mercado.

Tras el análisis de los indicadores de evaluación se encuentra que el proyecto puede ser aceptado ya que el IR es mayor a 1, teniendo que por cada sol de inversión retorna S/2.90 de rentabilidad.

3.1.2. Definición de requerimientos del negocio

3.1.2.1. Requerimientos del negocio

A continuación se presentan los requerimientos no funcionales que se deberá tener en cuenta en las áreas de comercialización y producción de la industria alimentaria Huacariz.

Tabla 15 Cuadro de requerimiento no funcional N°1

Identificador	R-01	Nombre	Elaborar modelo dimensional
Tipo	No funcional		
Descripción	Cada área contará con su modelo dimensional.		

Tabla 16 Cuadro de requerimiento no funcional N°2

Identificador	R-02	Nombre	Software
Tipo	No funcional		
Descripción	Para el proceso de ETL se hará uso del software con el que ya cuenta la empresa (SQL Server 2008, SQL Server Integration Services), y para consumir los modelos dimensionales se usará la versión gratuita de QlikView.		

Tabla 17 Cuadro de requerimiento no funcional N°3

Identificador	R-03	Nombre	Hardware
Tipo	No funcional		
Descripción	Los requisitos mínimos son SO Windows 7 x64 o Windows Server 2008 x64 en adelante, procesador Inter Core 2 Duo a más, 4GB ram, 300MB en disco.		

Los requerimientos que se presentan a continuación, están enfocados al área de comercialización de la industria alimentaria Huacariz.

Tabla 18 Cuadro de requerimiento funcional N°4

Identificador	R-04	Nombre	Venta anual
Tipo	Funcional		
Descripción	Conocer el monto vendido por año en meses.		

Tabla 19 Cuadro de requerimiento funcional N°5

Identificador	R-05	Nombre	Meta para las ventas
Tipo	Funcional		
Descripción	Establecer una meta para las ventas mensuales.		

Tabla 20 Cuadro de requerimiento funcional N°6

Identificador	R-06	Nombre	Ventas por documento
Tipo	Funcional		
Descripción	Conocer el monto vendido por año segmentado en meses según el tipo de documento.		

Tabla 21 Cuadro de requerimiento funcional N°7

Identificador	R-07	Nombre	Ventas por precio
Tipo	Funcional		
Descripción	Comparar las ventas por el precio del producto versus el precio al que se vendió el producto		

Tabla 22 Cuadro de requerimiento funcional N°8

Identificador	R-08	Nombre	Productos con más demanda
Tipo	Funcional		
Descripción	Conocer los productos con mayor demanda que representen aproximadamente el 50% de las ventas.		

Tabla 23 Cuadro de requerimiento funcional N°9

Identificador	R-09	Nombre	Cuentas por cobrar
Tipo	Funcional		
Descripción	Conocer las ventas que tengan deuda pendiente.		

Los requerimientos que se presentan a continuación, están enfocados al área de producción de la industria alimentaria Huacariz.

Tabla 24 Cuadro de requerimiento funcional N°10

Identificador	R-10	Nombre	Consumo de leche anual
Tipo	Funcional		
Descripción	Permitir conocer la cantidad de leche que se consume por año segmentado en meses.		

Tabla 25 Cuadro de requerimiento funcional N°11

Identificador	R-11	Nombre	Proyección de consumo de leche
Tipo	Funcional		
Descripción	Permitir establecer una proyección para el consumo de leche en cada mes.		

Tabla 26 Cuadro de requerimiento funcional N°12

Identificador	R-12	Nombre	Uso de cada insumo
Tipo	Funcional		
Descripción	Permitir conocer la cantidad utilizada de cada insumo por año segmentado en meses.		

Tabla 27 Cuadro de requerimiento funcional N°13

Identificador	R-13	Nombre	Proyección de uso por insumo
Tipo	Funcional		
Descripción	Permitir establecer una proyección para el uso de cada insumo en cada mes.		

Tabla 28 Cuadro de requerimiento funcional N°14

Identificador	R-14	Nombre	Productos más producidos
Tipo	Funcional		
Descripción	Permitir conocer los productos que representen aproximadamente el 50% de la producción.		

3.1.2.2. Entrevistas

Las entrevistas se realizaron a las siguientes personas:

✓ Ing. Casas Medina Renzo Jankarlo

Se orientó principalmente en dos lineamientos:

- Entender la forma de extraer reportes de la base de datos.
- Entender el funcionamiento de la base de datos transaccional.

✓ Ing. Rojas Gálvez Viviana

- Entender los procesos en el área de producción para entender qué información necesitan analizar.

✓ Lic. Bardales Castillo Wilda

- Entender los procesos en el área de comercialización para entender qué información necesitan analizar.

✓ Ing. Vergara Quiroz Carlos

- Entender el nivel de detalle en la que se mostrará la información.

3.1.2.3. Base de datos transaccional en SQL Server del área de comercialización



Figura 13 Base de datos - comercialización

tb_productosTerminados	
	id_productoTerminado
	nombreProducto
	precio
	stock
	id_categoria
	tipo
	descripcion
	unidad
	id_proveedor
	indicador
	stockbajo
	stockMedio
	abc
	cantPeso
	formula

Figura 14 Tabla - tb_productosTerminados

tb_detalleVenta	
	id_venta
	id_sucursal
	id_productoTerminado
	cantidad
	fecha
	saldo
	tipoMov
	tipoES
	precioVenta
	precioproducto
	tipo_ingre

Figura 15 Tabla - tb_detalleVenta

tb_tipoCliente	
	id_tipoCliente
	nombreTipoCliente
	descripcion

Figura 16 Tabla - tb_tipoCliente

tb_sucursalLugar	
	id_sucursal
	nombreSucursal
	direccion
	telefono
	tipo

Figura 17 Tabla - tb_sucursalLugar

tb_usuario	
	id_usuario
	usuario
	contraseña
	nombres
	apellidos
	dni
	celular
	telefono
	direccion
	email
	fechaNacimiento
	sexo
	nivelAcademico
	nivel
	id_cargo
	indicador
	id_sucursal

Figura 18 Tabla - tb_usuario

tb_tipoDocumento	
	id_Documento
	nombreDocumento
	iniciales
	descripcion

Figura 19 Tabla - tb_tipoDocumento

tb_cliente	
	id_cliente
	nombreCliente
	nombres1
	apellidos
	Ruc
	celular
	telefono
	direccion
	email
	departamento
	provincia
	distrito
	representante
	descuento
	id_tipoCliente
	documento
	activo
	fechamodificacion
	usuario

Figura 20 Tabla - tb_cliente

tb_Ventas	
	id_venta
	fechaVenta
	id_cliente
	id_Documento
	nroDocumento
	valor
	igv
	total
	id_sucursal
	estado
	tipo
	id_usuario
	serie
	nombres
	rucDni
	numTranferencia
	creditoFin
	valorTexto
	direccion
	est_pagado
	cantidadPagada
	est_Depositado
	fecha_EstDepositado
	tipoVenta
	totalContado
	totalTarjeta
	descuento
	fechamodificacion
	tipoTarjeta

Figura 21 Tabla - tb_Ventas

tb_Proveedores	
	id_proveedor
	razonSocial
	ruc
	direccion
	telefono
	celular
	email
	nombreRepresentante
	indicador
	id_categoria
	tipo
	moneda
	numCuenta

Figura 22 Tabla - tb_Proveedores

tb_calendar	
Bimestre	
BimestreID	
Date	
[Day (#)]	
DayFin	
DayInicio	
LastDayMonthPrev	
Month	
[Month (#)]	
MonthID	
NDate	
NombreDia	
NomWeek	
PER_SK_PERIOD	
Period	
[Period (#)]	
Periodo	
Quarter	
QuarterID	
QuarterName	
QuarterName1	
vtoday	
Week	
WeekDay	
Year	

Figura 23 Tabla - tb_calendar

tb_categoria	
 id_categoria	
nombreCategoria	
Descripcion	
indicador	
tipo	

Figura 24 Tabla - tb_categoria

UbiDepartamento	
idDepartamento	
nombreDepartamento	

Figura 25 Tabla - UbiDepartamento

UbiProvincia	
idDepartamento	
idProvincia	
nombreProvincia	

Figura 26 Tabla - UbiProvincia

UbiDistrito	
idDepartamento	
idProvincia	
idDistrito	
nombreDistrito	

Figura 27 Tabla - UbiDistrito

3.1.2.4. Diccionario de tablas del área de comercialización

Tabla 29 Tablas del área de comercialización

Tabla	Descripción
tb_productosTerminados	Productos que son producidos por Huacariz y mercadería que es adquirida para la venta.
tb_detalleVenta	Detalle de cada venta realizada
tb_tipoCliente	Son los diferentes tipos de clientes con los que cuenta la industria alimentaria Huacariz
tb_sucursalLugar	Sucursales pertenecientes a la industria alimentaria Huacariz
tb_usuario	Personas que tienen acceso al sistema
tb_tipoDocumento	Tipo de documento que se emite como medio de comprobante para una venta
tb_cliente	Cientes registrados
tb_Ventas	Información general de cada venta realizada
tb_Proveedores	Son los proveedores que brindan tanto productos terminado como insumos
tb_calendar	Calendario maestro
tb_categoria	Es la categoría a la que pertenece cada producto
UbiDepartamento	Departamentos del Perú
UbiProvincia	Provincias de cada departamento del Perú
UbiDistrito	Distritos de cada provincia del Perú

3.1.2.5.Diccionario de datos del área de comercialización

Tabla 30 Descripción tb_productosTerminados

Columna	Tipo de dato	Null option	Descripción	PK	FK
id_productoTerminado	int	No	Código de producto	PK	-
nombreProducto	varchar(150)	No	Nombre producto	-	-
precio	decimal(18, 2)	No	Precio producto	-	-
stock	decimal(18, 3)	No	Stock producto	-	-
id_categoria	int	No	Categoría producto	-	FK
tipo	varchar(50)	No	Tipo producto	-	-
descripcion	varchar(150)	No	Descripción	-	-
unidad	varchar(50)	No	Unidad del producto	-	-
id_proveedor	int	No	Proveedor producto	-	FK
indicador	char(2)	No	Indica si el producto está activo	-	-
stockbajo	int	No	Stock bajo del producto	-	-
stockMedio	int	No	Stock medio del producto	-	-
abc	varchar(2)	No	Indica la rotación del producto	-	-
cantPeso	decimal(18, 3)	No	Peso del producto	-	-
formula	int	No	Formula del producto	-	FK

Tabla 31 Descripción tb_detalleVenta

Columna	Tipo de dato	Null option	Descripción	PK	FK
id_venta	int	No	Código de venta	-	FK
id_sucursal	Int	No	Sucursal de la venta	-	FK
id_productoTerminado	int	No	Producto vendido	-	FK
cantidad	decimal(18, 3)	Si	Cantidad vendida	-	-
fecha	datetime	Si	Fecha de venta	-	-
saldo	decimal(18, 3)	Si	Stock del producto	-	-
tipoMov	varchar(50)	No	Indica si es venta o rotación	-	-
tipoES	varchar(15)	No	Indica si es ingreso o salida	-	-
precioVenta	decimal(18, 2)	No	Precio de venta del producto	-	-
precioproducto	decimal(18, 2)	No	Precio real al que se vendió el producto	-	-
tipo_ingre	varchar(1)	No	Tipo de ingreso	-	-

Tabla 32 Descripción tb_tipoCliente

Columna	Tipo de dato	Null option	Descripción	PK	FK
id_tipoCliente	int	No	Código de tipo de cliente	PK	-
nombreTipoCliente	varchar(150)	Si	Nombre de tipo de cliente	-	-
descripcion	varchar(150)	Si	Descripción de tipo de cliente	-	-

Tabla 33 Descripción tb_sucursalLugar

Columna	Tipo de dato	Null option	Descripción	PK	FK
id_sucursal	int	No	Código de sucursal	PK	-
nombreSucursal	varchar(150)	Si	Nombre de sucursal	-	-
direccion	varchar(50)	Si	Dirección de sucursal	-	-
telefono	varchar(50)	Si	Teléfono de sucursal	-	-
tipo	varchar(20)	Si	Tipo de sucursales	-	-

Tabla 34 Descripción UbiDepartamento

Columna	Tipo de dato	Null option	Descripción	PK	FK
idDepartamento	int	No	Código departamento	PK	-
nombreDepartamento	varchar(100)	Si	Nombre departamento	-	-

Tabla 35 Descripción tb_usuario

Columna	Tipo de dato	Null option	Descripción	PK	FK
id_usuario	int	No	Código usuario	PK	-
usuario	varchar(50)	No	Alias usuario	-	-
contraseña	varchar(50)	No	Contraseña de usuario	-	-
nombres	varchar(150)	No	Nombres del usuario	-	-
apellidos	varchar(150)	No	Apellidos del usuario	-	-
dni	varchar(10)	No	DNI de usuario	-	-
celular	varchar(20)	No	Celular usuario	-	-
telefono	varchar(15)	No	Teléfono de usuario	-	-
direccion	varchar(150)	No	Dirección de usuario	-	-
email	varchar(150)	No	Email usuario	-	-
fechaNacimiento	datetime	No	Fecha de nacimiento del usuario	-	-
sexo	char(1)	No	Sexo usuario	-	-
nivelAcademico	varchar(150)	No	Nivel académico del usuario	-	-
nivel	char(1)	No	Nivel de acceso al sistema	-	-
id_cargo	varchar(5)	No	Cargo usuario	-	-
indicador	char(2)	No	Estado de usuario	-	-
id_sucursal	int	No	Sucursal de usuario	-	FK

Tabla 36 Descripción tb_cliente

Columna	Tipo de dato	Null option	Descripción	PK	FK
id_cliente	int	No	Código cliente	PK	-
nombreCliente	varchar(150)	No	Razón social	-	-
nombres1	varchar(250)	No	Nombres del cliente	-	-
apellidos	varchar(250)	No	Apellidos del cliente	-	-
ruc	varchar(12)	No	Ruc cliente	-	-
celular	varchar(30)	No	Celular cliente	-	-
telefono	varchar(30)	No	Teléfono	-	-
direccion	varchar(250)	No	Dirección cliente	-	-
email	varchar(150)	No	Email cliente	-	-
departamento	int	No	Departamento cliente	-	-
provincia	int	No	Provincia cliente	-	-
distrito	int	No	Distrito cliente	-	-
representante	varchar(150)	No	Representante cliente	-	-
descuento	decimal(18, 3)	No	Descuento por cliente	-	-
id_tipoCliente	int	No	Tipo de cliente	-	FK
documento	varchar(50)	No	Documento del cliente	-	-
activo	varchar(2)	No	Estado cliente	-	-
fechamodificacion	datetime	No	Fecha de actualización	-	-
usuario	int	No	Usuario que registro al cliente	-	-

Tabla 37 Descripción tb_Ventas

Columna	Tipo de dato	Null option	Descripción	PK	FK
id_venta	int	No	Código de venta	PK	-
fechaventa	datetime	No	Fecha de venta	-	-
id_cliente	int	No	Código de cliente	-	FK
id_Documento	int	No	Código de documento	-	FK
nroDocumento	varchar(15)	No	Numero de documento	-	-
valor	decimal(18, 2)	No	Monto total de la venta sin IGV	-	-
igv	decimal(18, 2)	No	IGV del monto total de la venta	-	-
total	decimal(18, 2)	No	Monto total de la venta incluido IGV	-	-
id_sucursal	int	No	Código de sucursal	-	FK
estado	varchar(20)	No	Estado del producto en la venta	-	-
tipo	varchar(20)	No	Tipo de destino	-	-
id_usuario	int	No	Código de usuario	-	FK
serie	varchar(5)	No	Serie del documento	-	-

nombres	varchar(200)	No	Nombre del cliente	-	-
rucDni	varchar(15)	No	Ruc o DNI del cliente	-	-
numTranferencia	varchar(25)	No	Numero de transferencia	-	-
creditoFin	varchar(2)	No	Estado del crédito	-	-
valorTexto	varchar(250)	No	Mono total en texto	-	-
direccion	varchar(150)	No	Dirección del cliente	-	-
est_pagado	varchar(2)	No	Estado de pago de la venta	-	-
cantidadPagada	decimal(18, 2)	No	Monto pagado	-	-
est_Depositado	varchar(2)	No	Estado del deposito	-	-
fecha_EstDepositado	datetime	No	Fecha deposito	-	-
tipoVenta	varchar(20)	No	Tipo de venta	-	-
totalContado	decimal(18, 2)	No	Monto pagado en efectivo	-	-
totalTarjeta	decimal(18, 2)	No	Monto pagado con tarjeta	-	-
descuento	decimal(18, 3)	No	Descuento al cliente	-	-
fechamodificacion	datetime	No	Fecha de modificación de la venta	-	-
tipoTarjeta	varchar(50)	No	Tipo de tarjeta usada para el pago	-	-

Tabla 38 Descripción tb_Proveedores

Columna	Tipo de dato	Null option	Descripción	PK	FK
id_proveedor	int	No	Código de proveedor	Si	-
razonSocial	varchar(150)	No	Razón social de proveedor	-	-
ruc	varchar(15)	No	Ruc del proveedor	-	-
dirección	varchar(150)	No	Dirección del proveedor	-	-
telefono	varchar(150)	No	Teléfono del proveedor	-	-
celular	varchar(150)	No	Celular del proveedor	-	-
email	varchar(150)	No	Email del proveedor	-	-
nombreRepresentante	varchar(150)	No	Representante del proveedor	-	-
indicador	char(2)	No	Estado activo del proveedor	-	-
id_categoria	int	Si	Código de categoría	-	-
tipo	varchar(15)	No	Tipo de proveedor	-	-
moneda	varchar(15)	No	Moneda con la que se le paga al proveedor	-	-
numCuenta	varchar(150)	No	Cuenta del proveedor	-	-

Tabla 39 Descripción tb_calendar

Columna	Tipo de dato	Null option	Descripción	PK	FK
Bimestre	nvarchar(255)	Si	Bimestre	-	-
BimestreID	float	Si	Id de bimestre	-	-
Date	datetime	Si	Fecha	-	-
[Day (#)]	float	Si	Día del mes	-	-
DayFin	datetime	Si	Fecha final del mes	-	-
DayInicio	datetime	Si	Fecha inicial del mes	-	-
LastDayMonthPrev	float	Si	Días que trae el mes	-	-
Month	nvarchar(255)	Si	Nombre mes	-	-
[Month (#)]	float	Si	N° de mes en el año	-	-
MonthID	float	Si	Id de mes	-	-
NDate	datetime	Si	Fecha y hora	-	-
NombreDia	nvarchar(255)	Si	Nombre y N° del día en el año	-	-
NomWeek	nvarchar(255)	Si	Semana de año	-	-
PER_SK_PERIOD	float	Si	Código de la fecha	-	-
Period	nvarchar(255)	Si	Nombre del mes y año	-	-
[Period (#)]	nvarchar(255)	Si	Año y mes	-	-
Periodo	nvarchar(255)	Si	Año y mes	-	-
Quarter	nvarchar(255)	Si	Cuarto de año	-	-
QuarterID	float	Si	Id de cuarto	-	-
QuarterName	nvarchar(255)	Si	Mes inicial y final de cuarto	-	-

			y año		
QuarterName1	nvarchar(255)	Si	Mes inicial y final de cuarto	-	-
vtoday	datetime	Si	Fecha y hora	-	-
Week	float	Si	Numero de semana en el año	-	-
WeekDay	nvarchar(255)	Si	Día de la semana	-	-
Year	float	Si	año	-	-

Tabla 40 Descripción tb_categoria

Columna	Tipo de dato	Null option	Descripción	PK	FK
id_categoria	int	No	Código categoría	PK	-
nombreCategoria	varchar(150)	Si	Nombre de la categoría	-	-
Descripcion	varchar(150)	Si	Descripción de la categoría	-	-
indicador	char(2)	Si	Indicador de la categoría	-	-
tipo	varchar(15)	Si	Tipo categoría	-	-

Tabla 41 Descripción UbiProvincia

Columna	Tipo de dato	Null option	Descripción	PK	FK
idDepartamento	int	No	Código departamento	-	-
idProvincia	int	No	Código provincia	-	-
nombreProvincia	varchar(100)	No	Nombre provincia	-	-

Tabla 42 Descripción UbiDistrito

Columna	Tipo de dato	Null option	Descripción	PK	FK
idDepartamento	int	No	Código departamento	-	-
idProvincia	int	No	Código provincia	-	-
idDistrito	int	No	Código de distrito	-	-
nombreDistrito	varchar(100)	No	Nombre de distrito	-	-

Tabla 43 Descripción tb_tipoDocumento

Columna	Tipo de dato	Null option	Descripción	PK	FK
id_Documento	int	No	Código de documento	PK	-
nombreDocumento	varchar(50)	Si	Nombre del documento	-	-
iniciales	char(2)	Si	Iniciales del documento	-	-
descripcion	varchar(150)	No	Descripción del documento	-	-

3.1.2.6. Base de datos transaccional en SQL Server del área de producción



Figura 28 Base de datos - Producción

tb_formulas	
id_formula	
producto	
cant_leche	
cant_producida	
indicador	
tipo	

Figura 29 Tabla - tb_formulas

tb_ordenProduccion	
id_ordenProduccion	
id_productoTerminado	
cant_producida	
fecha_produc	
observaciones	
id_usuario	
est_finalizado	
responsable	
fehcaregistro	
est_transformado	
lecheUtilizada	

Figura 30 Tabla - tb_ordenProduccion

tb_usuario	
	id_usuario
	usuario
	contraseña
	nombres
	apellidos
	dni
	celular
	telefono
	direccion
	email
	fechaNacimiento
	sexo
	nivelAcademico
	nivel
	id_cargo
	indicador
	id_sucursal

Figura 31 Tabla - tb_usuario

tb_detalleOrdenProduccion	
	id_ordenProduccion
	id_insumo
	cantidad
	tipoIngreso

Figura 32 Tabla - tb_detalleOrdenProduccion

tb_insumos	
	id_insumo
	nombreInsumo
	stock
	valorUnitario
	unidad
	id_categoria
	id_proveedor
	indicador
	stockBajo
	StockMedio
	tipo
	abc
	peso

Figura 33 Tabla - tb_insumos

tb_categoria	
	id_categoria
	nombreCategoria
	Descripcion
	indicador
	tipo

Figura 34 Tabla - tb_categoria

tb_calendar	
	Bimestre
	BimestreID
	Date
	[Day (#)]
	DayFin
	DayInicio
	LastDayMonthPrev
	Month
	[Month (#)]
	MonthID
	NDate
	NombreDia
	NomWeek
	PER_SK_PERIOD
	Period
	[Period (#)]
	Periodo
	Quarter
	QuarterID
	QuarterName
	QuarterName1
	vtoday
	Week
	WeekDay
	Year

Figura 35 Tabla - tb_calendar

3.1.2.7. Diccionario de tablas del área de producción

Tabla 44 Tablas del área de producción

Tabla	Descripción
tb_formulas	Guarda la fórmula para cada producto que se produce en la industria
tb_ordenProduccion	Orden emitida por un usuario del sistema para solicitar la producción de un determinado producto
tb_usuario	Personas que tienen acceso al sistema
tb_detalleOrdenProduccion	Detalle de la orden de producción
tb_insumos	Insumo requerido para la elaboración de productos
tb_categoria	Categoría de los productos elaborados
tb_calendar	Calendario maestro

3.1.2.8. Diccionario de datos del área de producción

Tabla 45 Descripción tb_formulas

Columna	Tipo de dato	Null option	Descripción	PK	FK
id_formula	int	No	Código formula	PK	-
producto	varchar(150)	Si	Nombre producto	-	-
cant_leche	float	Si	Cantidad de leche usada	-	-
cant_producida	decimal(18, 3)	Si	Cantidad producida del producto	-	-
indicador	char(2)	Si	Estado activo de la formula	-	-
tipo	varchar(15)	Si	Tipo de formula a producir	-	-

Tabla 46 Descripción tb_usuario

Columna	Tipo de dato	Null option	Descripción	PK	FK
id_usuario	int	No	Código usuario	PK	-
usuario	varchar(50)	No	Alias usuario	-	-
contraseña	varchar(50)	No	Contraseña de usuario	-	-
nombres	varchar(150)	No	Nombres del usuario	-	-
apellidos	varchar(150)	No	Apellidos del usuario	-	-
dni	varchar(10)	No	DNI de usuario	-	-
celular	varchar(20)	No	Celular usuario	-	-
telefono	varchar(15)	No	Teléfono de usuario	-	-
direccion	varchar(150)	No	Dirección de usuario	-	-
email	varchar(150)	No	Email usuario	-	-
fechaNacimiento	datetime	No	Fecha de nacimiento del usuario	-	-
sexo	char(1)	No	Sexo usuario	-	-
nivelAcademico	varchar(150)	No	Nivel académico del usuario	-	-
nivel	char(1)	No	Nivel de acceso al sistema	-	-
id_cargo	varchar(5)	No	Cargo usuario	-	-
indicador	char(2)	No	Estado de usuario	-	-
id_sucursal	int	No	Sucursal de usuario	-	FK

Tabla 47 Descripción tb_calendar

Columna	Tipo de dato	Null option	Descripción	PK	FK
Bimestre	nvarchar(255)	Si	Bimestre	-	-
BimestreID	float	Si	Id de bimestre	-	-
Date	datetime	Si	Fecha	-	-
[Day (#)]	float	Si	Día del mes	-	-
DayFin	datetime	Si	Fecha final del mes	-	-
DayInicio	datetime	Si	Fecha inicial del mes	-	-
LastDayMonthPrev	float	Si	Días que trae el mes	-	-
Month	nvarchar(255)	Si	Nombre mes	-	-
[Month (#)]	float	Si	N° de mes en el año	-	-
MonthID	float	Si	Id de mes	-	-
NDate	datetime	Si	Fecha y hora	-	-
NombreDia	nvarchar(255)	Si	Nombre y N° del día en el año	-	-
NomWeek	nvarchar(255)	Si	Semana de año	-	-
PER_SK_PERIOD	float	Si	Código de la fecha	-	-
Period	nvarchar(255)	Si	Nombre del mes y año	-	-
[Period (#)]	nvarchar(255)	Si	Año y mes	-	-
Periodo	nvarchar(255)	Si	Año y mes	-	-
Quarter	nvarchar(255)	Si	Cuarto de año	-	-
QuarterID	float	Si	Id de cuarto	-	-
QuarterName	nvarchar(255)	Si	Mes inicial y final de cuarto	-	-

			y año		
QuarterName1	nvarchar(255)	Si	Mes inicial y final de cuarto	-	-
vtoday	datetime	Si	Fecha y hora	-	-
Week	float	Si	Numero de semana en el año	-	-
WeekDay	nvarchar(255)	Si	Día de la semana	-	-
Year	float	Si	año	-	-

Tabla 48 Descripción tb_categoria

Columna	Tipo de dato	Null option	Descripción	PK	FK
id_categoria	int	No	Código categoría	PK	-
nombreCategoria	varchar(150)	Si	Nombre categoría	-	-
Descripcion	varchar(150)	Si	Descripcion categoría	-	-
indicador	char(2)	Si	Estado activo de la categoría	-	-
tipo	varchar(15)	Si	Tipo de categoría	-	-

Tabla 49 Descripción tb_detalleOrdenProduccion

Columna	Tipo de dato	Null option	Descripción	PK	FK
id_ordenProduccion	int	No	Código orden	PK	-
id_insumo	int	No	Código insumo	PK	-
cantidad	decimal(18, 3)	No	Cantidad del insumo	-	-
tipoIngreso	varchar(50)	No	Tipo de ingreso	-	-

Tabla 50 Descripción tb_insumos

Columna	Tipo de dato	Null option	Descripción	PK	FK
id_insumo	int	No	Código de insumo	PK	-
nombreInsumo	varchar(150)	No	Nombre de insumo	-	-
stock	decimal(18, 3)	No	Stock de insumo	-	-
valorUnitario	decimal(18, 2)	No	Valor unitario del insumo	-	-
unidad	varchar(50)	No	Unidad de insumo	-	-
id_categoria	int	No	Código categoría	-	-
id_proveedor	int	No	Código proveedor	-	-
indicador	char(2)	No	Estado activo del insumo	-	-
stockBajo	int	No	Stock bajo insumo	-	-
StockMedio	int	No	Stock medio insumo	-	-
tipo	varchar(50)	No	Tipo de insumo	-	-
abc	varchar(2)	No	Abc del insumo	-	-
peso	decimal(18, 3)	No	Peso del insumo	-	-

Tabla 51 Descripción tb_ordenProduccion

Columna	Tipo de dato	Null option	Descripción	PK	FK
id_ordenProduccion	int	No	Código de orden	PK	-
id_productoTerminado	int	Si	Código de formula	-	FK
cant_producida	decimal(18, 3)	Si	Cantidad producida	-	-
fecha_produc	datetime	Si	Fecha de producción	-	-
observaciones	varchar(150)	Si	Observaciones de la orden	-	-
id_usuario	int	No	Código usuario que realizo la orden	-	FK
est_finalizado	char(2)	No	Estado de finalización de la orden	-	-
responsable	char(18)	Si	Responsable de la orden	-	-
fecharegistro	datetime	Si	Fecha de registro de la orden	-	-
est_transformado	char(2)	No	Estado pendiente o finalizada de la orden	-	-
lecheUtilizada	decimal(18, 3)	No	Leche usada para la orden de producción	-	-

3.1.3. Modelo dimensional

Habiendo realizado el análisis de los requerimientos basándonos en la metodología de Ralph Kimball pasaremos a identificar las medidas y dimensiones que servirán para analizar la información en sus diferentes niveles.

3.1.3.1. Elección del esquema a utilizar

Un **esquema de estrella** es un tipo de esquema de base de datos relacional que consta de una sola tabla de hechos central rodeada de tablas de dimensiones [27]. 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 [28].



Figura 36 Esquema estrella [27]

Un **esquema en copo de nieve** es una estructura algo más compleja que el esquema en estrella. Se da cuando alguna de las dimensiones se implementa con más de una tabla de datos. La finalidad es normalizar las tablas y así reducir el espacio de almacenamiento al eliminar la redundancia de datos; pero tiene la contrapartida de generar peores rendimientos al tener que crear más tablas de dimensiones y más relaciones entre las tablas lo que tiene un impacto directo sobre el rendimiento [29].

En la siguiente figura se muestra un esquema de copo de nieve con dos dimensiones, cada una con tres niveles. Un esquema de copo de nieve puede tener varias dimensiones y cada dimensión puede tener varios niveles [30].

Esquema de copo de nieve

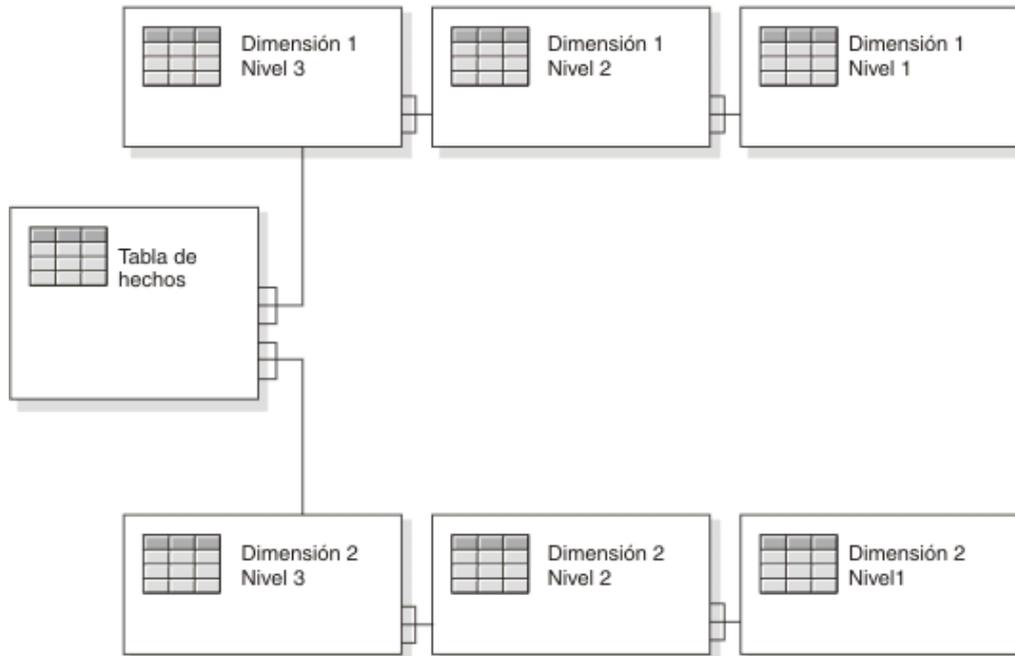


Figura 37 Esquema de copo de nieve [30]

El único argumento a favor de los esquemas en copo de nieve es que al estar normalizadas las tablas de dimensiones, se evita la redundancia de datos y con ello se ahorra espacio [29].

El diseño de esquemas en estrella permite implementar la funcionalidad de una base de datos multidimensional utilizando una clásica base de datos relacional. Este esquema es ideal por su simplicidad y velocidad para ser usado en análisis multidimensionales (OLAP, datamarts, EIS, etc). Permite acceder tanto a datos agregados como de detalle. Otra razón para utilizar los esquemas en estrella es su simplicidad desde el punto de vista del usuario final. Las consultas no son complicadas, ya que las condiciones y las uniones (join) necesarias sólo involucran a la tabla de hechos y a las de dimensiones, no haciendo falta que se encadenen uniones y condiciones a dos o más niveles como ocurriría en un esquema en copo de nieve. En la mayoría de los casos son

preferibles los de estrellas por su simplicidad respecto a los de copo de nieve por ser más fáciles de manejar.

Finalmente, el diseño de esquemas en estrella es la opción que se usará ya que presenta un mejor rendimiento y velocidad pues permite indexar las dimensiones de forma individualizada sin que repercuta en el rendimiento de la base de datos en su conjunto [31].

3.1.3.2. Elección de las dimensiones para el área de comercialización

Para determinar las dimensiones del data mart en el área de comercialización, se identificó las variables de análisis por las cuales el usuario suele elaborar sus reportes. Entre las más destacadas se encontraron:

- ✓ Producto vendido
- ✓ Categoría del producto
- ✓ Sucursal que realizo la venta
- ✓ Tipo de venta
- ✓ Estado de pago de la venta
- ✓ Tipo de sucursal
- ✓ Tipo de documentos emitidos por la venta
- ✓ Proveedores del producto
- ✓ Ruc del proveedor
- ✓ Usuario que realizo la venta
- ✓ Cliente que compro el producto
- ✓ Estado del cliente
- ✓ Fecha de venta del producto
- ✓ Mes de venta del producto

Una vez identificadas las variables usadas para la elaboración de reportes, se agrupan por afinidad para dar origen a las dimensiones que se usaran en el modelo dimensional del data mart de comercialización.

Tabla 52 Elección de dimensiones para el área de comercialización

Dimensiones	Variables identificadas
Dim_Producto	Producto vendido
	Categoría del producto
Dim_Sucursal	Sucursal que realizo la venta
	Tipo de sucursal
Dim_Documento	Tipo de documentos emitidos por venta
Dim_Proveedor	Proveedores del producto
	Ruc del proveedor
Dim_Usuario	Usuario que realizo la venta
Dim_Cliente	Cliente que compro el producto
	Estado del cliente
Dim_Calendario	Fecha de venta del producto
	Mes de venta del producto
Dim_Venta	Tipo de venta
	Estado de pago de la venta

3.1.3.3. Elección de las dimensiones para el área de producción

Para determinar las dimensiones del data mart en el área de producción, se identificó las variables de análisis por las cuales el usuario suele elaborar sus reportes. Entre las más destacadas se encontraron:

- ✓ Formula necesaria para la producción
- ✓ Usuario que solicite la producción
- ✓ Fecha de la orden de producción
- ✓ Numero de orden de producción
- ✓ Insumos utilizados para la producción

Una vez identificadas las variables usadas para la elaboración de reportes, se agrupan por afinidad para dar origen a las dimensiones que se usaran en el modelo dimensional del data mart de producción.

Tabla 53 Elección de dimensiones para el área de producción

Dimensiones	Variables identificadas
Dim_Formula	Formula necesaria para la producción
Dim_Usuario	Usuario que solicite la producción
Dim_Calendario	Fecha de la orden de producción
Dim_Orden	Numero de orden de producción
Dim_Insumo	Insumos utilizados

3.1.3.4.Total de dimensiones encontradas

Tras el análisis anterior, se pudo encontrar que el data warehouse contará con las siguientes dimensiones:

- ✓ Dim_Producto
- ✓ Dim_Sucursal
- ✓ Dim_Documento
- ✓ Dim_Proveedor
- ✓ Dim_Usuario
- ✓ Dim_Cliente
- ✓ Dim_Calendario
- ✓ Dim_Formula
- ✓ Dim_Orden
- ✓ Dim_Insumo
- ✓ Dim_Venta

3.1.3.5.Medidas encontradas para el área de comercialización

Según el análisis realizado, se encontraron las siguientes medidas:

- ✓ Cantidad vendida por producto
- ✓ Precio por producto
- ✓ Precio de venta por producto

3.1.3.6. Medidas encontradas para el área de producción

Según el análisis realizado, se encontraron las siguientes medidas:

- ✓ Cantidad producida de cada producto
- ✓ Cantidad utilizada de un insumo
- ✓ Cantidad de leche usada

3.1.3.7. Diseño del modelo dimensional para el área de producción

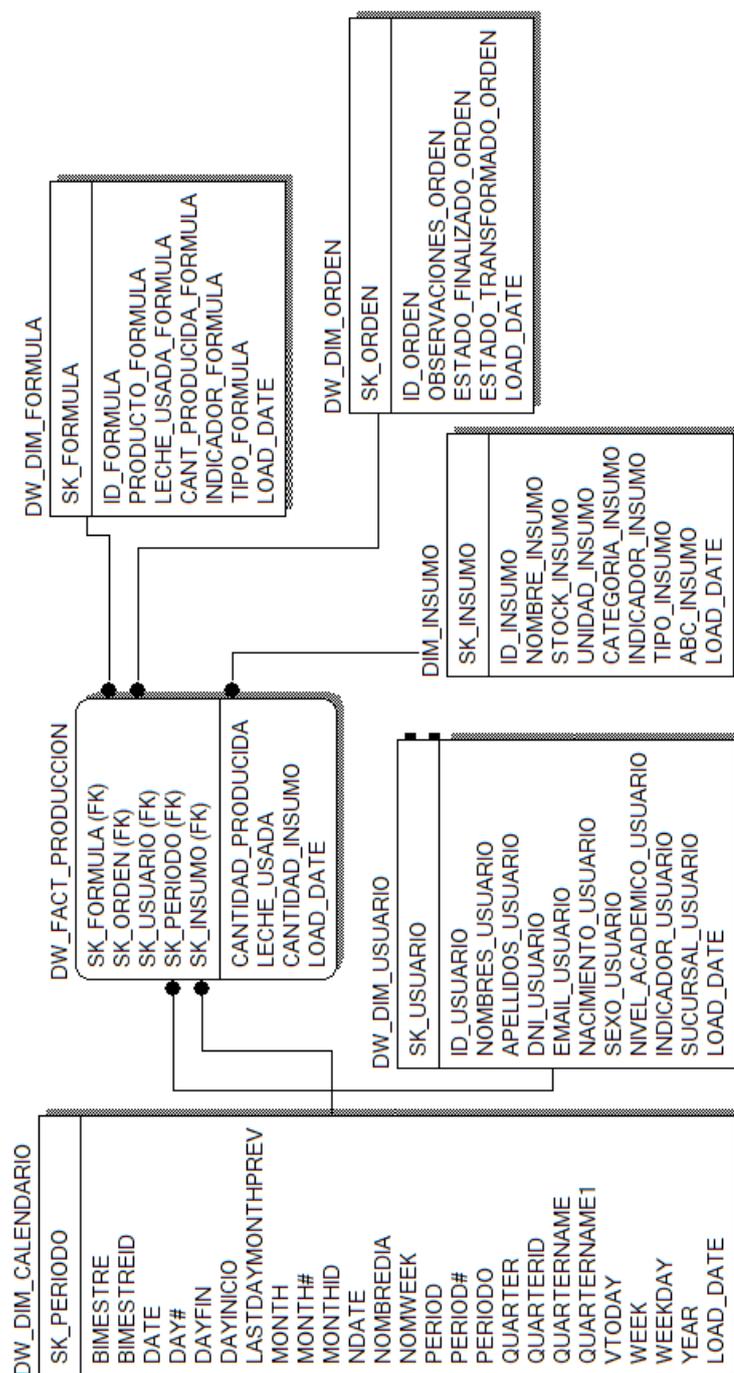


Figura 38 Modelo dimensional para el área de producción

3.1.3.8. Diseño del modelo dimensional para el área de comercialización

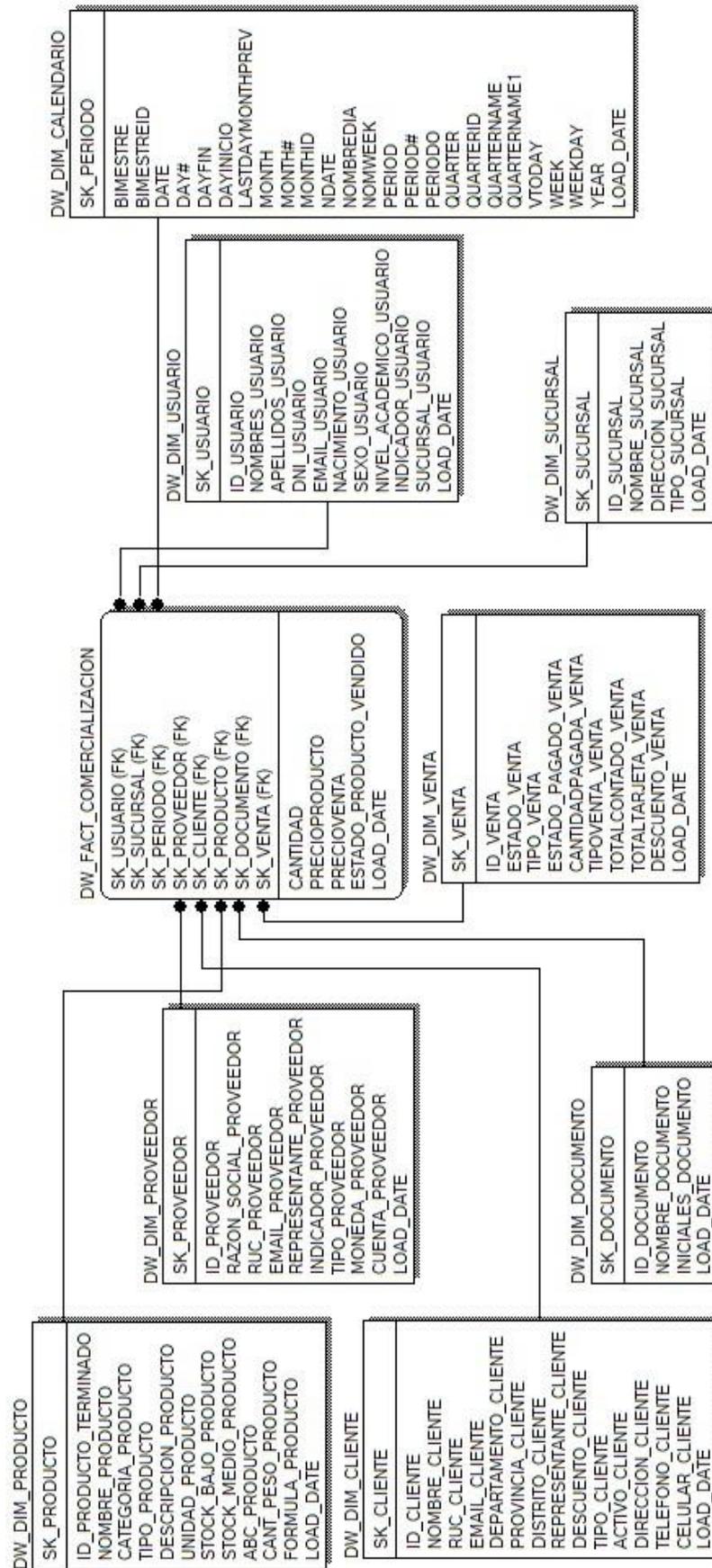


Figura 39 Modelo dimensional para el área de comercialización

3.1.3.9. Modelo general del data warehouse

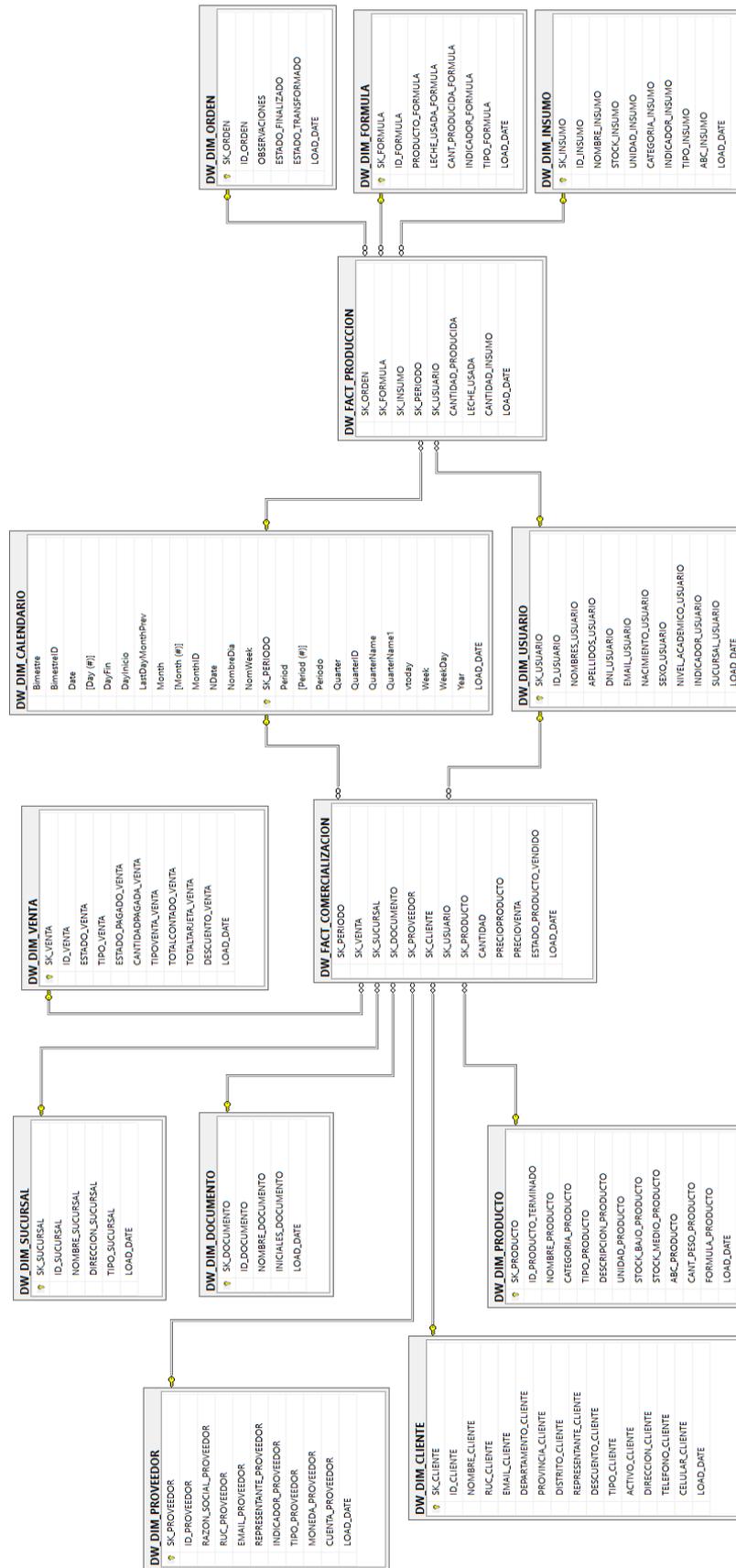


Figura 40 Modelo general del data warehouse

3.1.4.Diseño físico

3.1.4.1.Dimensión Dim_Producto

Tabla 54 Diseño físico de Dim_Producto

Columna	Tipo de dato	Descripción
SK_PRODUCTO	int	Clave surrogada
ID_PRODUCTO_TERMINADO	int	Código del producto
NOMBRE_PRODUCTO	varchar(150)	Nombre del producto
CATEGORIA_PRODUCTO	varchar(150)	Categoría del producto
TIPO_PRODUCTO	varchar(150)	Tipo de producto
DESCRIPCION_PRODUCTO	varchar(150)	Descripción del producto
UNIDAD_PRODUCTO	varchar(150)	Unidad del producto
STOCK_BAJO_PRODUCTO	int	Stock bajo del producto
STOCK_MEDIO_PRODUCTO	int	Stock medio del producto
ABC_PRODUCTO	varchar(150)	Abc del producto
CANT_PESO_PRODUCTO	decimal(18, 3)	Peso del producto
FORMULA_PRODUCTO	int	Formula del producto
LOAD_DATE	datetime	Fecha de carga

3.1.4.2.Dimensión Dim_Formula

Tabla 55 Diseño físico de Dim_Formula

Columna	Tipo de dato	Descripción
SK_FORMULA	int	Clave surrogada
ID_FORMULA	int	Código venta
PRODUCTO_FORMULA	varchar(150)	Producto a producir
LECHE_USADA_FORMULA	float	Cantidad de leche usada
CANT_PRODUCIDA_FORMULA	decimal(18, 3)	Cantidad producida del producto
INDICADOR_FORMULA	varchar(150)	Indicador de la formula
TIPO_FORMULA	varchar(150)	Tipo de formula
LOAD_DATE	datetime	Fecha de carga

3.1.4.3. Dimensión Dim_Venta

Tabla 56 Diseño físico de Dim_Venta

Columna	Tipo de dato	Descripción
SK_PRODUCTO	int	Clave surrogada
ID_VENTA	int	Código del producto
ESTADO_VENTA	varchar(150)	Categoría del producto
TIPO_VENTA	varchar(150)	Tipo de producto
ESTADO_PAGADO_VENTA	varchar(150)	Descripción del producto
CANTIDADPAGADA_VENTA	varchar(150)	Unidad del producto
TIPOVENTA_VENTA	int	Stock bajo del producto
TOTALCONTADO_VENTA	int	Stock medio del producto
TOTALTARJETA_VENTA	varchar(150)	Abc del producto
DESCUENTO_VENTA	decimal(18, 3)	Peso del producto
LOAD_DATE	datetime	Fecha de carga

3.1.4.4. Dimensión Dim_Proveedor

Tabla 57 Diseño físico de Dim_Proveedor

Columna	Tipo de dato	Descripción
SK_PROVEEDOR	int	Clave surrogada
ID_PROVEEDOR	int	Código proveedor
RAZON_SOCIAL_PROVEEDOR	varchar(150)	Razón social proveedor
RUC_PROVEEDOR	varchar(150)	Ruc del proveedor
EMAIL_PROVEEDOR	varchar(150)	Email del proveedor
REPRESENTANTE_ PROVEEDOR	varchar(150)	Representante del proveedor
INDICADOR_PROVEEDOR	varchar(150)	Indicador del proveedor
TIPO_PROVEEDOR	varchar(150)	Tipo de proveedor
MONEDA_PROVEEDOR	varchar(150)	Moneda del proveedor
CUENTA_PROVEEDOR	varchar(150)	Cuenta del proveedor
LOAD_DATE	datetime	Fecha de carga

3.1.4.5. Dimensión Dim_Sucursal

Tabla 58 Diseño físico de Dim_Sucursal

Columna	Tipo de dato	Descripción
SK_SUCURSAL	int	Clave surrogada
ID_SUCURSAL	int	Código sucursal
NOMBRE_SUCURSAL	varchar(150)	Nombre de sucursal
DIRECCION_SUCURSAL	varchar(150)	Dirección de sucursal
TIPO_SUCURSAL	varchar(150)	Tipo de sucursal
LOAD_DATE	datetime	Fecha de carga

3.1.4.6. Dimensión Dim_Cliente

Tabla 59 Diseño físico de Dim_Cliente

Columna	Tipo de dato	Descripción
SK_CLIENTE	int	Clave surrogada
ID_CLIENTE	int	Código cliente
NOMBRE_CLIENTE	varchar(150)	Nombre de cliente
RUC_CLIENTE	varchar(150)	Ruc del cliente
EMAIL_CLIENTE	varchar(150)	Email del cliente
DEPARTAMENTO_CLIENTE	varchar(150)	Departamento del cliente
PROVINCIA_CLIENTE	varchar(150)	Provincia del cliente
DISTRITO_CLIENTE	varchar(150)	Distrito del cliente
REPRESENTANTE_CLIENTE	varchar(150)	Representante del cliente
DESCUENTO_CLIENTE	decimal(18, 3)	Descuento al cliente
TIPO_CLIENTE	varchar(150)	Tipo de cliente
ACTIVO_CLIENTE	varchar(150)	Indicador del cliente
DIRECCION_CLIENTE	varchar(150)	Dirección del cliente
TELEFONO_CLIENTE	varchar(150)	Teléfono del cliente
CELULAR_CLIENTE	varchar(150)	Celular del cliente
LOAD_DATE	datetime	Fecha de carga

3.1.4.7. Dimensión Dim_Calendario

Tabla 60 Diseño físico de Dim_Calendario

Columna	Tipo de dato	Descripción
SK_PERIODO	int	Clave surrogada
Bimestre	nvarchar(255)	Bimestre
BimestreID	float	Id de bimestre
Date	date	Fecha
[Day (#)]	float	Día del mes
DayFin	datetime	Fecha final del mes
DayInicio	datetime	Fecha inicial del mes
LastDayMonthPrev	float	Días que trae el mes
Month	nvarchar(255)	Nombre mes
[Month (#)]	float	N° de mes en el año
MonthID	float	Id de mes
NDate	datetime	Fecha y hora
NombreDia	nvarchar(255)	Nombre y N° del día en el año
NomWeek	nvarchar(255)	Semana de año
Period	nvarchar(255)	Nombre del mes y año
[Period (#)]	nvarchar(255)	Año y mes
Periodo	nvarchar(255)	Año y mes
Quarter	nvarchar(255)	Cuarto de año
QuarterID	float	Id de cuarto
QuarterName	nvarchar(255)	Mes inicial y final de cuarto y año
QuarterName1	nvarchar(255)	Mes inicial y final de cuarto
vtoday	datetime	Fecha y hora
Week	float	Numero de semana en el año
WeekDay	nvarchar(255)	Día de la semana
Year	float	año
LOAD_DATE	datetime	Fecha de carga

3.1.4.8. Dimensión Dim_Usuario

Tabla 61 Diseño físico de Dim_Usuario

Columna	Tipo de dato	Descripción
SK_USUARIO	int	Clave surrogada
ID_USUARIO	int	Código usuario
NOMBRES_USUARIO	varchar(150)	Nombres del usuario
APELLIDOS_USUARIO	varchar(150)	Apellidos del usuario
DNI_USUARIO	varchar(150)	DNI del usuario
EMAIL_USUARIO	varchar(150)	Email del usuario
NACIMIENTO_USUARIO	datetime	Fecha de nacimiento
SEXO_USUARIO	varchar(150)	Sexo del usuario
NIVEL_ACADEMICO_USUARIO	varchar(150)	Nivel académico
INDICADOR_USUARIO	varchar(150)	Indicador de usuario
SUCURSAL_USUARIO	varchar(150)	Sucursal del usuario
LOAD_DATE	datetime	Fecha de carga

3.1.4.9. Dimensión Dim_Insumo

Tabla 62 Diseño físico de Dim_Insumo

Columna	Tipo de dato	Descripción
SK_INSUMO	int	Clave surrogada
ID_INSUMO	int	Código orden
NOMBRE_INSUMO	varchar(150)	Nombre del insumo
STOCK_INSUMO	decimal(18, 3)	Stock del insumo
UNIDAD_INSUMO	varchar(150)	Unidad del insumo
CATEGORIA_INSUMO	varchar(150)	Categoría del insumo
INDICADOR_INSUMO	varchar(150)	Indicador del insumo
TIPO_INSUMO	varchar(150)	Tipo de insumo
ABC_INSUMO	varchar(150)	Abc del insumo
LOAD_DATE	datetime	Fecha de carga

3.1.4.10. Dimensión Dim_Documento

Tabla 63 Diseño físico de Dim_Documento

Columna	Tipo de dato	Descripción
SK_DOCUMENTO	int	Clave surrogada
ID_DOCUMENTO	int	Código documento
NOMBRE_DOCUMENTO	varchar(150)	Nombre documento
INICIALES_DOCUMENTO	varchar(150)	
LOAD_DATE	datetime	Fecha de carga

3.1.4.11. Dimensión Dim_Orden

Tabla 64 Diseño físico de Dim_Orden

Columna	Tipo de dato	Descripción
SK_ORDEN	int	Clave surrogada
ID_ORDEN	int	Código formula
OBSERVACIONES	varchar(150)	Observaciones de orden
ESTADO_FINALIZADO	varchar(150)	Estado de finalización de la orden
ESTADO_TRANSFORMADO	varchar(150)	Estado de transformación de la orden
LOAD_DATE	datetime	Fecha de carga

3.1.4.12. Tabla hechos Fact_Producción

Tabla 65 Diseño físico de Fact_Produccion

Columna	Tipo de dato	Descripción
SK_ORDEN	int	Clave surrogada orden
SK_FORMULA	int	Clave surrogada formula
SK_INSUMO	int	Clave surrogada insumo
SK_PERIODO	int	Clave surrogada periodo
SK_USUARIO	int	Clave surrogada usuario
CANTIDAD_PRODUCIDA	decimal(18, 3)	Cantidad producida
LECHE_USADA	decimal(18, 3)	Leche utilizada
CANTIDAD_INSUMO	decimal(18, 2)	Cantidad insumo usada
LOAD_DATE	datetime	Fecha de carga

3.1.4.13.Tabla hechos Fact_Comercialización

Tabla 66 Diseño físico de Fact_Comercializacion

Columna	Tipo de dato	Descripción
SK_PERIODO	int	Clave surrogada periodo
SK_SUCURSAL	int	Clave surrogada sucursal
SK_DOCUMENTO	int	Clave surrogada documento
SK_PROVEEDOR	int	Clave surrogada proveedor
SK_CLIENTE	int	Clave surrogada cliente
SK_USUARIO	int	Clave surrogada usuario
SK_PRODUCTO	int	Clave surrogada producto
SK_VENTA	int	Código de venta
CANTIDAD	decimal(18, 3)	Cantidad del producto
PRECIOPRODUCTO	decimal(18, 2)	Precio del producto
PRECIOVENTA	decimal(18, 2)	Precio al que se vendió el producto
ESTADO_PRODUCTO_VENDIDO	varchar(150)	Estado del producto en la venta
LOAD_DATE	datetime	Fecha de carga

3.1.5.Microsoft SQL Server Integration Services (SSIS)

Es una plataforma para la creación de soluciones de integración y de transformación de datos. Utilizar servicios de integración para resolver problemas complejos de negocios mediante la copia o la descarga de archivos, el envío de mensajes de correo electrónico en respuesta a eventos, la actualización de los almacenes de datos, limpieza y extracción de datos y la gestión de objetos y datos de SQL Server. Los paquetes pueden trabajar solos o en concierto con otros paquetes para hacer frente a las necesidades empresariales complejas. Integration Services pueden extraer y transformar datos de una amplia variedad de fuentes, tales como archivos XML de datos, archivos planos, y fuentes de datos relacionales, y luego cargar los datos en uno o más destinos [32].

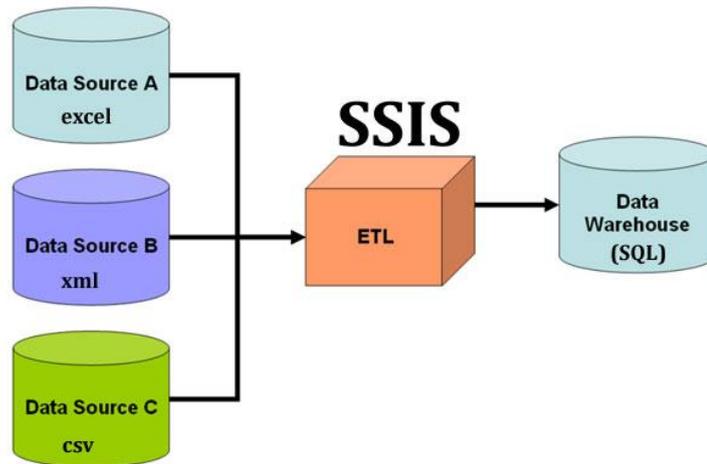


Figura 41 Microsoft SQL Server Integration Services (SSIS) [33]

3.1.6. Diseño e implementación del subsistema de ETL en SSIS

Dada la cantidad de transformaciones que se han de realizar, es recomendable crear una arquitectura que se encuentre dividida en tres capas, por las que han de pasar los datos en su camino desde el sistema origen hacia el data warehouse, esto permitirá garantizar la integridad de la información sumada con la limpieza de los datos. Para la extracción de datos se dispone de un origen de datos, el cual se encuentra en SQL Server.

3.1.5.1. Capa stage

Es un área temporal donde se recogen los datos que se necesitan del sistema origen. Se recogen los datos de las tablas estrictamente necesarias para las cargas, y se aplica el mínimo de transformaciones. No se aplican restricciones de integridad ni se utilizan claves, los datos se tratan como si las tablas fueran ficheros planos y se le añade un prefijo al nombre de las tablas para diferenciarlas de las originales. De esta manera se minimiza la afectación a los sistemas origen y la carga suele ser rápida. Una vez que los datos están traspasados, el data warehouse se independiza del sistema origen hasta la siguiente carga. Lo único que se suele añadir es un campo que almacene la fecha de la carga. Estos datos no van a dar servicio a ninguna aplicación de reportes, son datos temporales que una vez hayan cumplido su función serán eliminados [34].

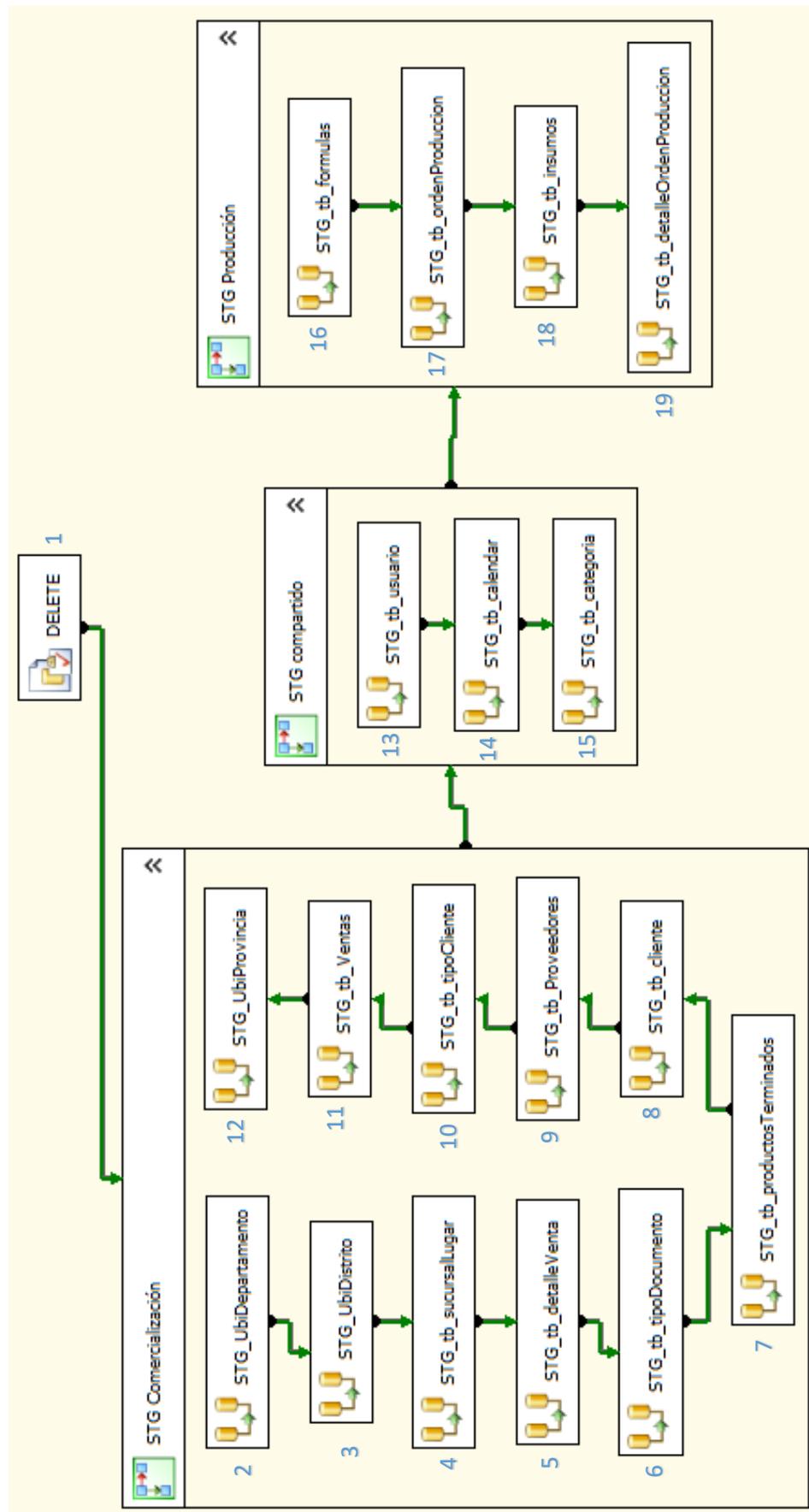


Figura 42 Capa stage

Ejemplo de cómo se realiza la carga a la tabla stage STG_tb_usuario desde la tabla origen tb_usuario en el flujo de datos STG_tb_usuario de la imagen anterior:

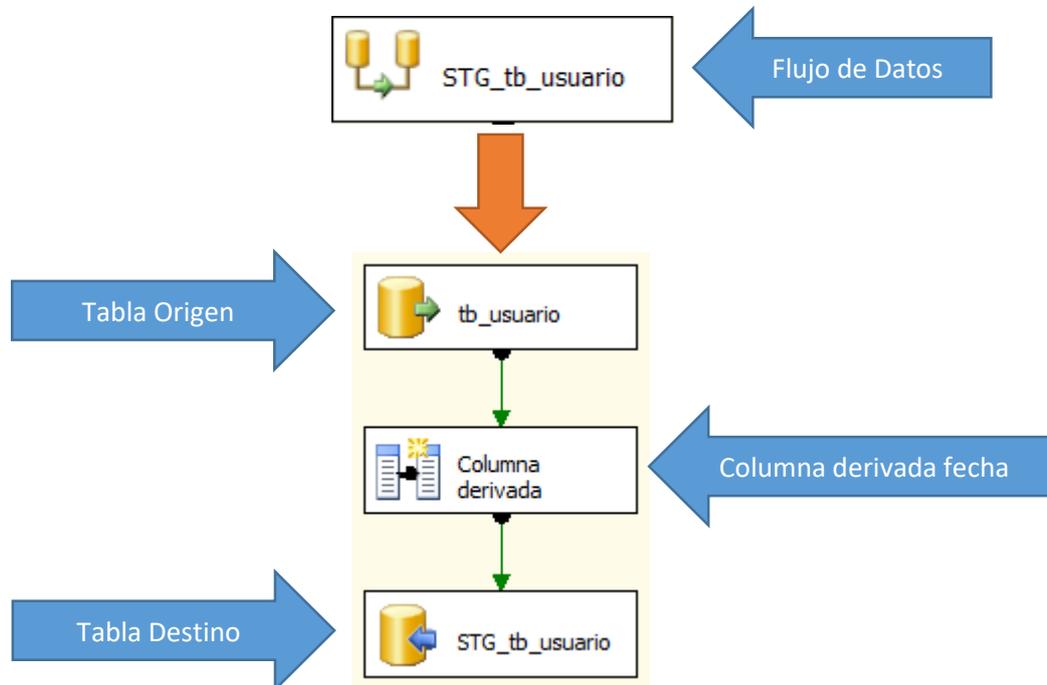


Figura 43 Flujo de datos STG_tb_usuario

De esta misma manera, se realiza la carga en todos los flujos de datos de la capa stage, consumiendo la tabla original tal cual y añadiéndole un campo que indique la fecha de carga.

3.1.5.2. Capa ODS (Operational Data Stores)

Es la capa intermedia que servirá como almacenamiento entre los sistemas fuente y el Data Warehouse. Su misión es ofrecer la posibilidad de gestionar los datos con el formato y estructura de origen para facilitar la integración y transformación hacia el modelo desnormalizado que poseen los data marts y el data warehouse. Hay que tener en cuenta que los sistemas fuente presentan toda la información separada por tablas, mientras que en el destino la información se compila en una estructura común, que facilita en gran medida, no sólo las búsquedas, sino también la comprensión de los datos que contienen [35].

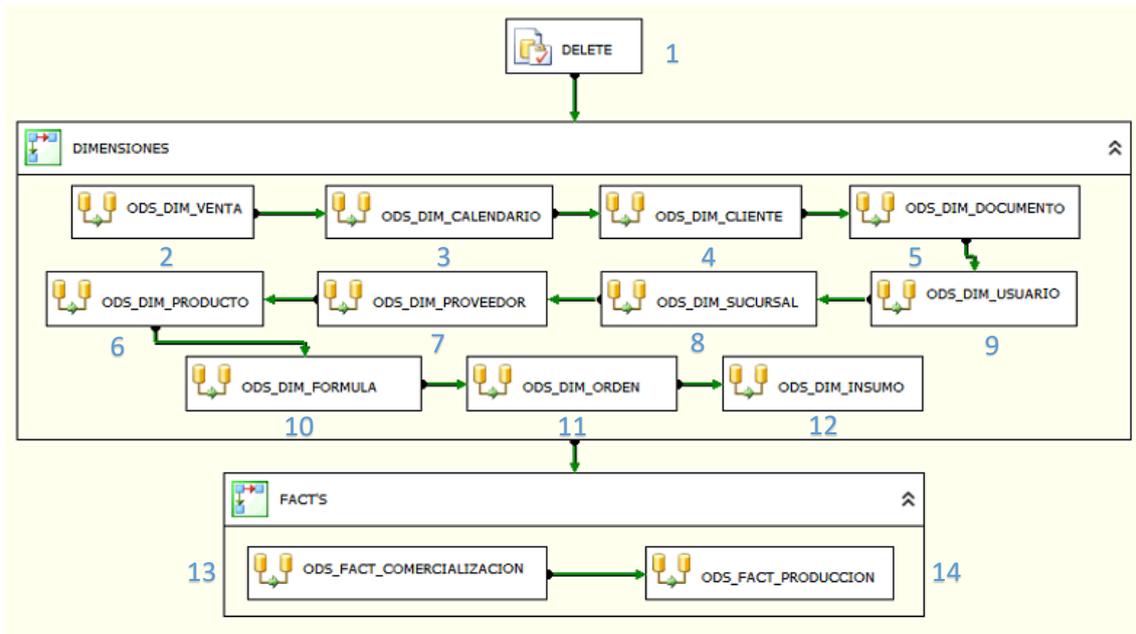


Figura 44 Capa ODS (Operational Data Store)

ODS_DIM_VENTA

Orígenes:

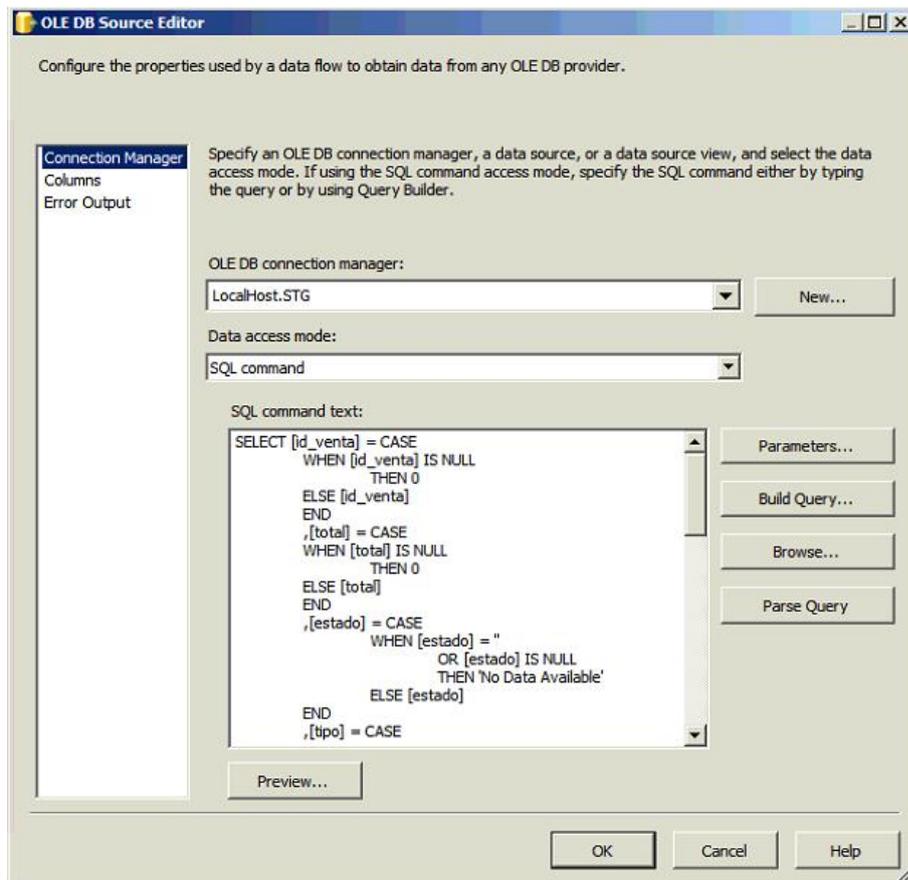


Figura 45 ODS_DIM_VENTA - Origen STG_tb_Ventas

Flujo de datos:

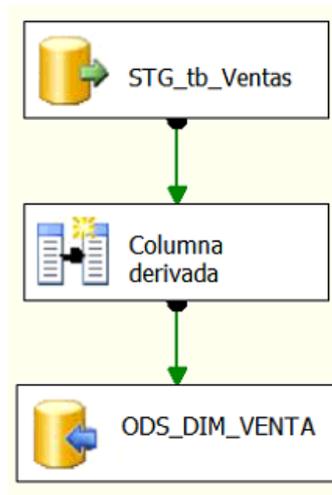


Figura 46 ODS_DIM_VENTA - Flujo de Datos

Destino:

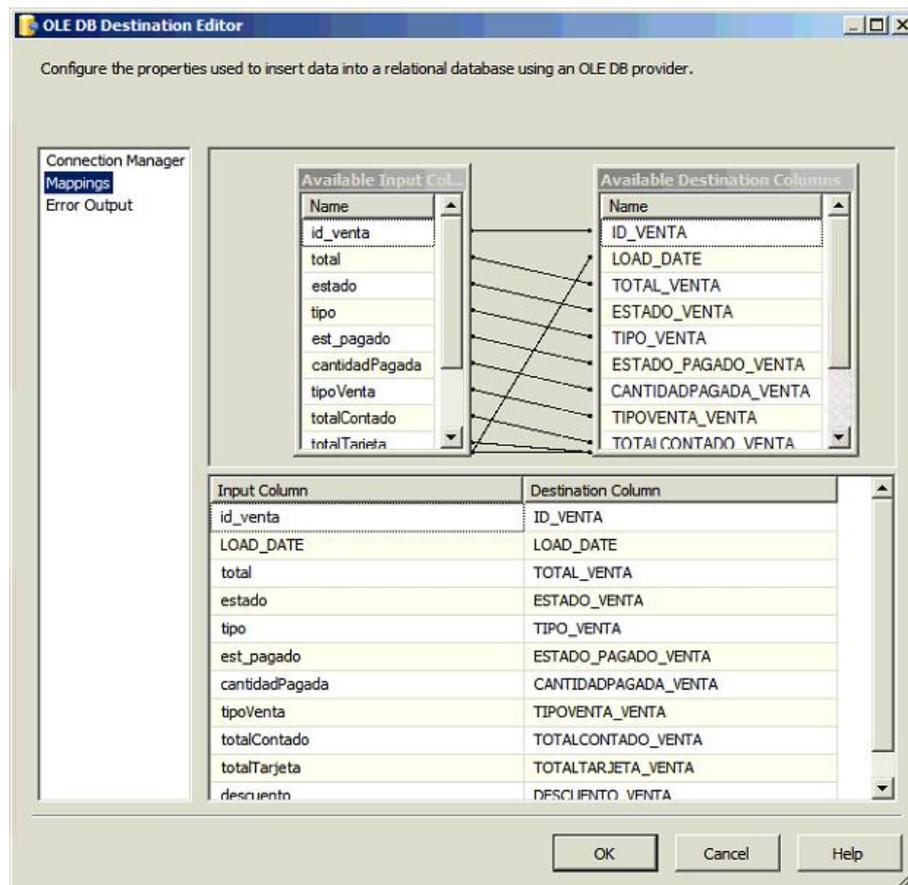


Figura 47 ODS_DIM_VENTA - Destino ODS_DIM_VENTA

ODS_DIM_CALENDARIO

Orígenes:

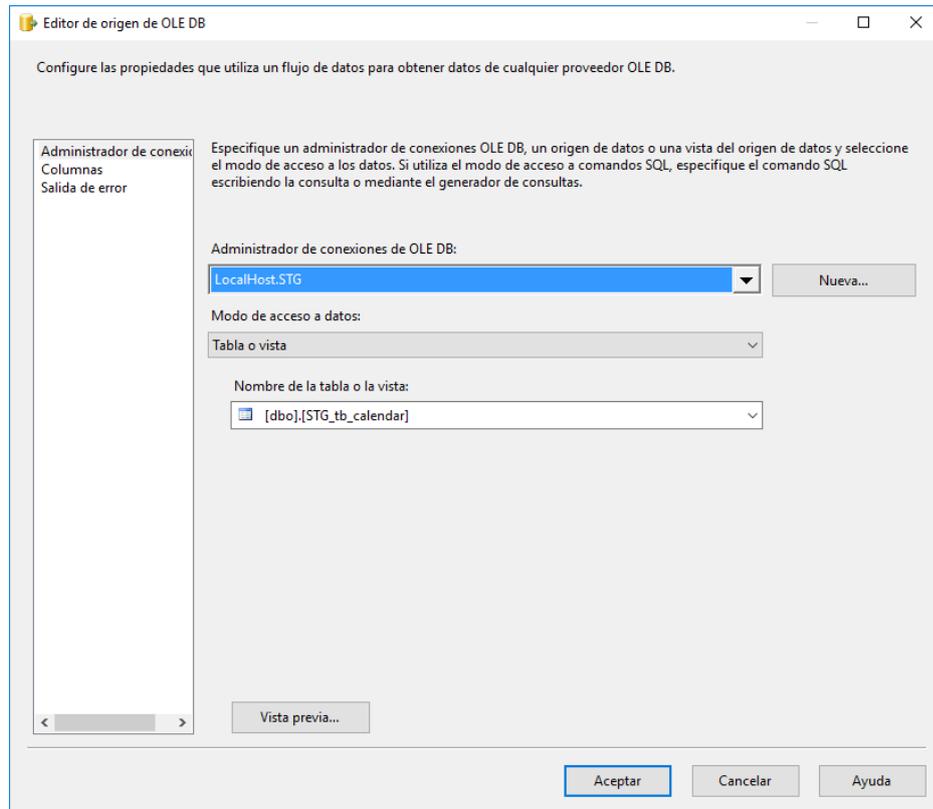


Figura 48 ODS_DIM_CALENDARIO - Origen STG_tb_calendar

Flujo de datos:

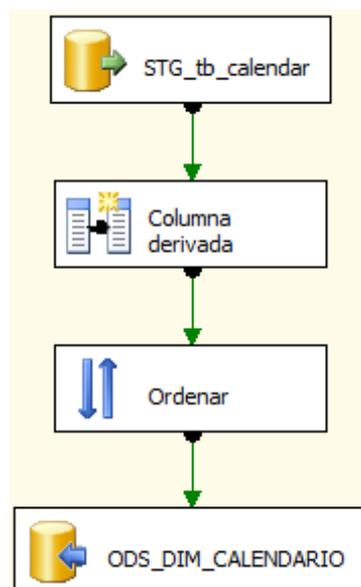


Figura 49 ODS_DIM_CALENDARIO - Flujo de Datos

Destino:

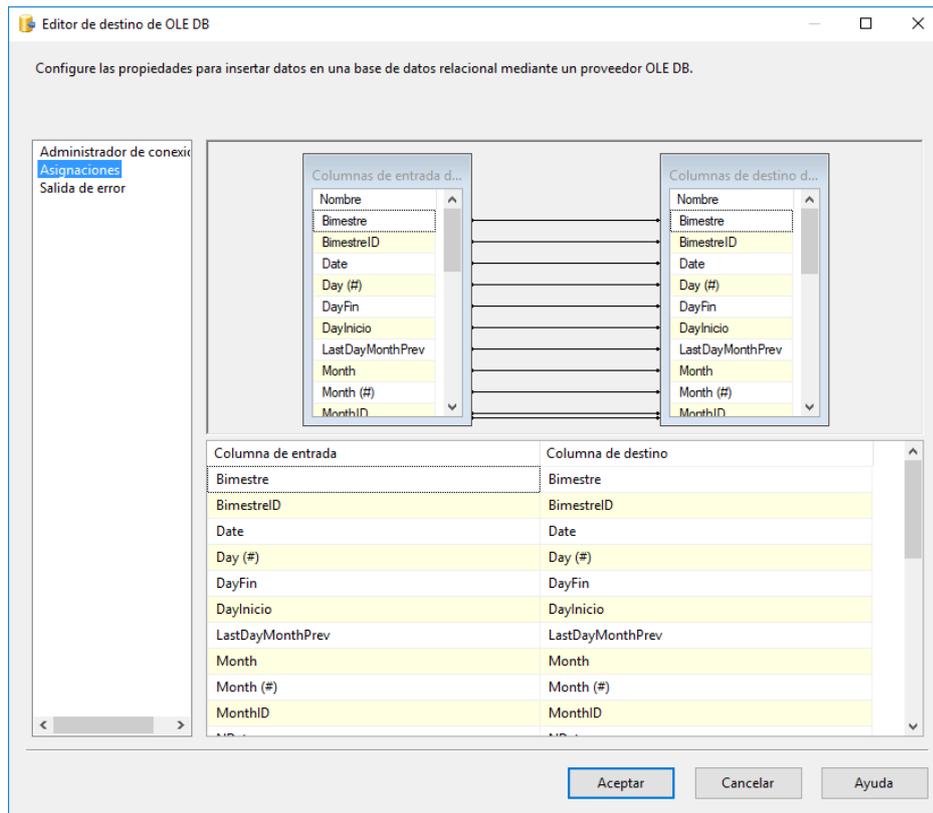


Figura 50 ODS_DIM_CALENDARIO - Destino ODS_DIM_CALENDARIO

ODS_DIM_CLIENTE

Orígenes:

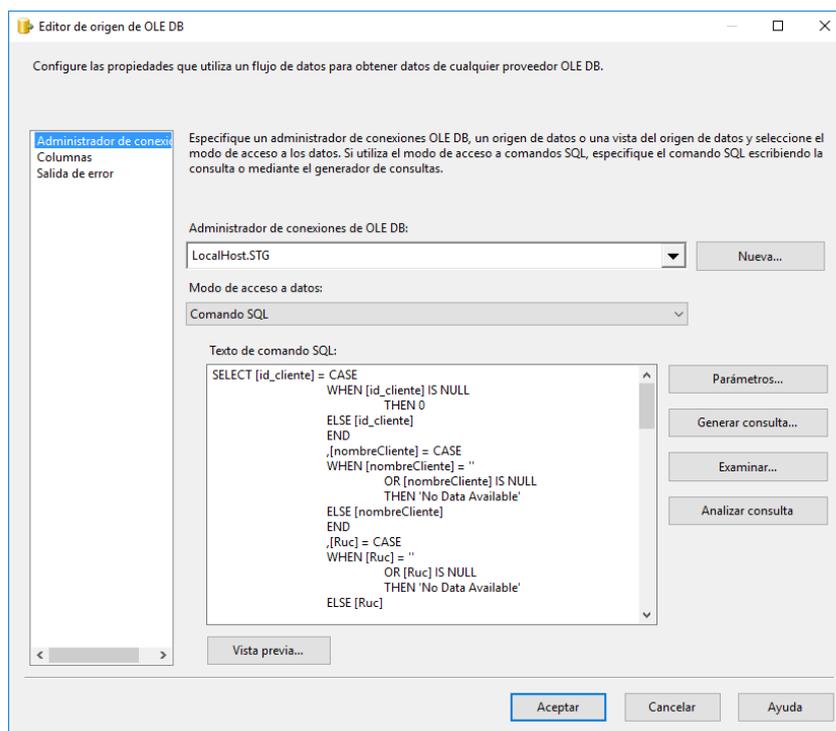


Figura 51 ODS_DIM_CLIENTE - Origen STG_tb_cliente

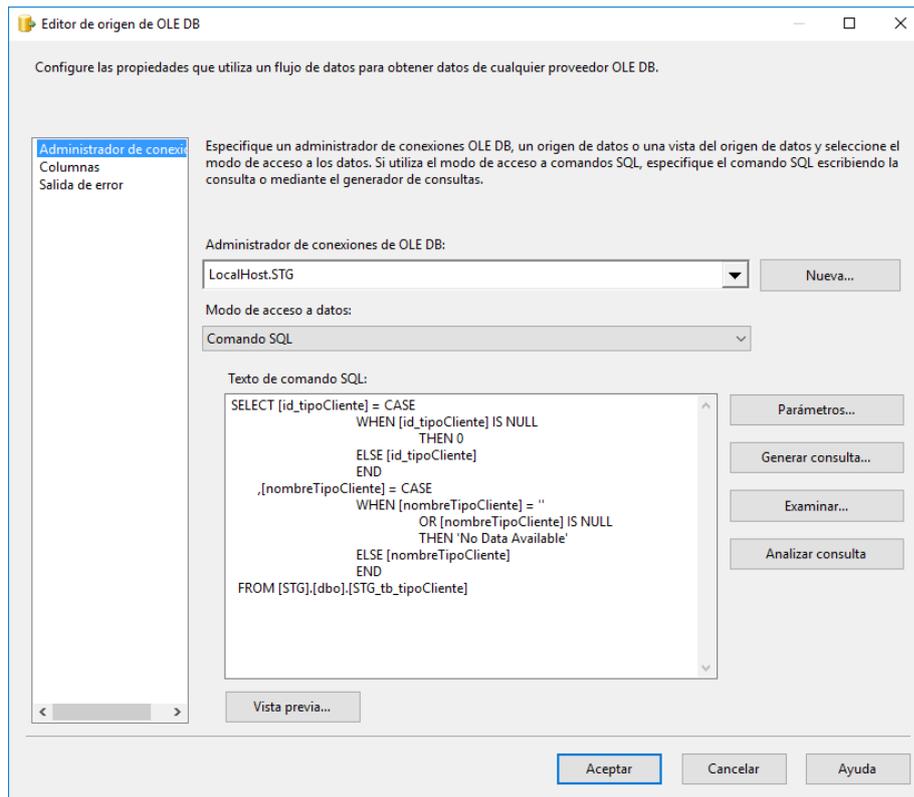


Figura 52 ODS_DIM_CLIENTE - Origen STG_tb_tipoCliente

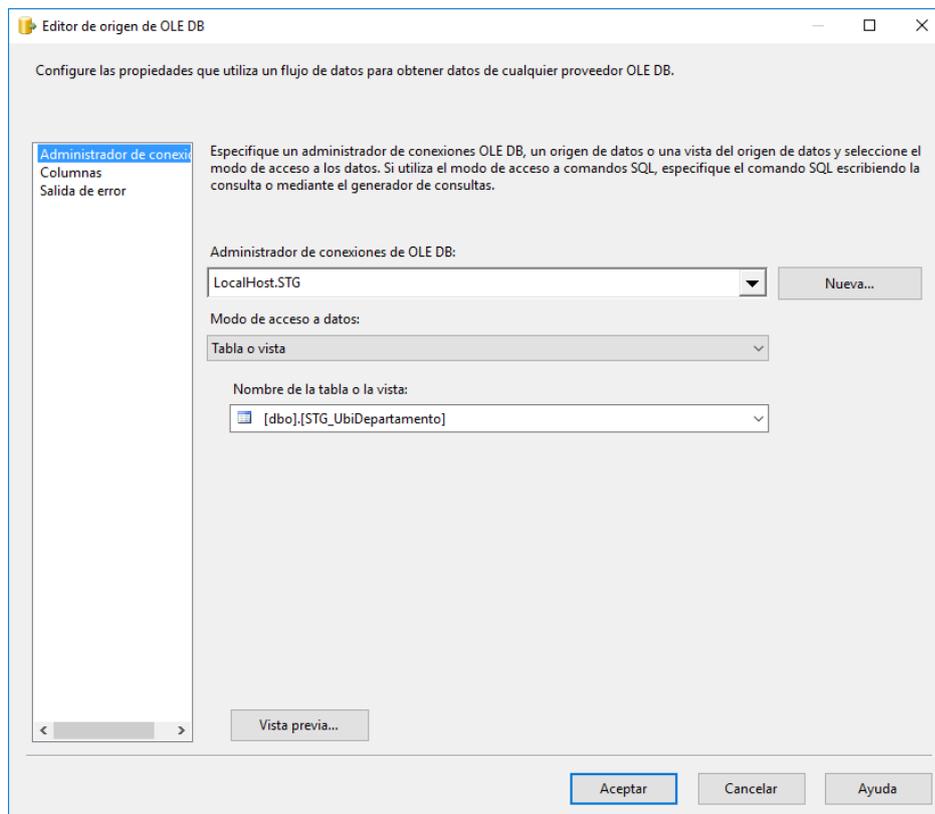


Figura 53 ODS_DIM_CLIENTE - Origen STG_UbiDepartamento

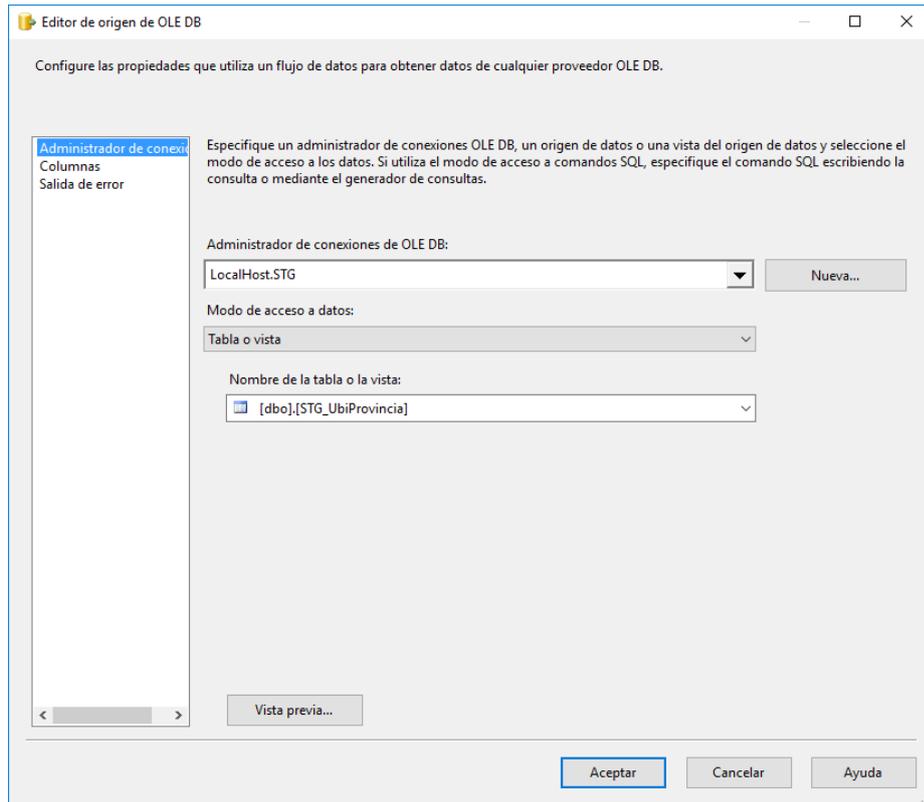


Figura 54 ODS_DIM_CLIENTE - Origen STG_UbiProvincia

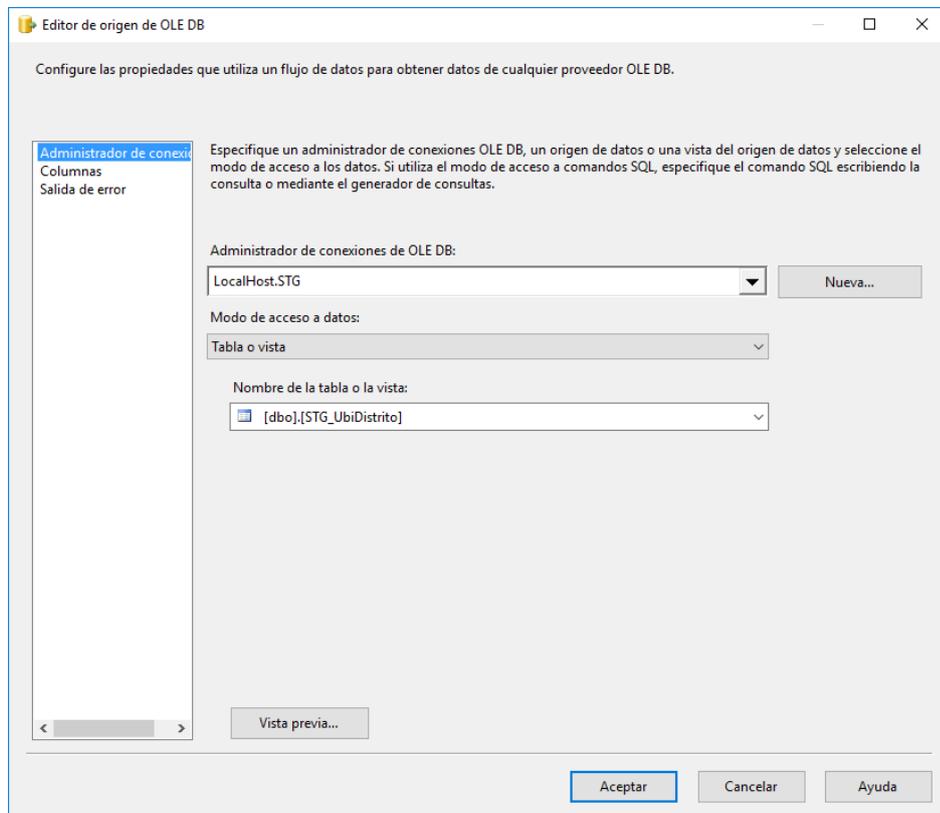


Figura 55 ODS_DIM_CLIENTE - Origen STG_UbiDistrito

Flujo de datos:

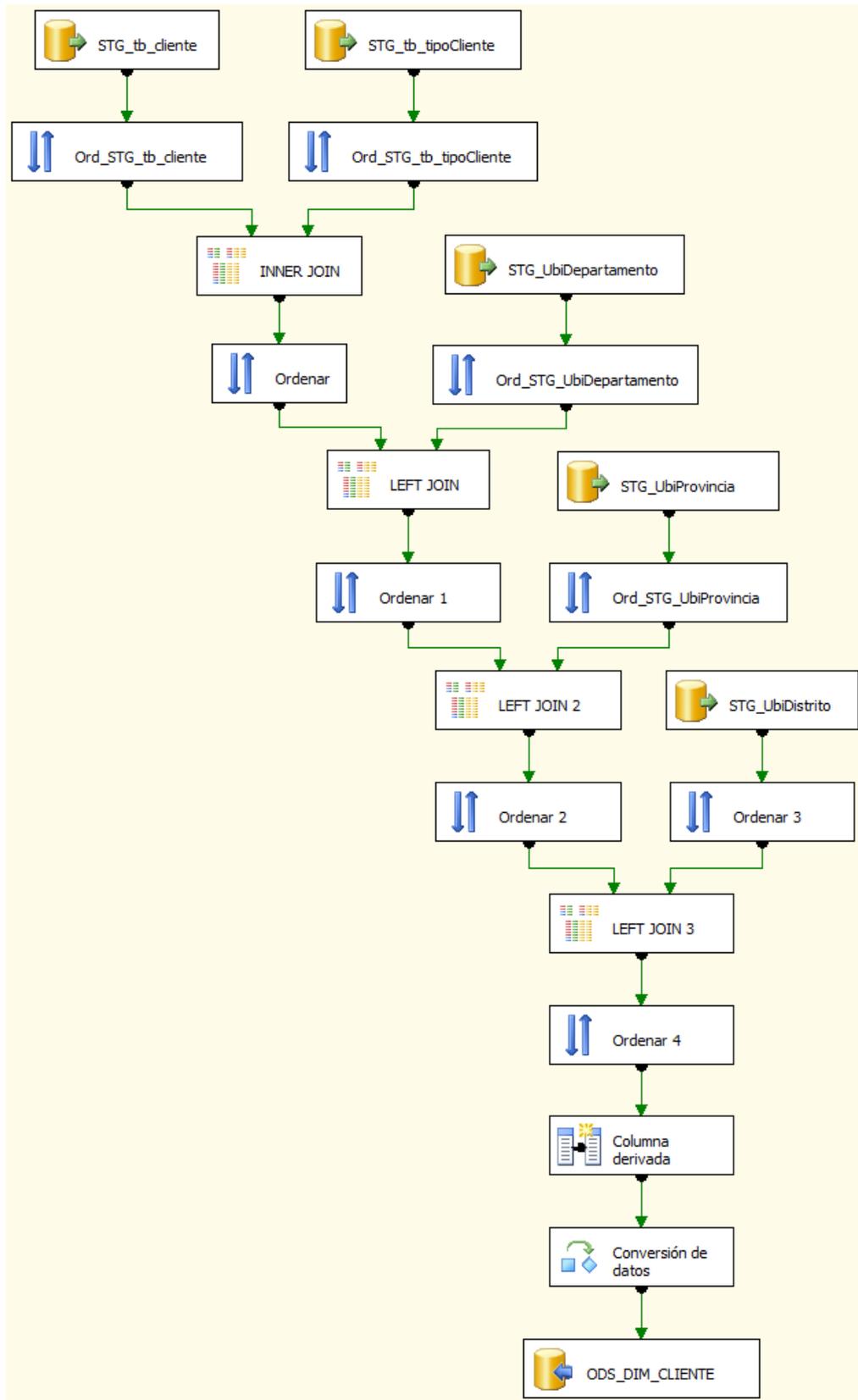


Figura 56 ODS_DIM_CLIENTE - Flujo de Datos

Destino:

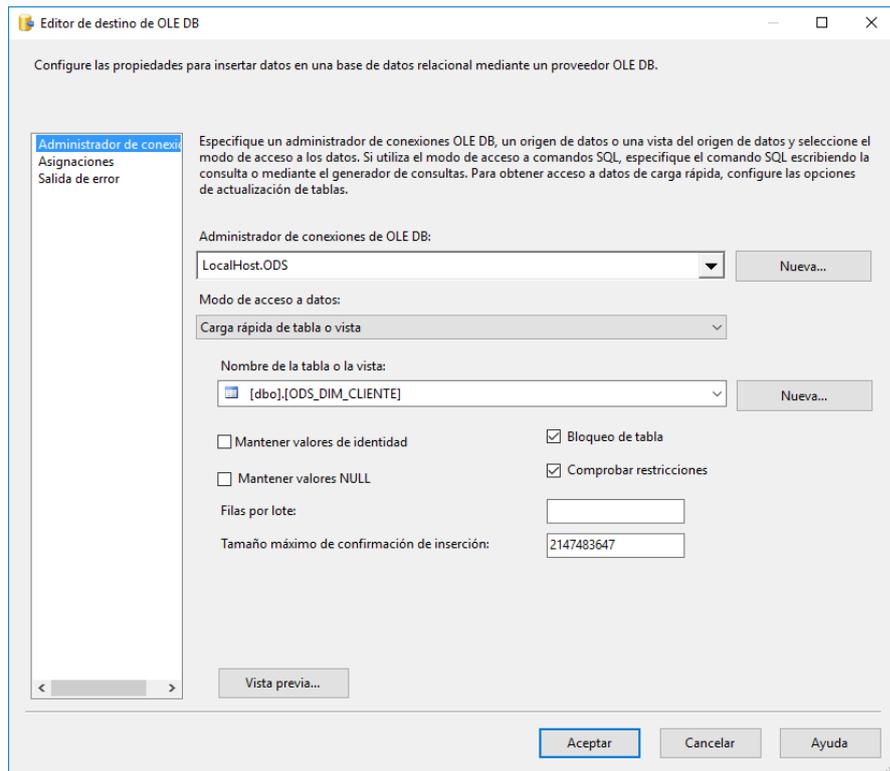


Figura 57 ODS_DIM_CLIENTE - Destino ODS_DIM_CLIENTE

ODS_DIM_DOCUMENTO

Orígenes:

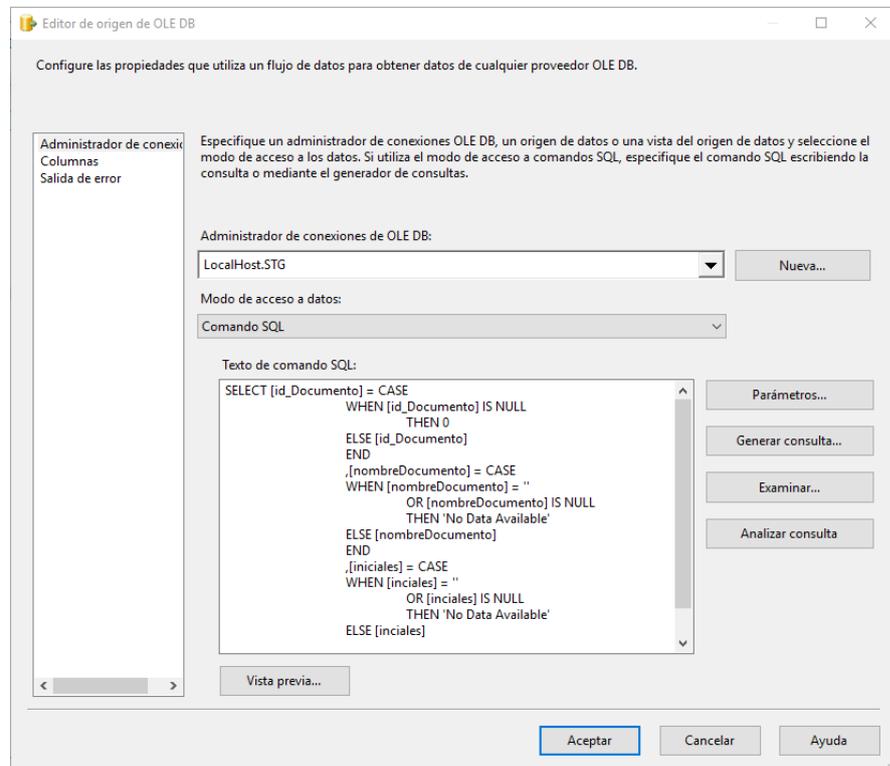


Figura 58 ODS_DIM_DOCUMENTO - Origen STG_tb_tipoDocumento

Flujo de datos:

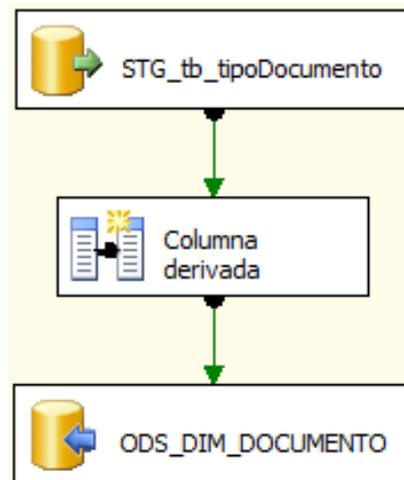


Figura 59 ODS_DIM_DOCUMENTO - Flujo de Datos

Destino:

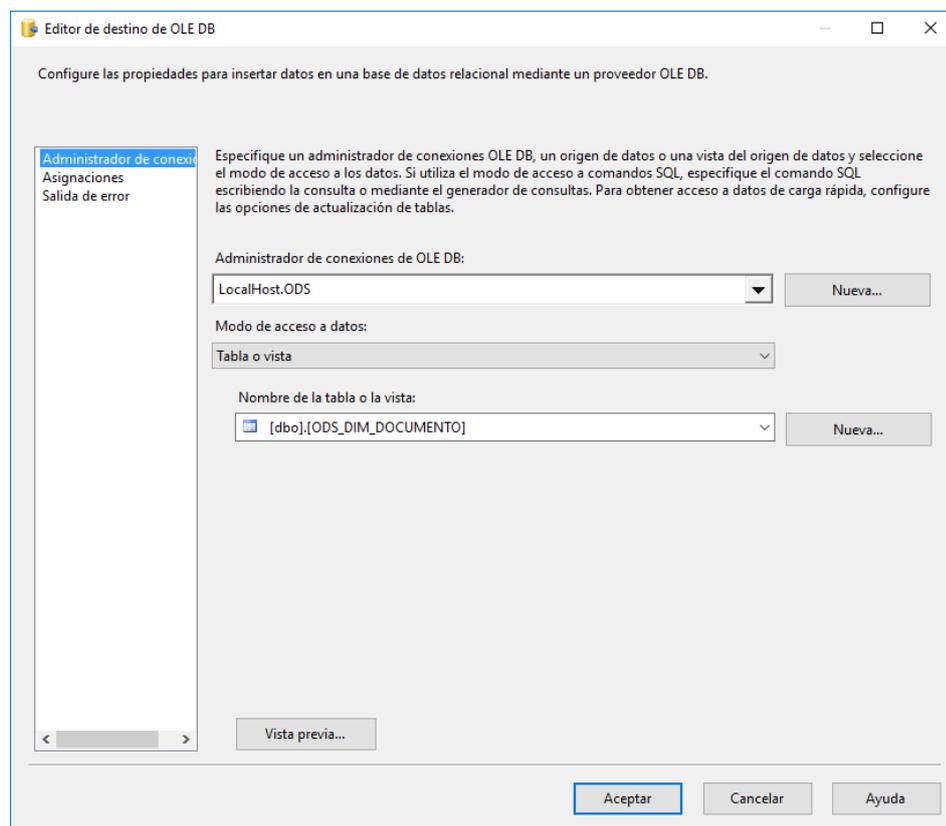


Figura 60 ODS_DIM_DOCUMENTO - Destino ODS_DIM_DOCUMENTO

ODS_DIM_PRODUCTO

Orígenes:

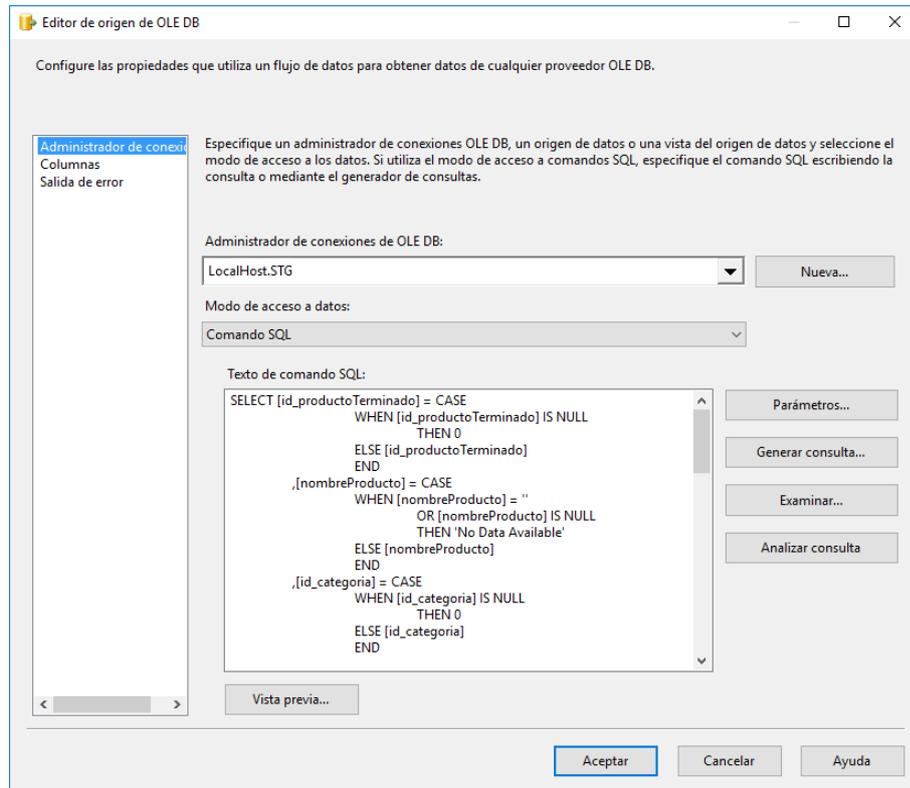


Figura 61 ODS_DIM_PRODUCTO - Origen STG_tb_productosTerminados

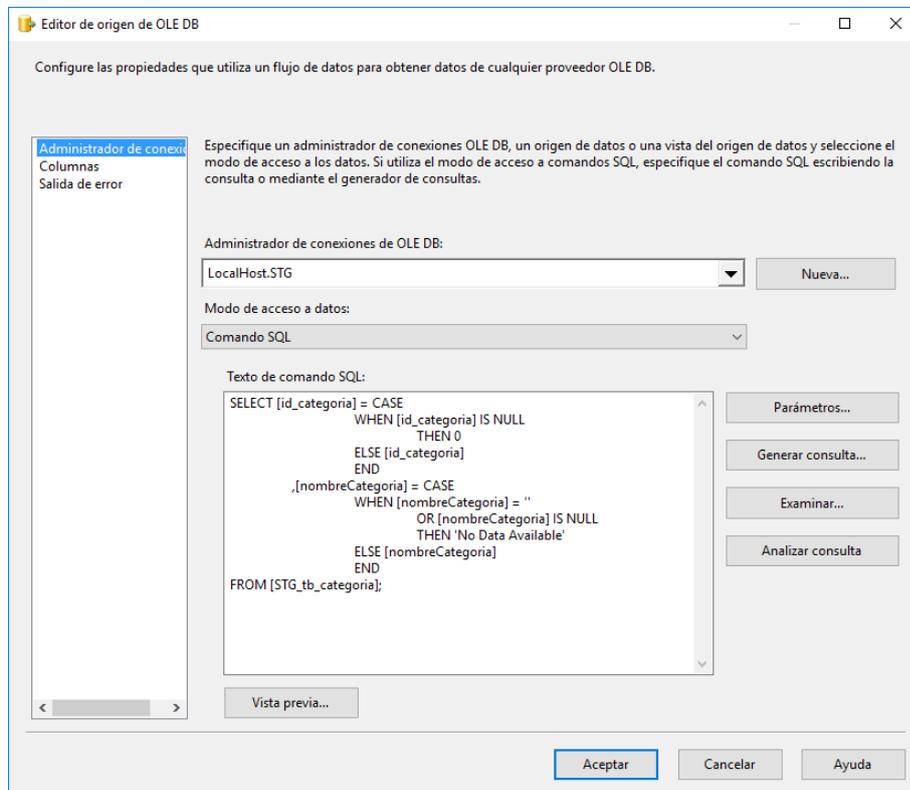


Figura 62 ODS_DIM_PRODUCTO - Origen STG_tb_categoria

Flujo de datos:

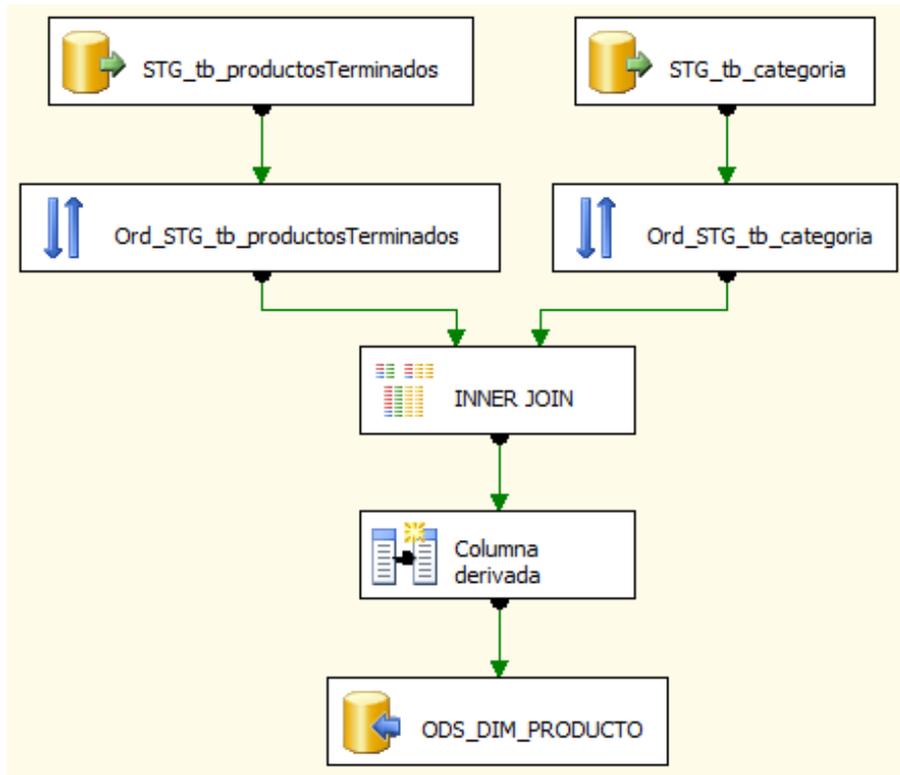


Figura 63 ODS_DIM_PRODUCTO - Flujo de Datos

Destino:

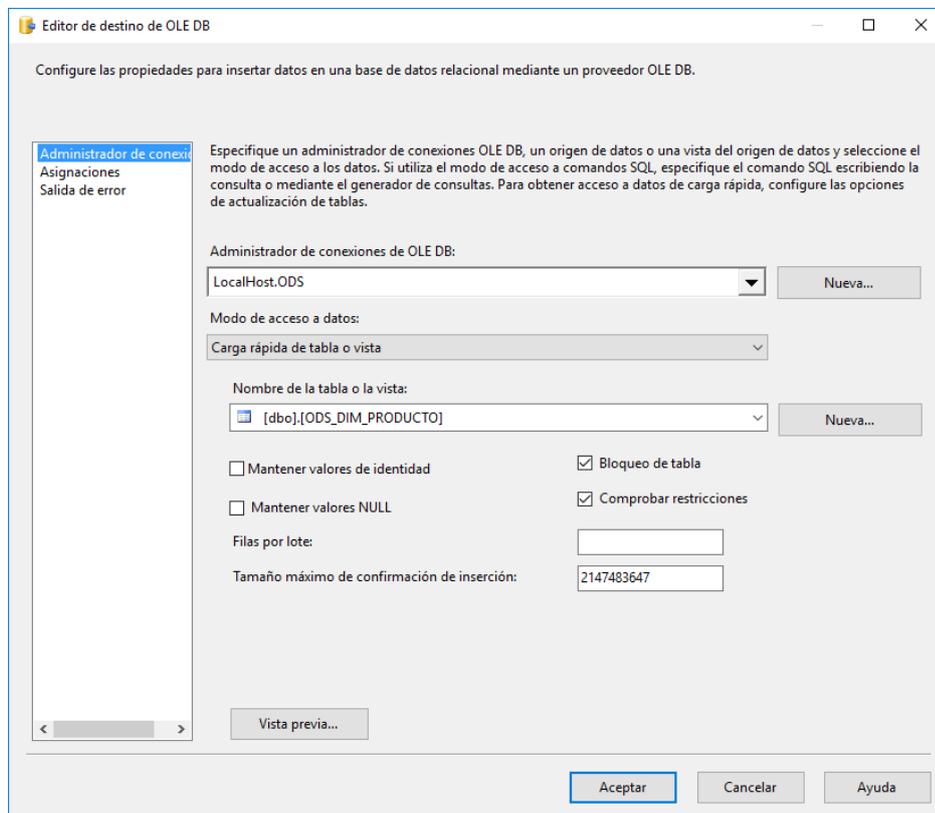


Figura 64 ODS_DIM_PRODUCTO - Destino ODS_DIM_PRODUCTO

ODS_DIM_PROVEEDOR

Orígenes:

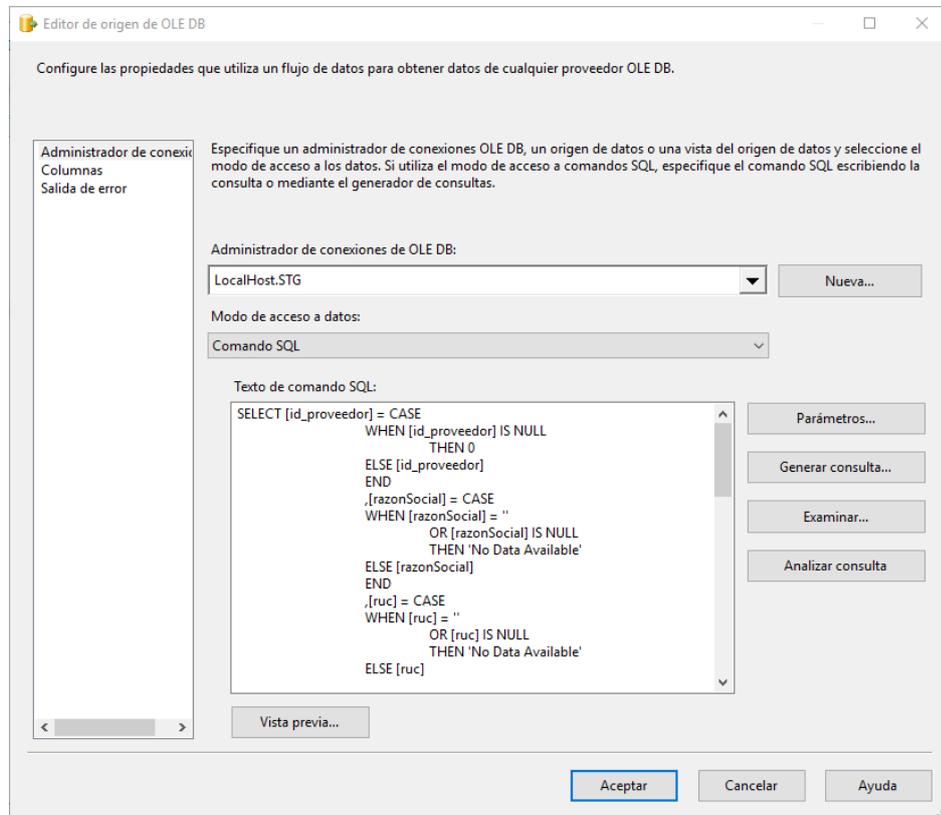


Figura 65 ODS_DIM_PROVEEDOR - Origen STG_tb_Proveedores

Flujo de datos:

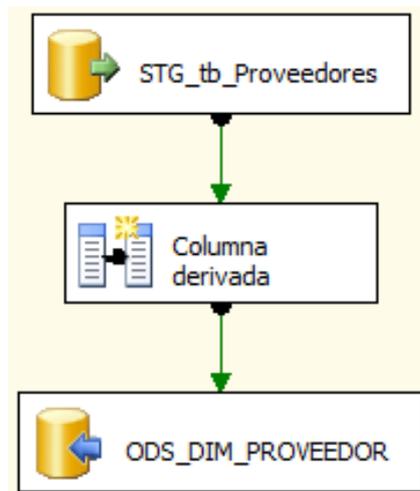


Figura 66 ODS_DIM_PROVEEDOR - Flujo de Datos

Destino:

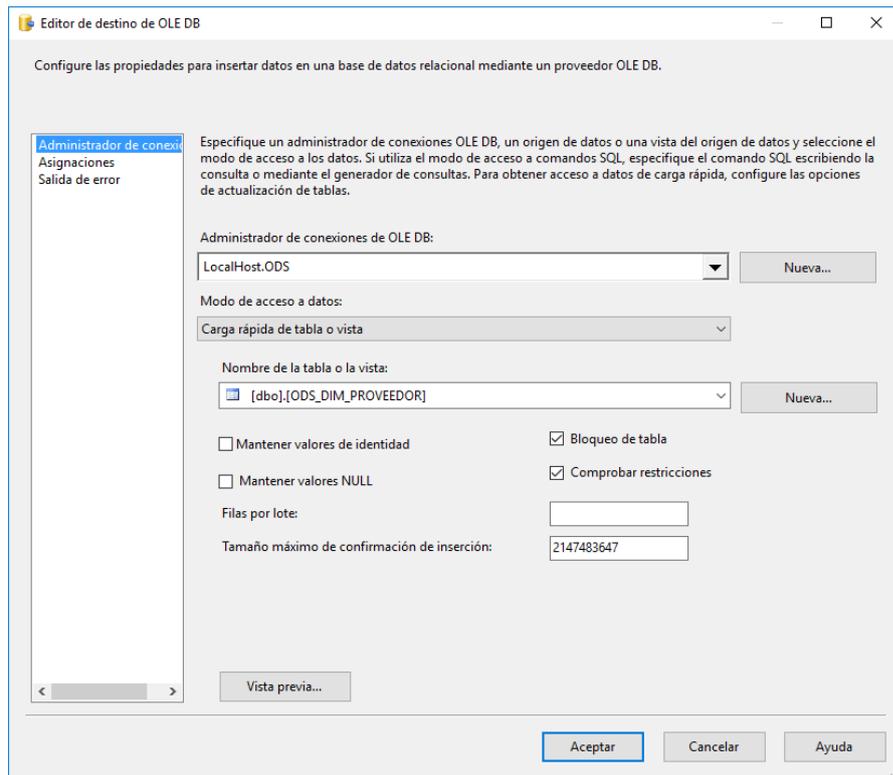


Figura 67 ODS_DIM_PROVEEDOR - Destino ODS_DIM_PROVEEDOR

ODS_DIM_SUCURSAL

Orígenes:

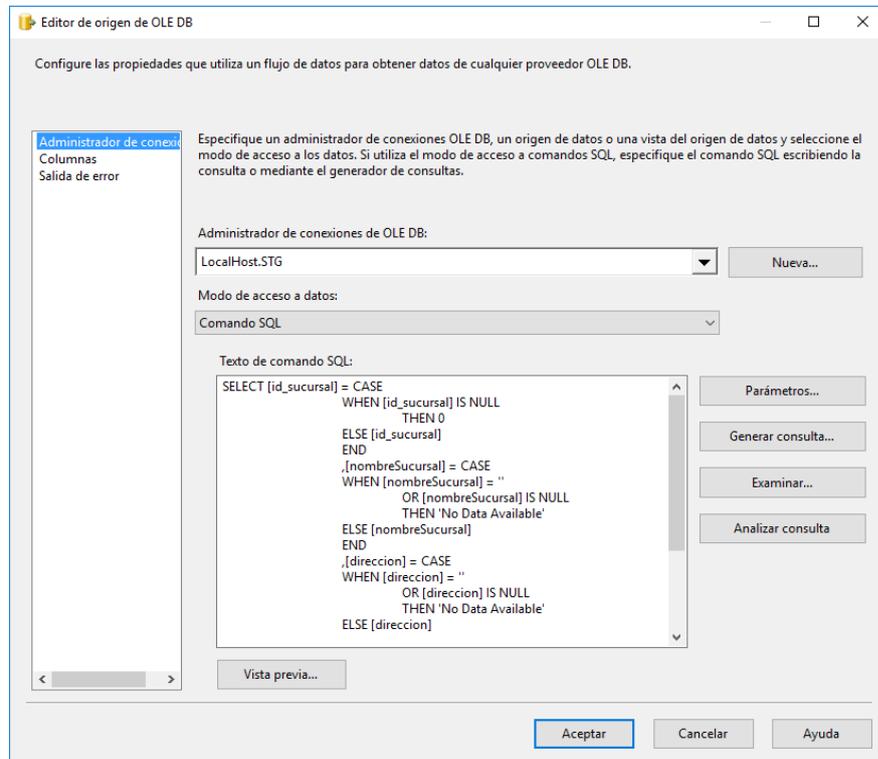


Figura 68 ODS_DIM_SUCURSAL - Origen STG_tb_sucursalLugar

Flujo de datos:

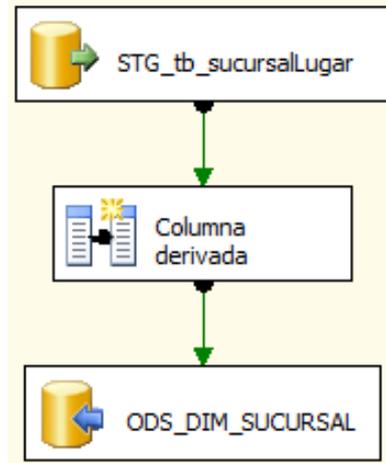


Figura 69 ODS_DIM_SUCURSAL - Flujo de Datos

Destino:

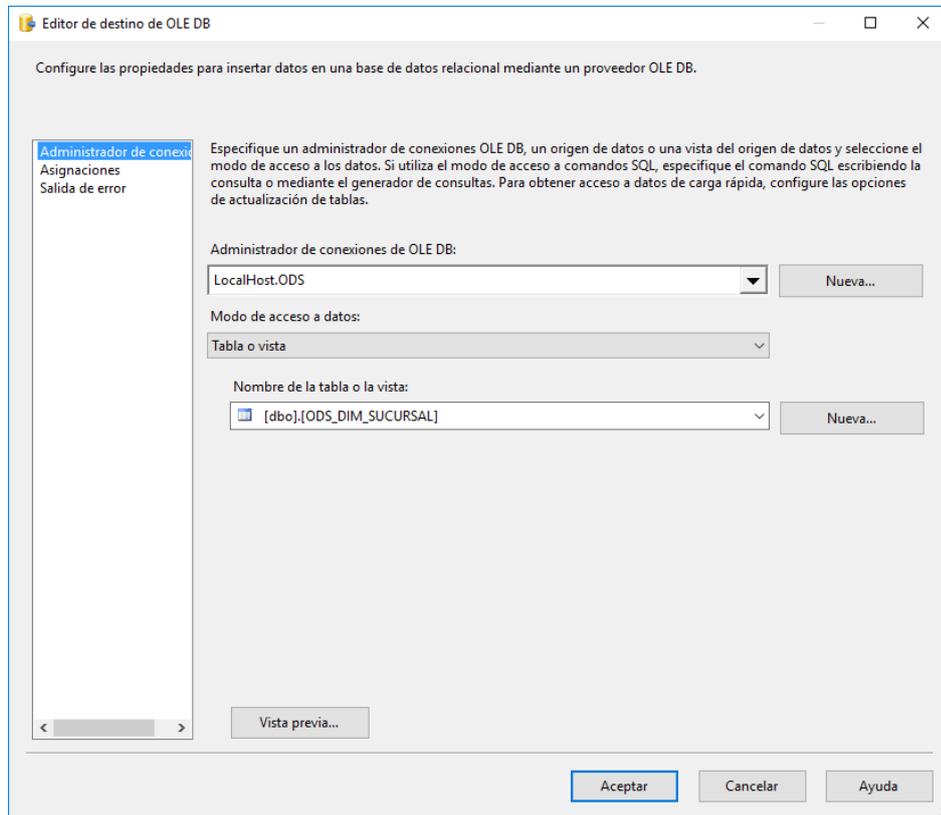


Figura 70 ODS_DIM_SUCURSAL - Destino ODS_DIM_SUCURSAL

ODS_DIM_USUARIO

Orígenes:

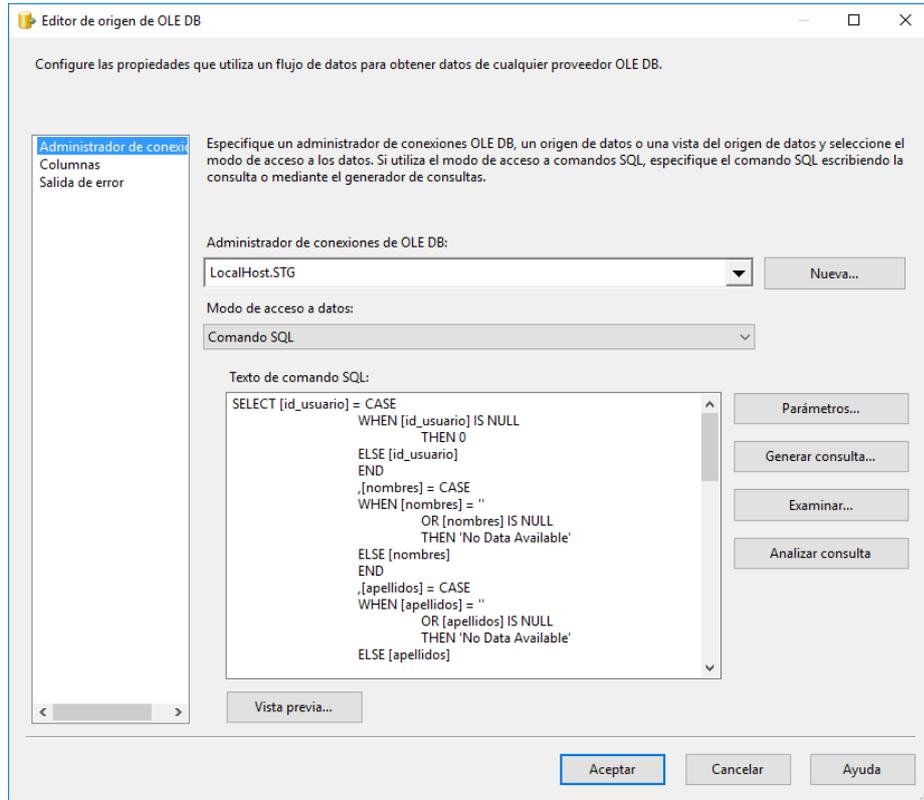


Figura 71 ODS_DIM_USUARIO - Origen STG_tb_usuario

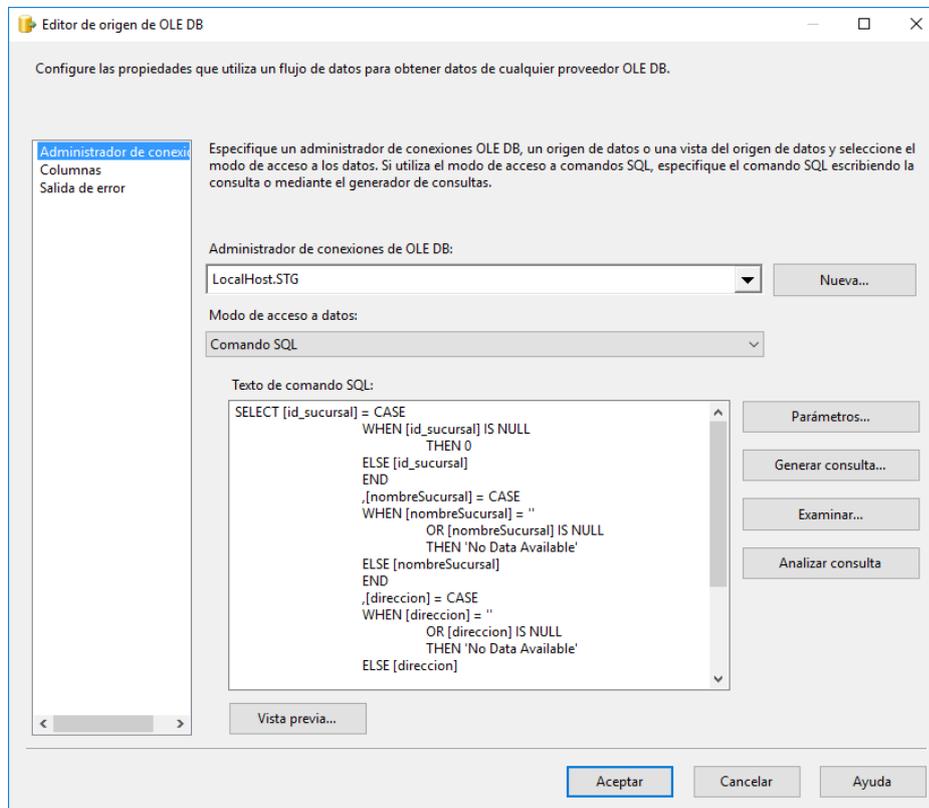


Figura 72 ODS_DIM_USUARIO - Origen STG_tb_sucursalLugar

Flujo de datos:

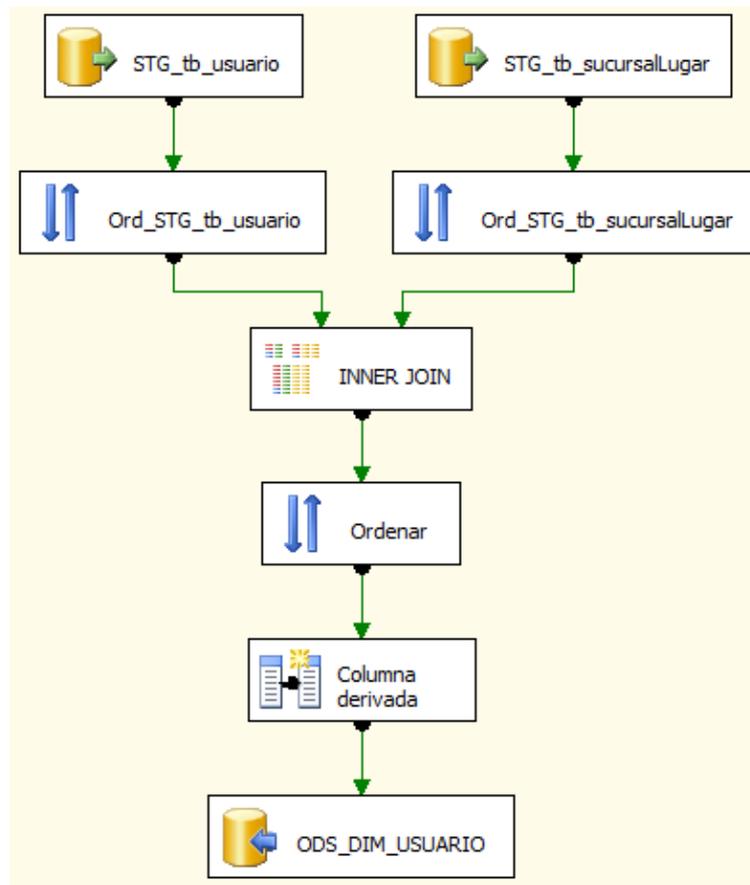


Figura 73 ODS_DIM_USUARIO - Flujo de Datos

Destino:

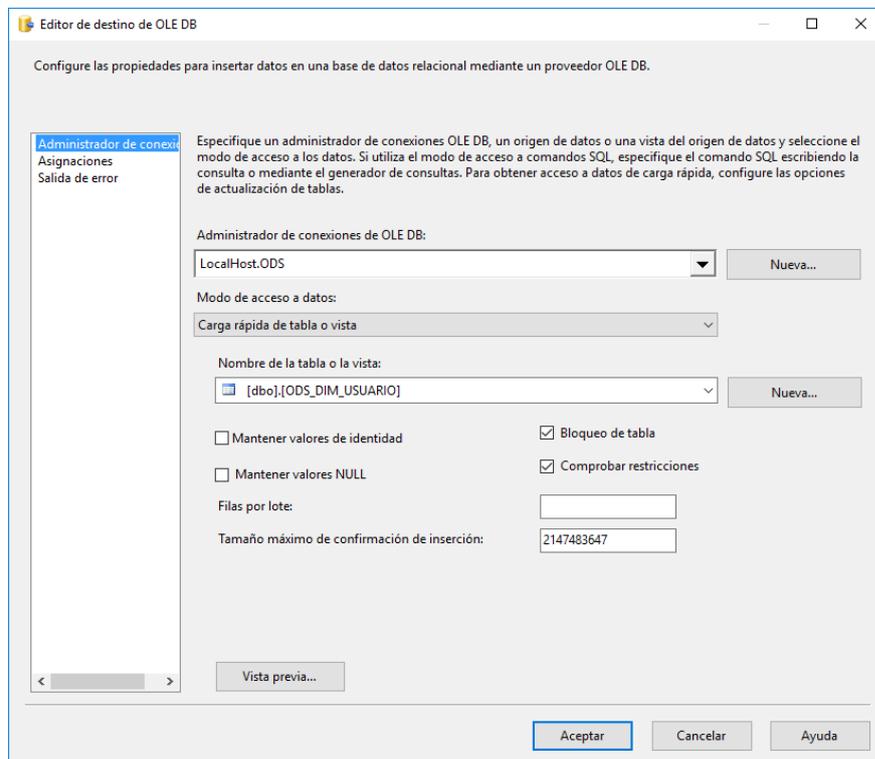


Figura 74 ODS_DIM_USUARIO - Destino ODS_DIM_USUARIO

ODS_DIM_FORMULA

Orígenes:

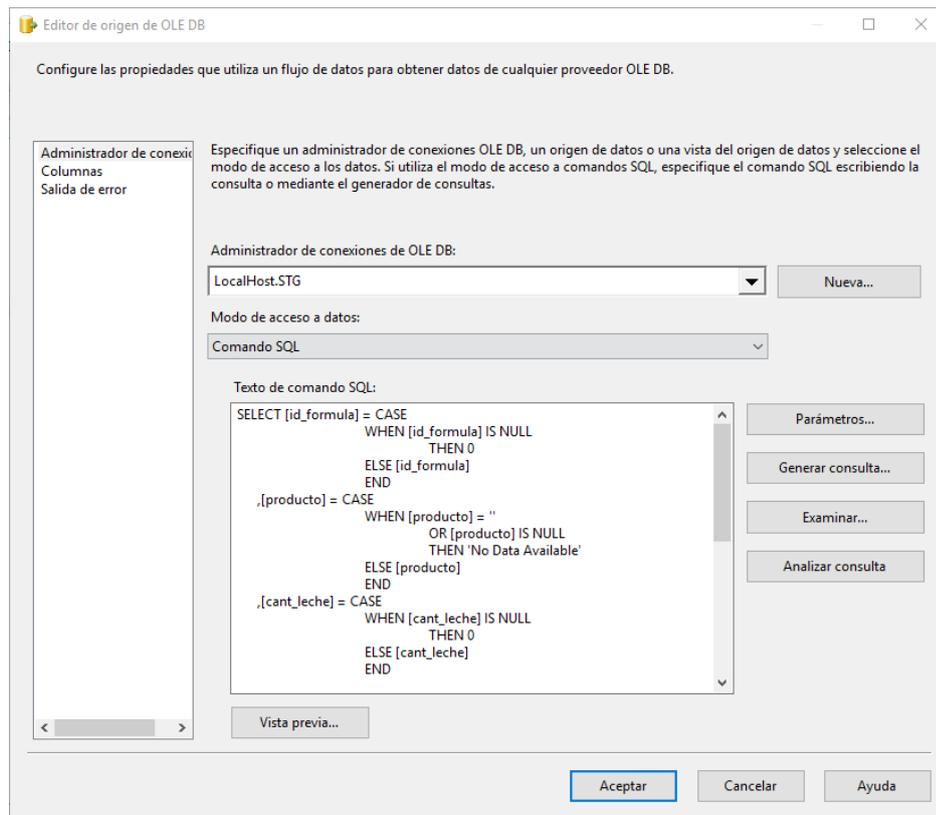


Figura 75 ODS_DIM_FORMULA - Origen STG_tb_formulas

Flujo de datos:

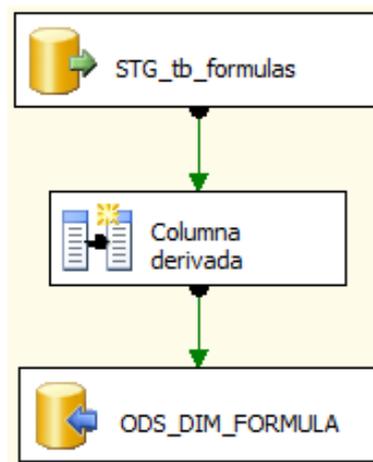


Figura 76 ODS_DIM_FORMULA - Flujo de Datos

Destino:

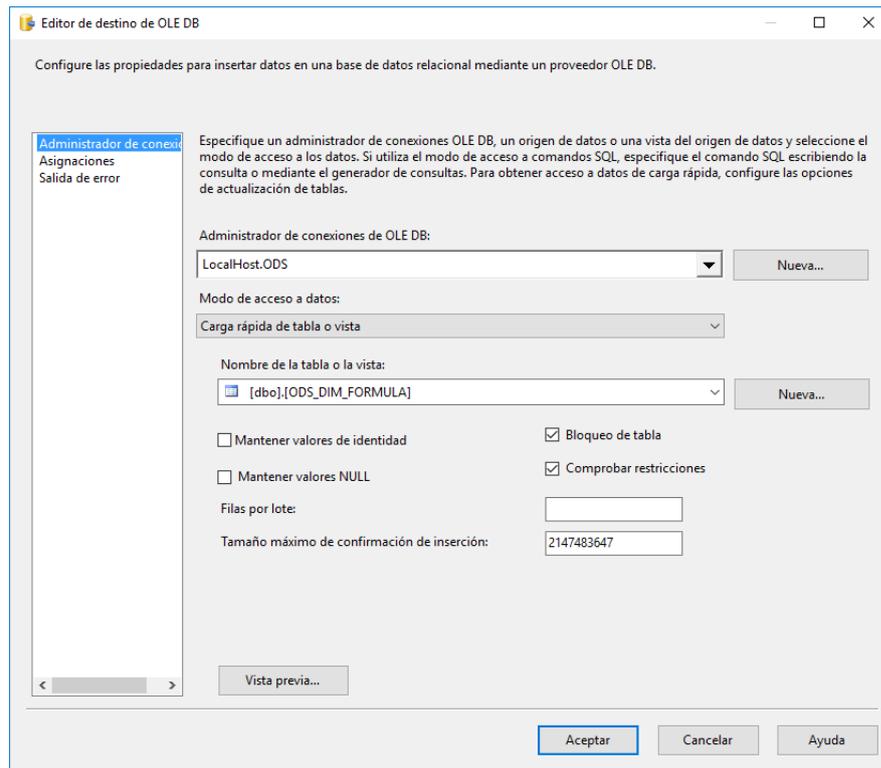


Figura 77 ODS_DIM_FORMULA - Destino ODS_DIM_FORMULA

ODS_DIM_ORDEN

Orígenes:

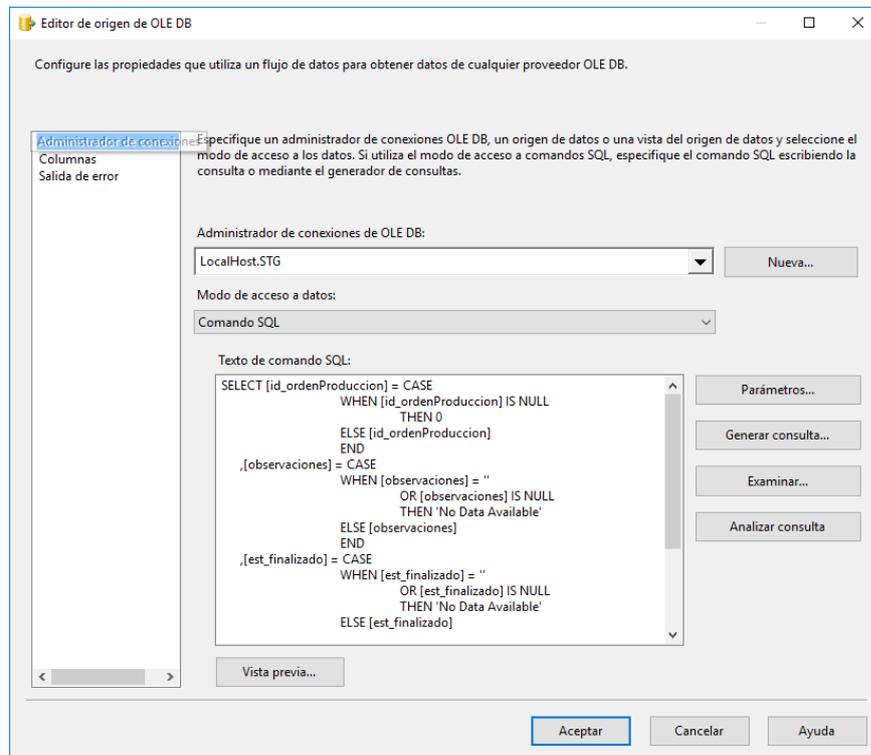


Figura 78 ODS_DIM_ORDEN - Origen STG_tb_ordenProduccion

Flujo de datos:

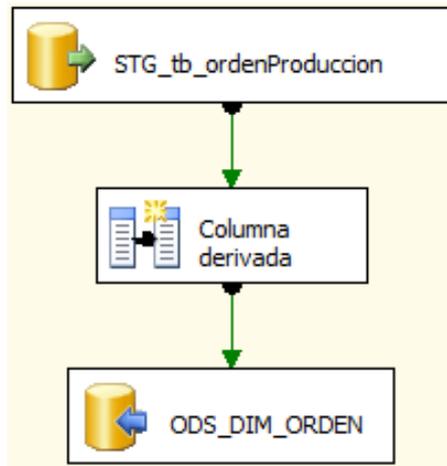


Figura 79 ODS_DIM_ORDEN - Flujo de Datos

Destino:

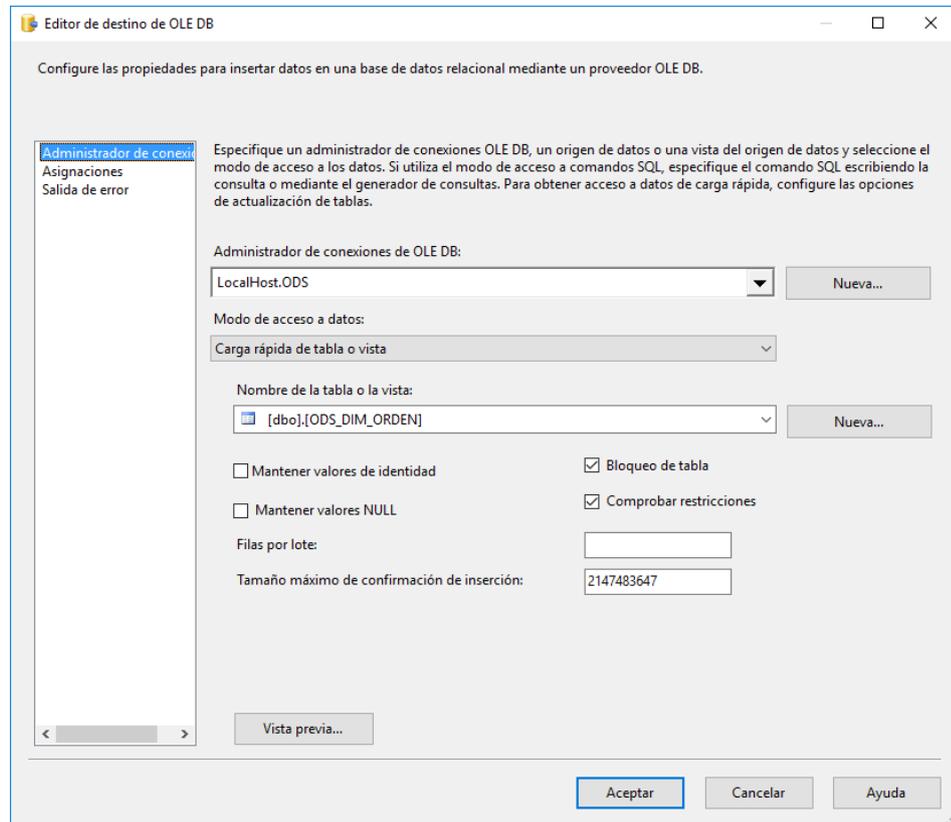


Figura 80 ODS_DIM_ORDEN - Destino ODS_DIM_ORDEN

ODS_DIM_INSUMO

Orígenes:

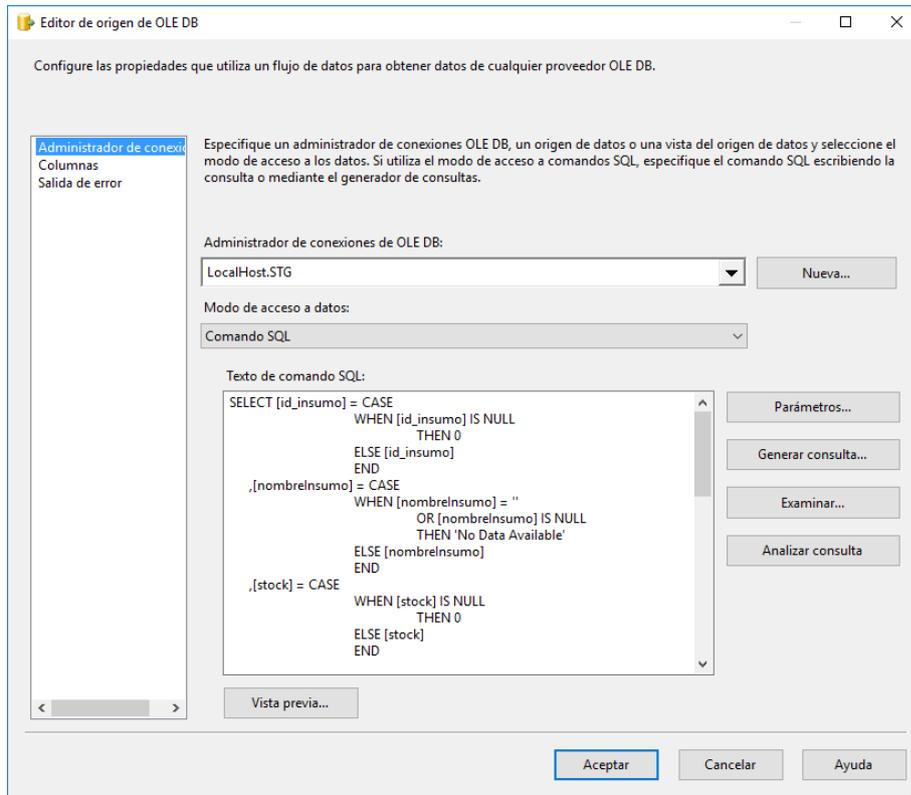


Figura 81 ODS_DIM_INSUMO - Origen STG_tb_insumos

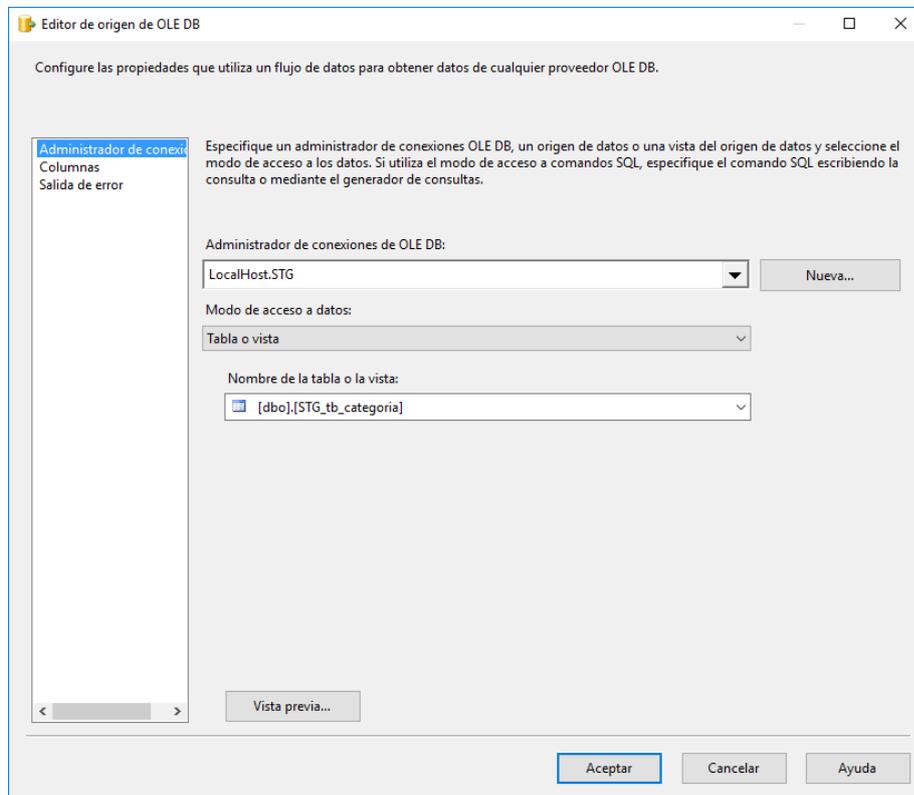


Figura 82 ODS_DIM_INSUMO - Origen STG_tb_categoria

Flujo de datos:

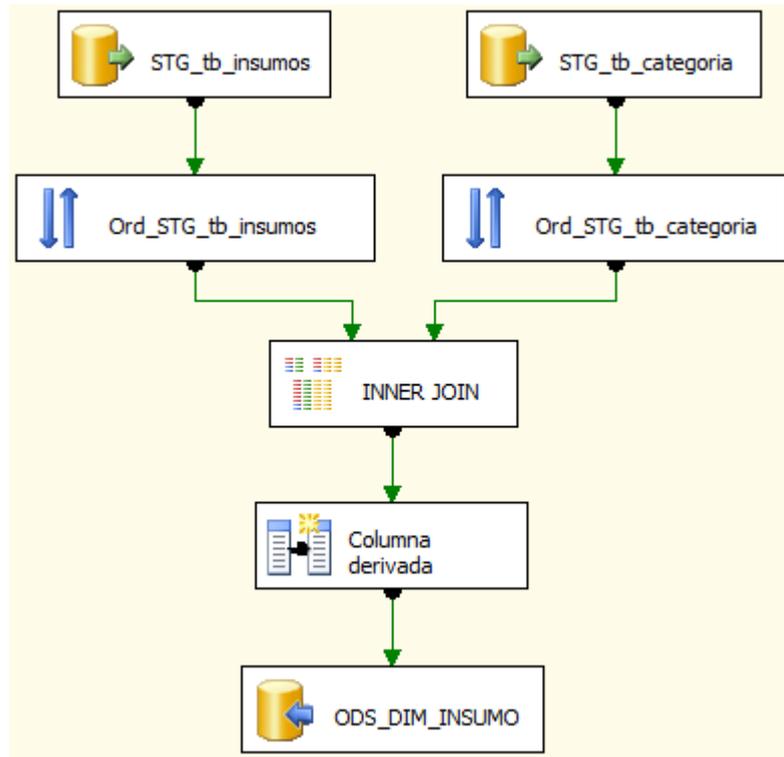


Figura 83 ODS_DIM_INSUMO - Flujo de Datos

Destino:

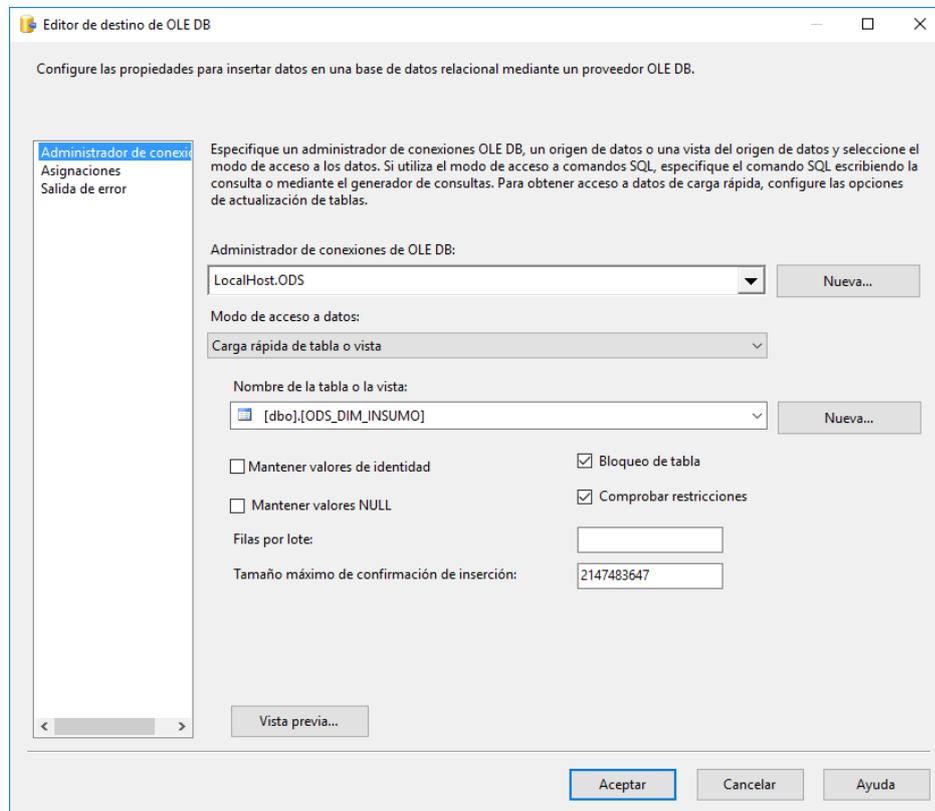


Figura 84 ODS_DIM_INSUMO - Destino ODS_DIM_INSUMO

ODS_FACT_COMERCIALIZACION

Orígenes:

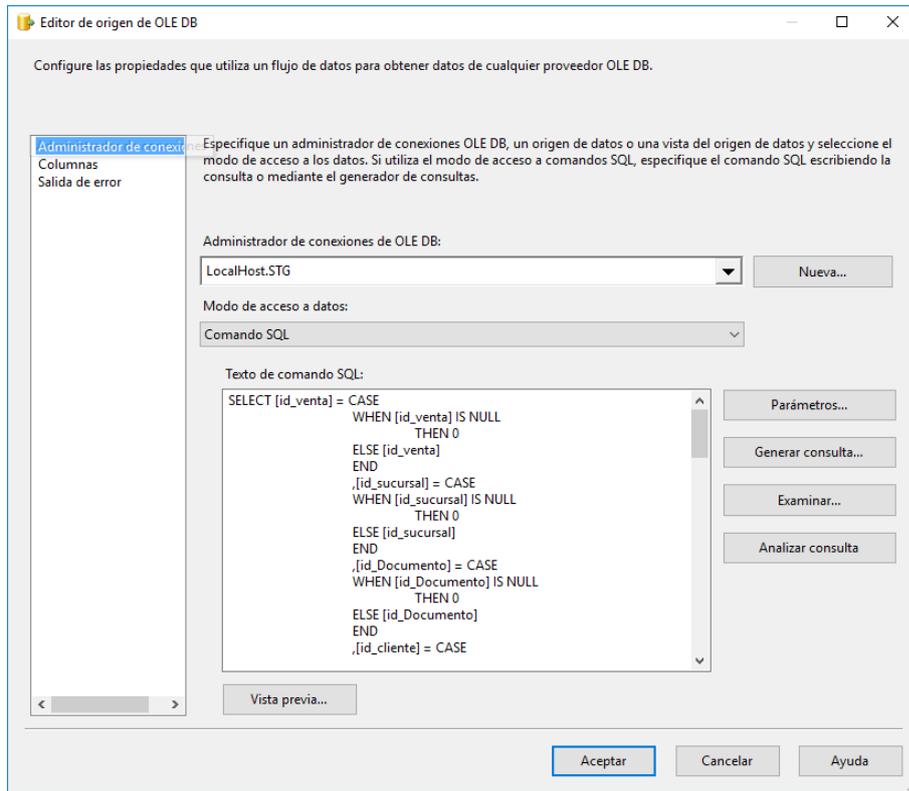


Figura 85 ODS_FACT_COMERCIALIZACION - Origen STG_tb_Ventas

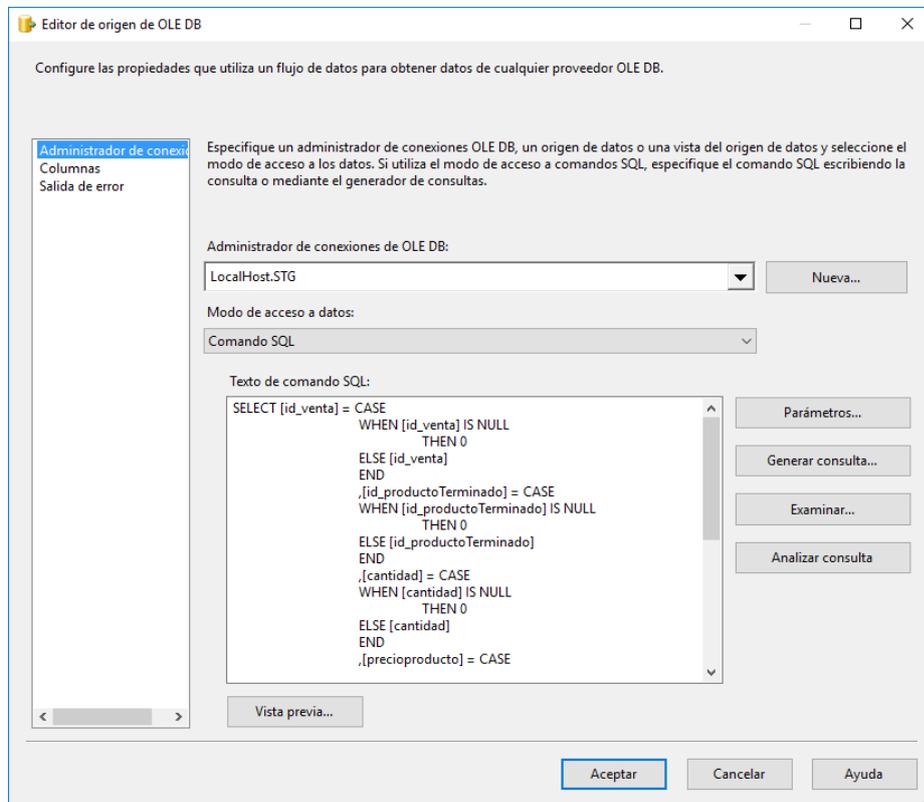


Figura 86 ODS_FACT_COMERCIALIZACION - Origen STG_tb_detalleVenta

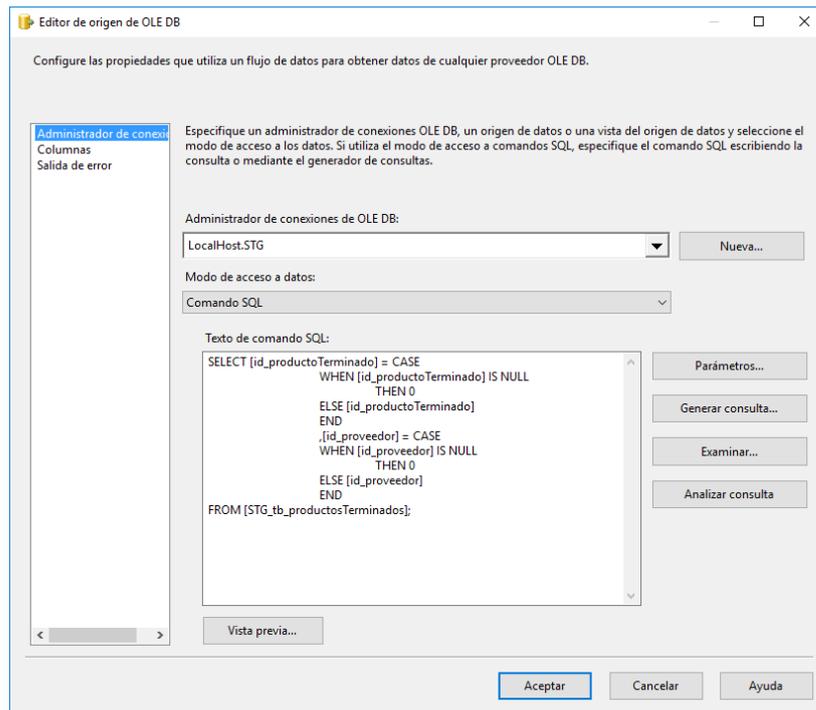


Figura 87 ODS_FACT_COMERCIALIZACION - Origen STG_tb_productosTerminados

Flujo de datos:

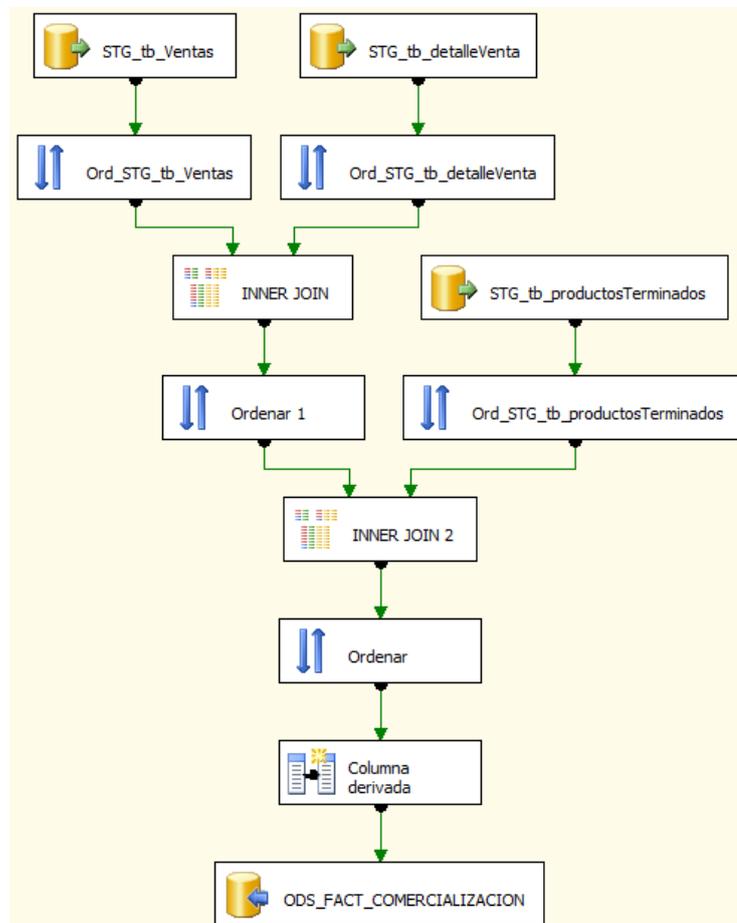


Figura 88 ODS_FACT_COMERCIALIZACION - Flujo de Datos

Destino:

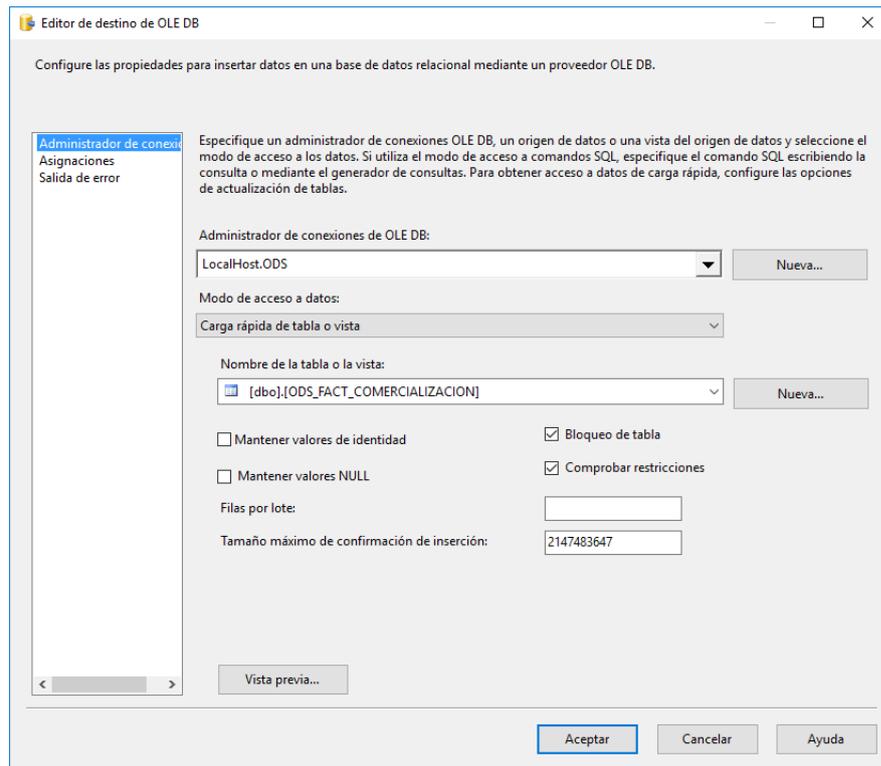


Figura 89 ODS_FACT_COMERCIALIZACION - Destino ODS_FACT_COMERCIALIZACION

ODS_FACT_PRODUCCION

Orígenes:

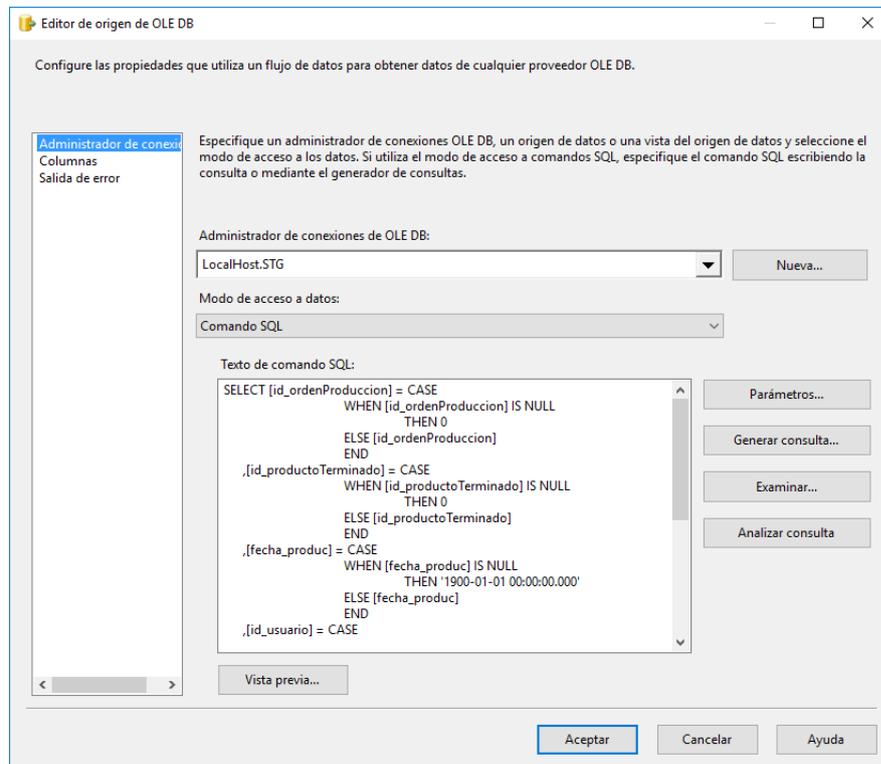


Figura 90 ODS_FACT_PRODUCCION - Origen STG_tb_ordenProduccion

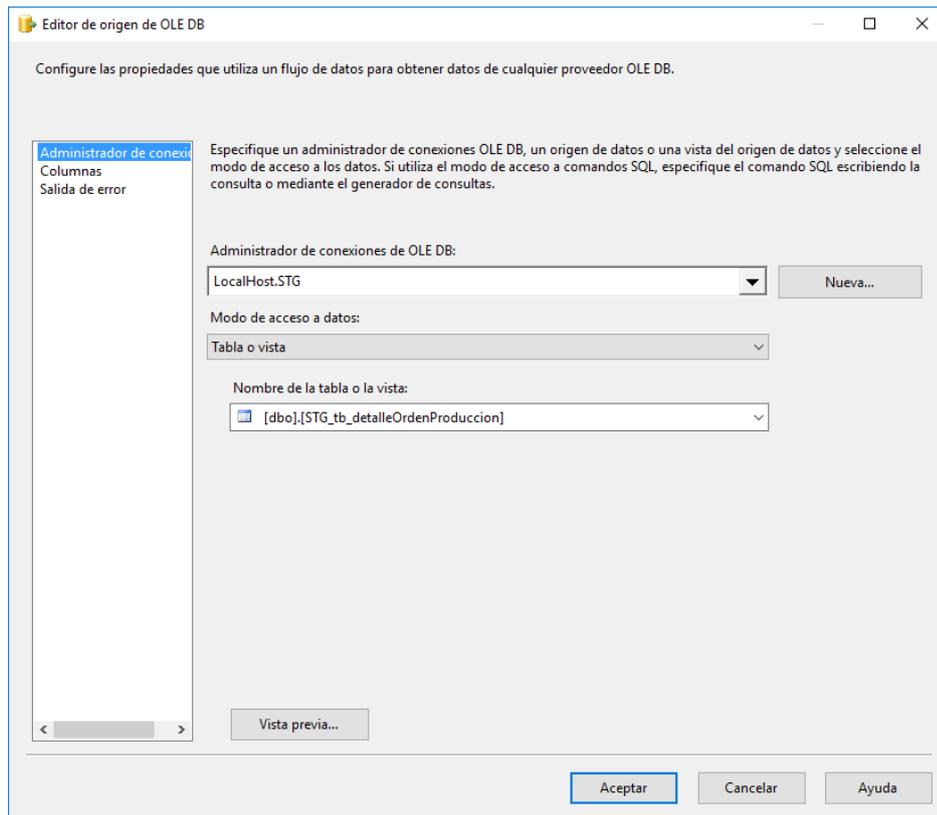


Figura 91 ODS_FACT_PRODUCCION - Origen STG_tb_detalleOrdenProduccion

Flujo de datos:

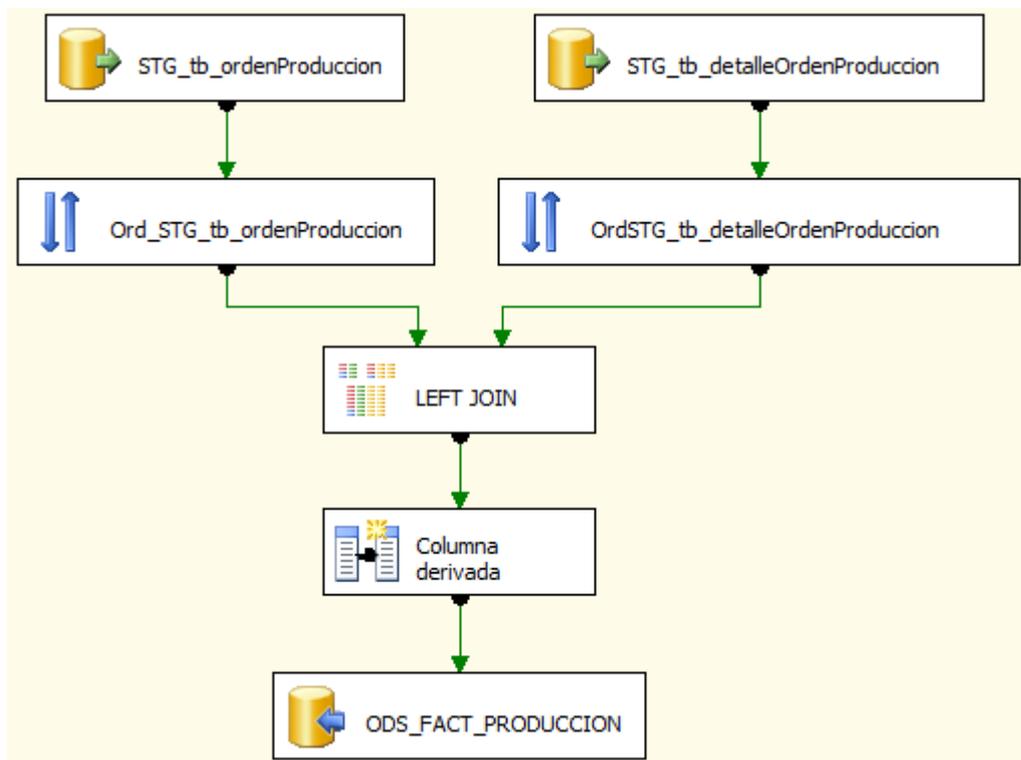


Figura 92 ODS_FACT_PRODUCCION - Flujo de Datos

Destino:

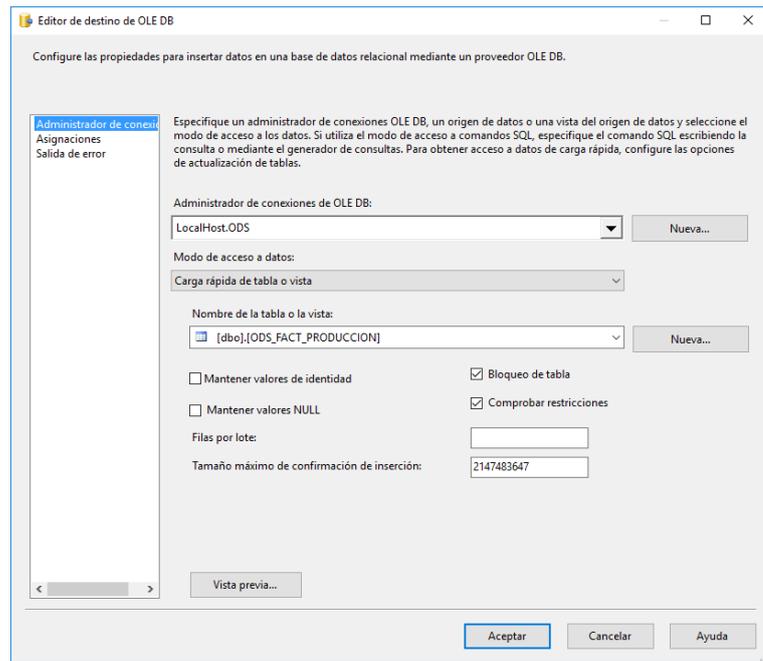


Figura 93 ODS_FACT_PRODUCCION - Destino ODS_FACT_PRODUCCION

3.1.5.3.Capa DW (Data Warehouse)

En esta última capa, se le añade los campos SK (Surrogate Key) a las dimensiones y se reemplaza los campos PK (Primary Key) por sus respectivos SK en las tablas de hechos, la información procesada en esta capa será cargada en los modelos dimensionales que han sido propuestos como solución.

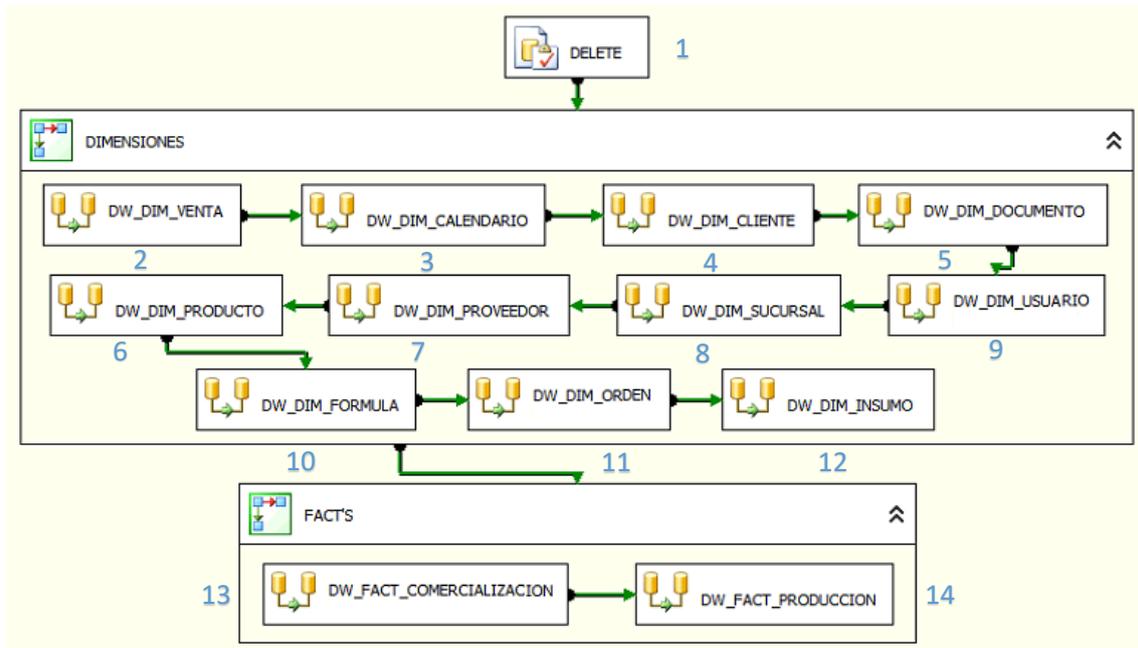


Figura 94 Capa DW (Data Warehouse)

El proceso de carga a las dimensiones es repetitivo, por lo cual sólo se mostrará la carga de una dimensión a manera de ejemplo, a continuación se muestra como se realiza la carga a la tabla DW_CLIENTE desde la tabla origen ODS_CLIENTE en el flujo de datos DW_DIM_CLIENTE de la imagen anterior:

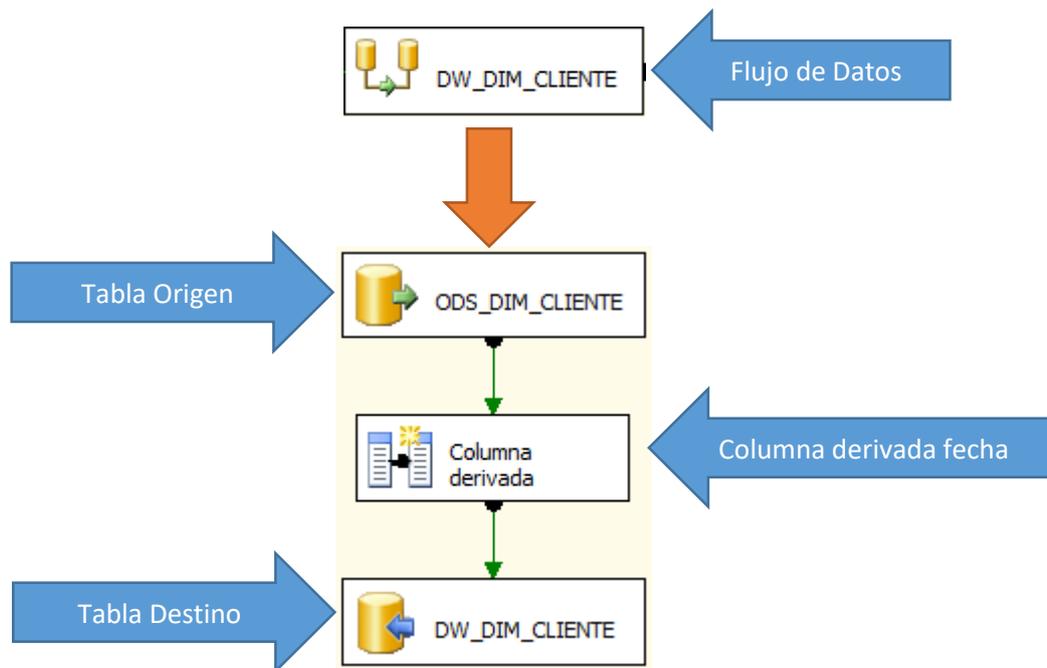


Figura 95 Flujo de Datos DW_DIM_CLIENTE

DW_FACT_COMERCIALIZACION

Orígenes:

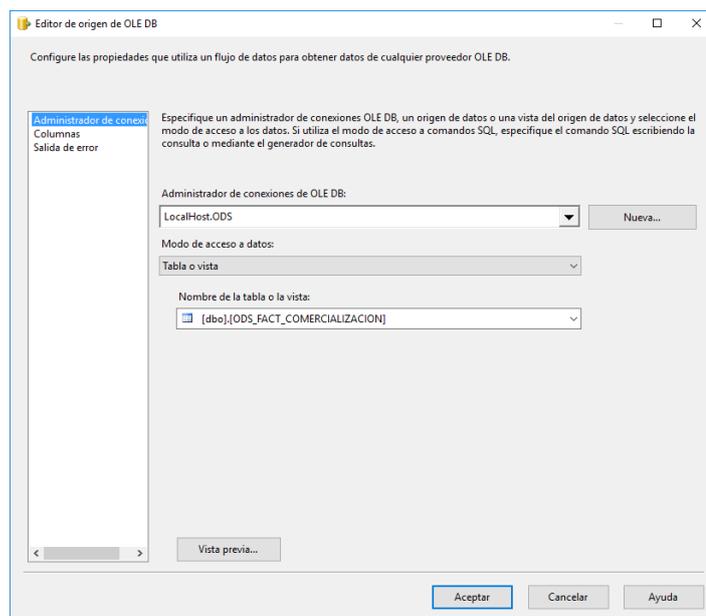


Figura 96 DW_FACT_COMERCIALIZACION - Origen ODS_FACT_COMERCIALIZACION

Flujo de datos:

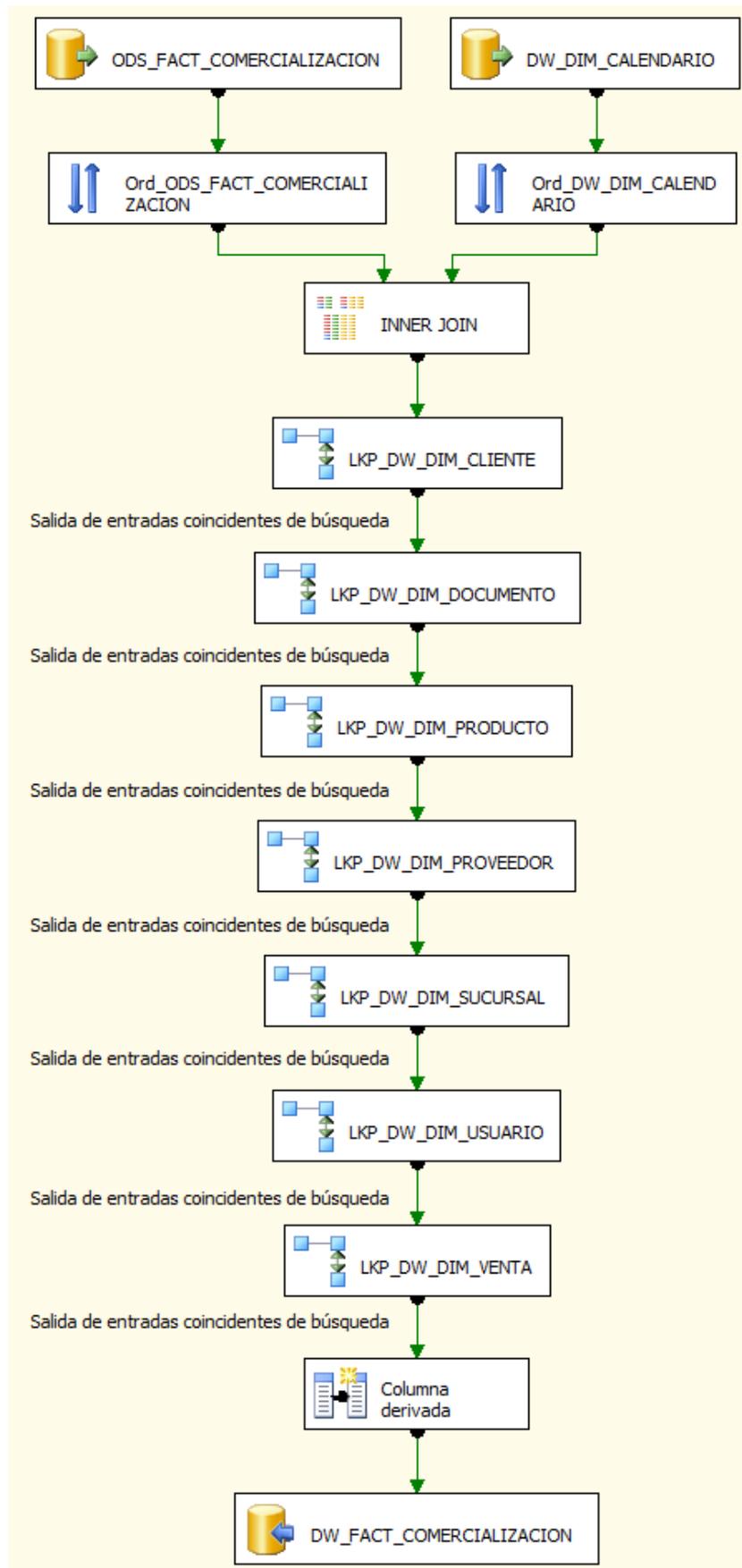


Figura 97 DW_FACT_COMERCIALIZACION - Flujo de Datos

Destino:

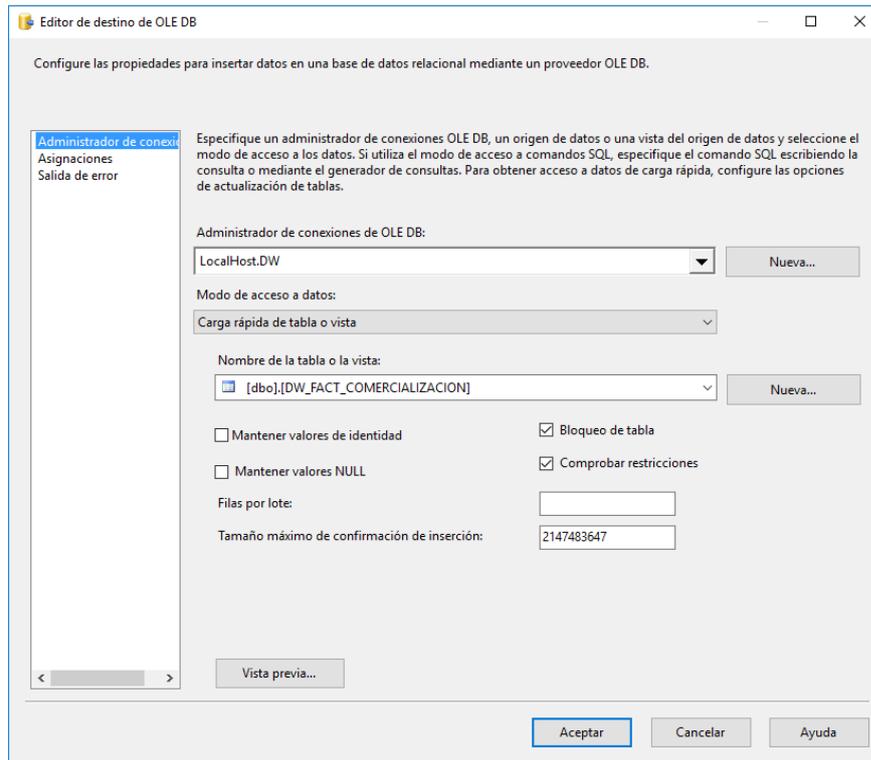


Figura 98 DW_FACT_COMERCIALIZACION - Destino DW_FACT_COMERCIALIZACION

DW_FACT_PRODUCCION

Orígenes:

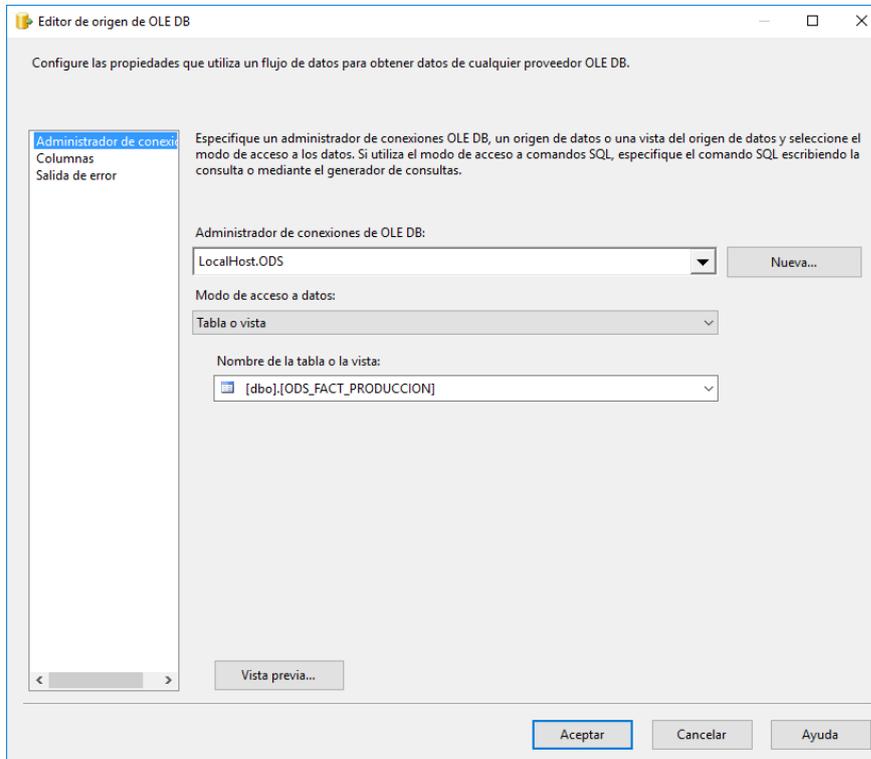


Figura 99 DW_FACT_COMERCIALIZACION - Origen ODS_FACT_COMERCIALIZACION

Flujo de datos:

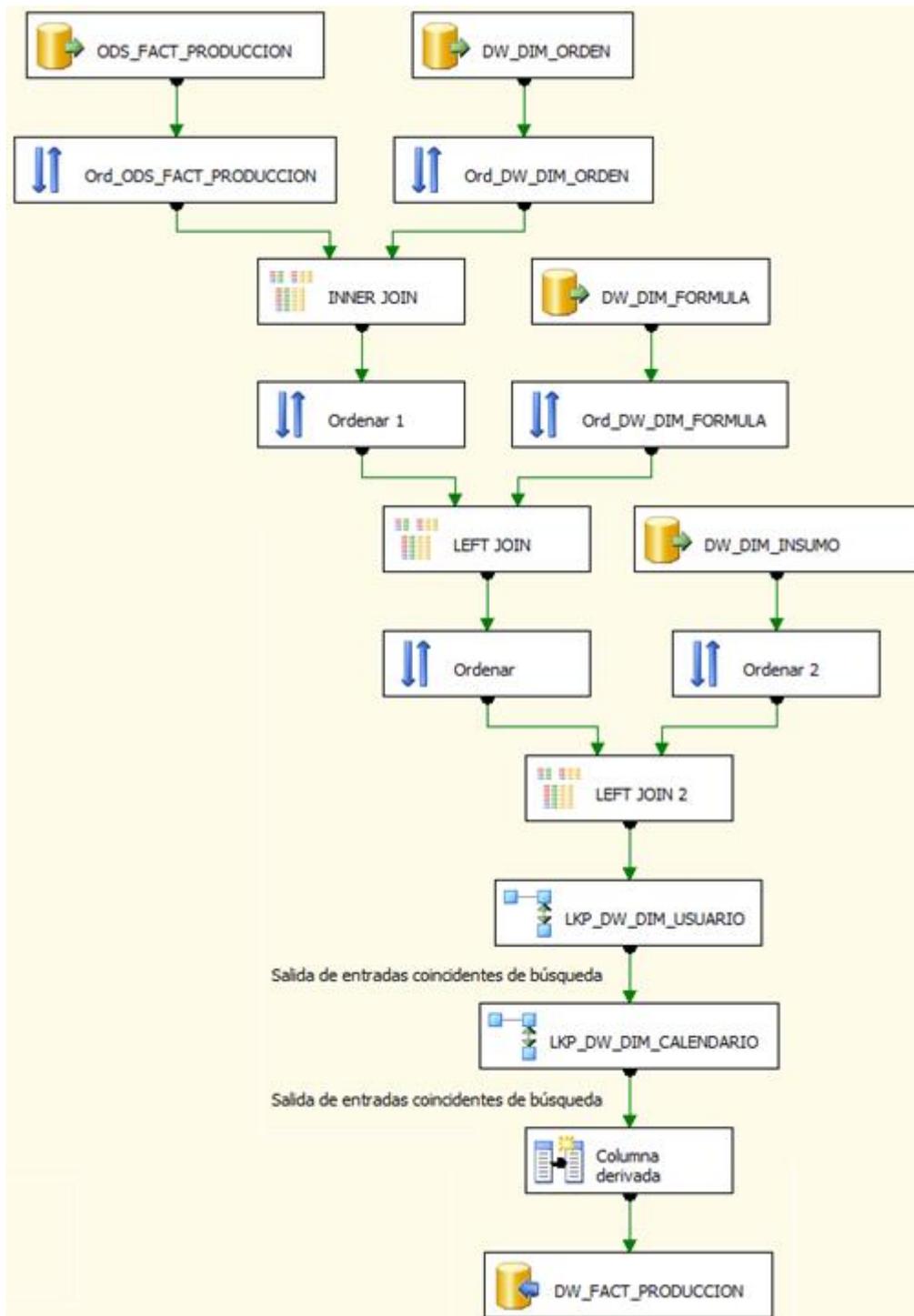


Figura 100 DW_FACT_COMERCIALIZACION - Flujo de Datos

Destino:

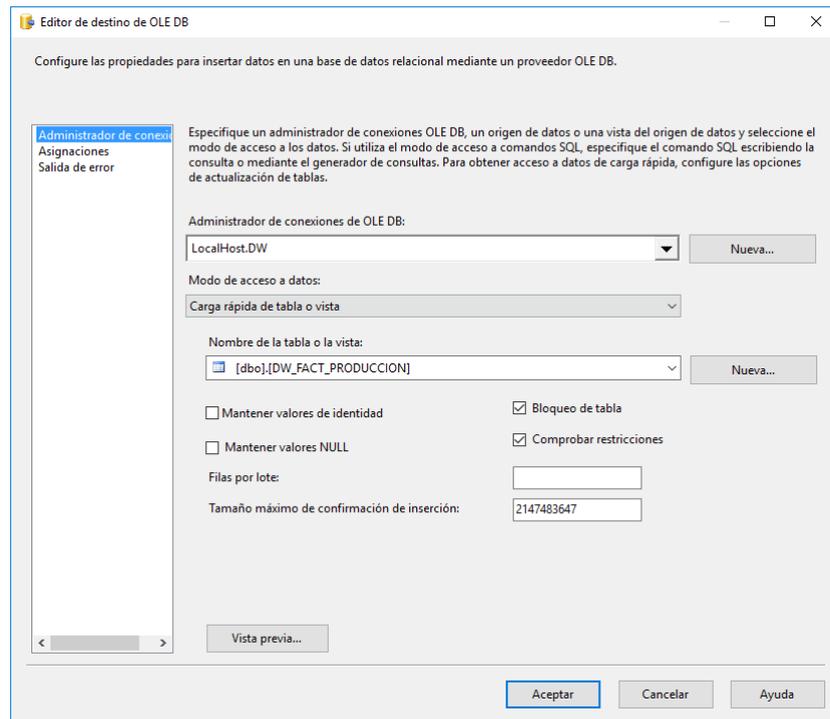


Figura 101 DW_FACT_COMERCIALIZACION - Destino DW_FACT_COMERCIALIZACION

3.1.5.4. Tipo de carga

No existe una regla general para establecer un tipo de carga en un data warehouse, esto suele ajustarse a las necesidades de cada proyecto, en esta ocasión se realizó una carga total de las 3 capas, lo cual tuvo una duración de 1:20 minutos, determinando que el tiempo y recursos requeridos para una carga total son mínimos por no ser un gran volumen de datos.

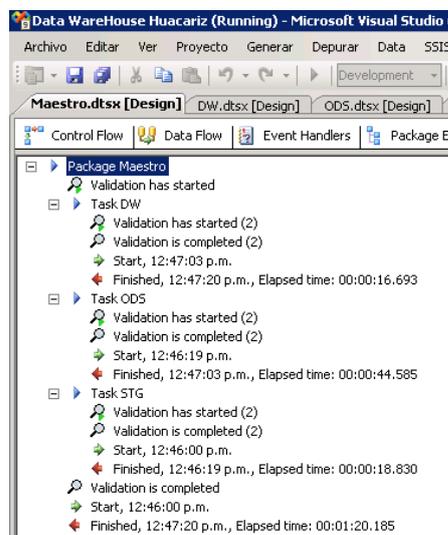


Figura 102 Ejecución Package Maestro

En vista de lo obtenido, la estrategia de carga a utilizar consistirá en programar cargas totales (borrar todo y volver a cargar desde cero) de las 3 capas todos los días a la 1 am, horario en el que el servidor tiene menos carga de trabajo con la finalidad de evitar cualquier inconveniente, para lo cual se siguieron los siguientes pasos:

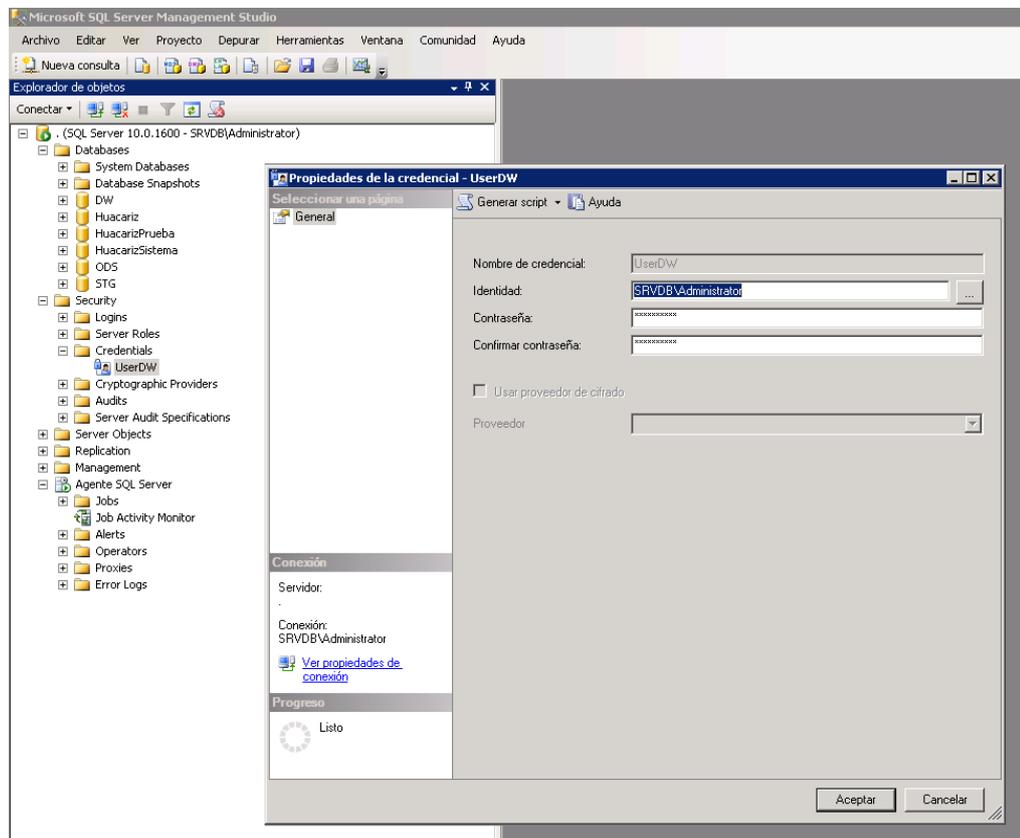


Figura 103 Crear una credencial

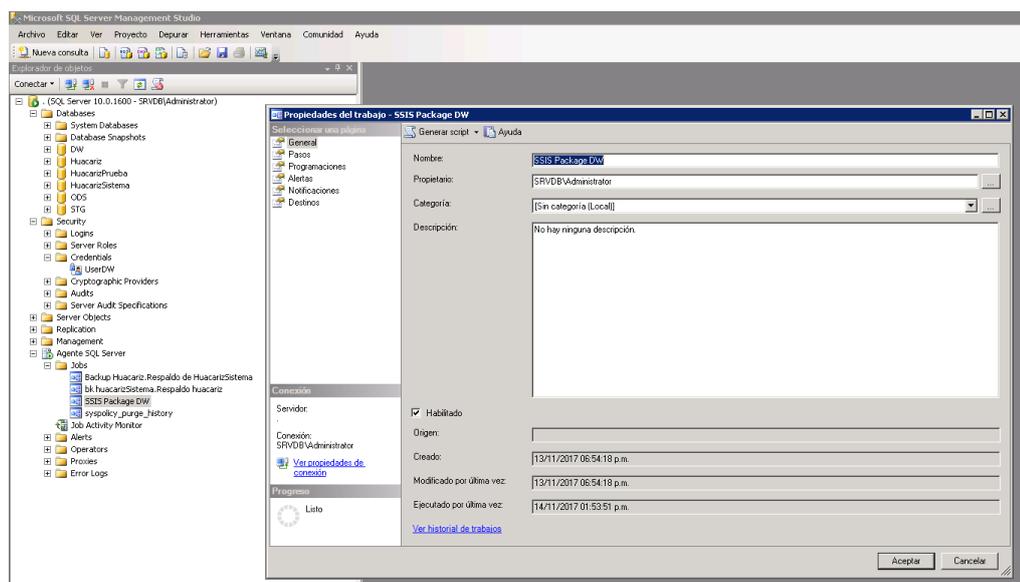


Figura 104 Crear un Job en el Agente SQL

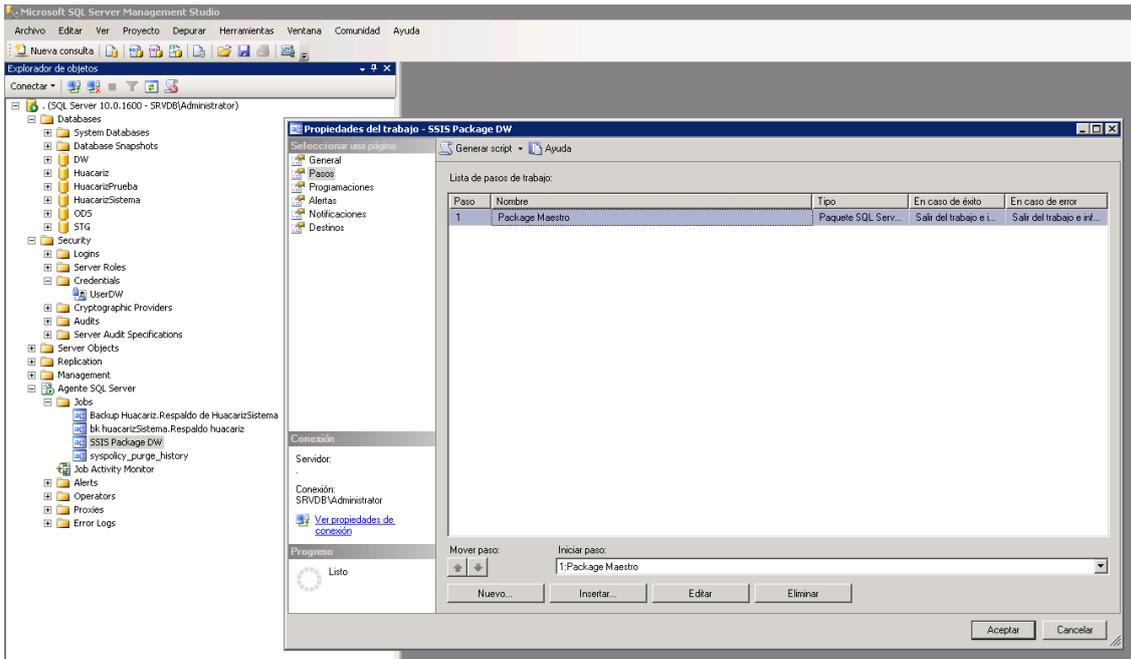


Figura 105 Crear los pasos del Job

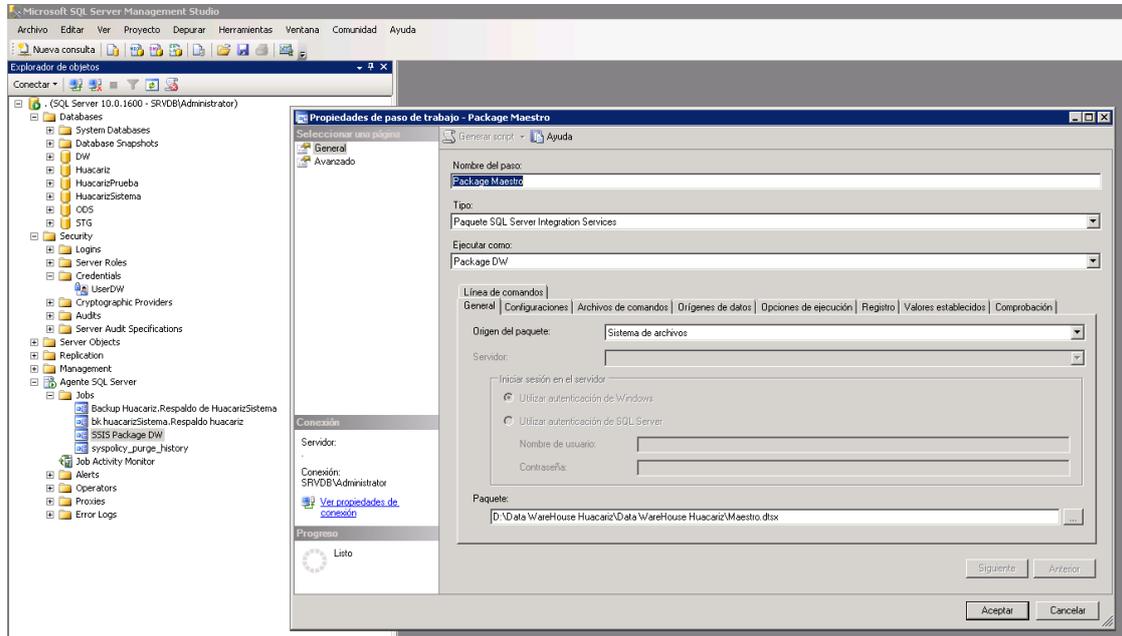


Figura 106 Configurar el paso para el Job

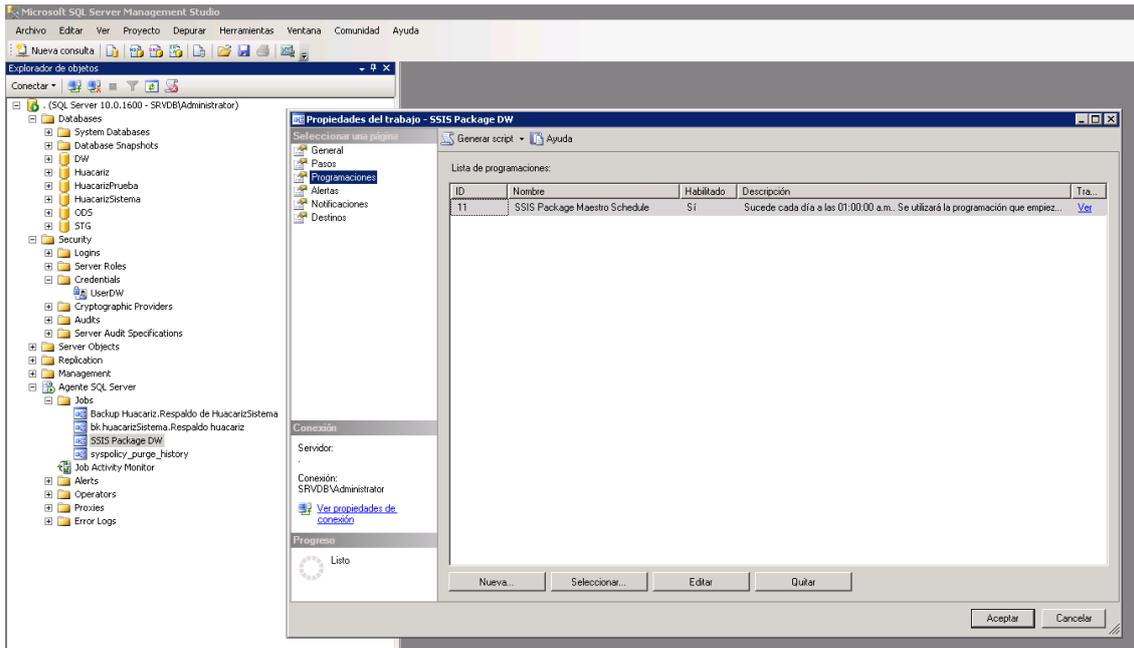


Figura 107 Crear la programación de ejecución para el Job

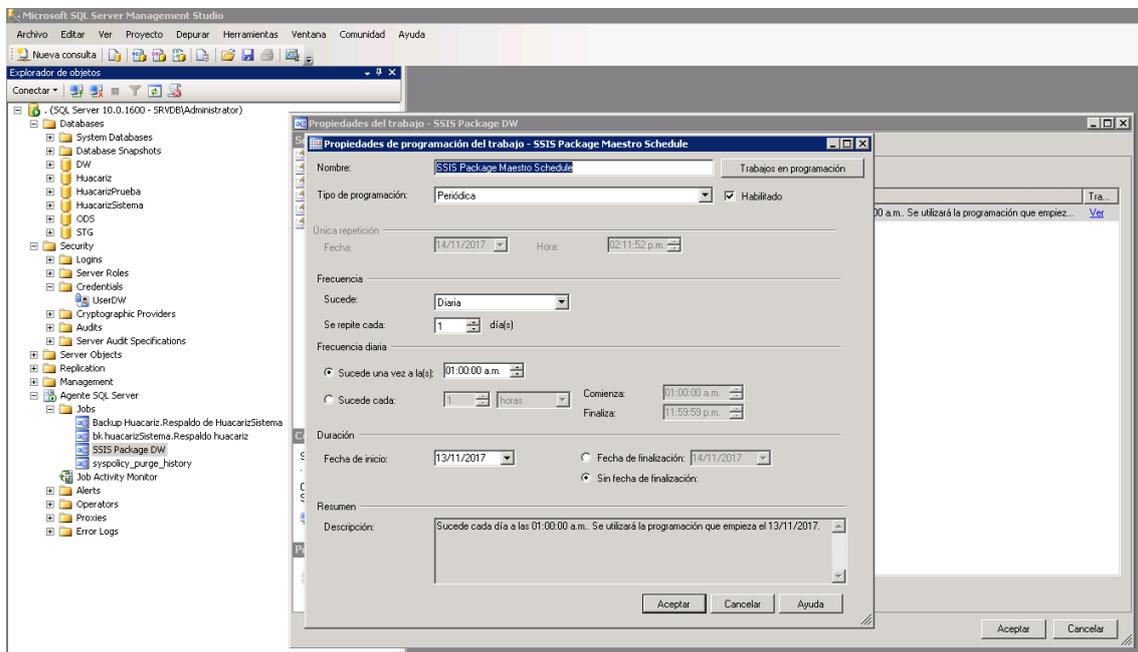


Figura 108 Configuración de la programación del Job

3.1.7. QlikView

Es una herramienta computacional desarrollada por QlikTech que permite el acceso a la información y posibilita el análisis de los datos, lo cual a su vez mejora y optimiza el proceso de toma de decisiones de negocio y por ende también el desempeño del mismo.

Históricamente, la inteligencia de negocios ha sido comandada principalmente por los departamentos de tecnologías de información en las empresas. Los departamentos de TI eran responsables de todo el ciclo de vida de una solución de inteligencia de negocios, desde extraer los datos hasta entregar los reportes finales, análisis y cuadros de mando. Aunque este modelo funciona bien para la distribución de reportes estáticos predefinidos, la mayoría de las empresas se han ido dando cuenta que no cumple con las necesidades de sus usuarios de negocio. Como TI controla de cerca los datos y herramientas, los usuarios comúnmente experimentan largos tiempos de espera cuando surgen nuevas preguntas de negocio que no pueden ser respondidas con los reportes estándar [36].

¿Cómo se diferencia QlikView de herramientas tradicionales de BI?

QlikTech se enorgullece de abordar la Inteligencia de Negocios de una manera distinta a lo que compañías como Oracle, SAP, e IBM (descritas por QlikTech como proveedores tradicionales de BI) ofrecen. QlikTech busca poner las herramientas en manos del usuario de negocio, permitiéndoles ser autosuficientes, ya que así puede realizar sus propios análisis.

Las firmas independientes de análisis de la industria han notado también este acercamiento distinto. En 2011, Gartner creó una subcategoría para herramientas de descubrimiento de datos en su evaluación anual de mercado, el cuadrante mágico de plataformas de inteligencia de negocios. QlikView fue el abanderado en esta nueva categoría de herramientas de BI.

QlikTech prefiere describir su producto como una herramienta de descubrimiento del negocio en lugar de descubrimiento de datos. Sostiene que descubrir cosas sobre el negocio es mucho más importante que descubrir datos. El siguiente diagrama ilustra este paradigma [36].



Figura 109 Descubrir datos VS Descubrir el negocio [36]

Además de la diferencia en quién usa la herramienta – usuarios de TI contra usuarios de negocio – hay algunas otras funcionalidades que diferencian a QlikView de otras soluciones [36].

Experiencia de usuario asociativa

La principal diferencia entre QlikView y otras soluciones de BI es la experiencia de usuario asociativa. Mientras que las soluciones de BI tradicionales usan caminos predefinidos para navegar y explorar datos, QlikView permite a los usuarios tomar cualquier ruta que deseen para realizar análisis. Esto resulta en una manera mucho más intuitiva de navegar los datos. QlikTech describe esto como “trabajar de la forma en que trabaja la mente humana”.

En la siguiente imagen se muestra un ejemplo. Mientras que en una solución típica de BI tendríamos que comenzar seleccionando una región para después entrar paso a paso en el camino jerárquico definido, en QlikView podemos elegir cualquier punto de entrada que deseemos – región, estado, producto, o vendedor. Al ir navegando los datos, se nos presenta solo la información relacionada a nuestra selección y, para nuestra siguiente selección, podemos elegir cualquier camino que deseemos. La navegación es infinitamente flexible [36].

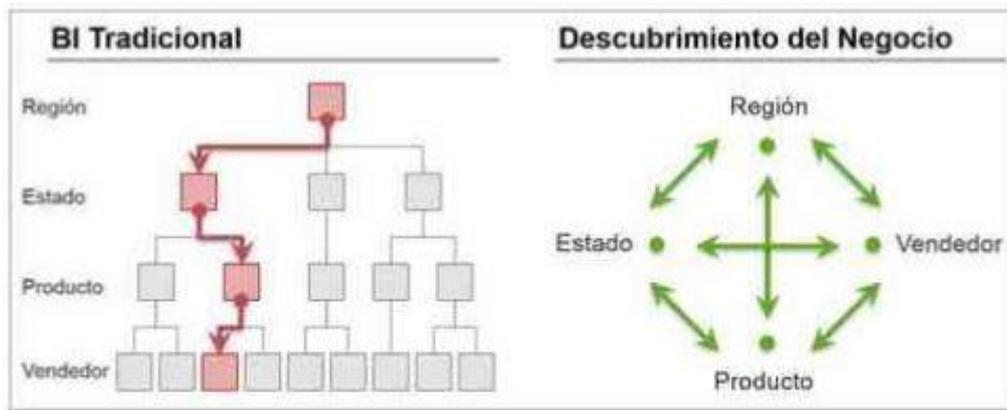


Figura 110 Descubrir datos VS descubrir el negocio (2) [36]

3.1.8. Desarrollo de la aplicación BI en QlikView

Una vez que el proceso de carga del ETL ha finalizado, le toca a la herramienta BI hacer su trabajo, en cual consiste en consumir el modelo dimensional, para ello se escogió a QlikView como herramienta.

Aplicación BI para el área de comercialización

Solo consume las dimensiones y hechos que correspondan a esta área.

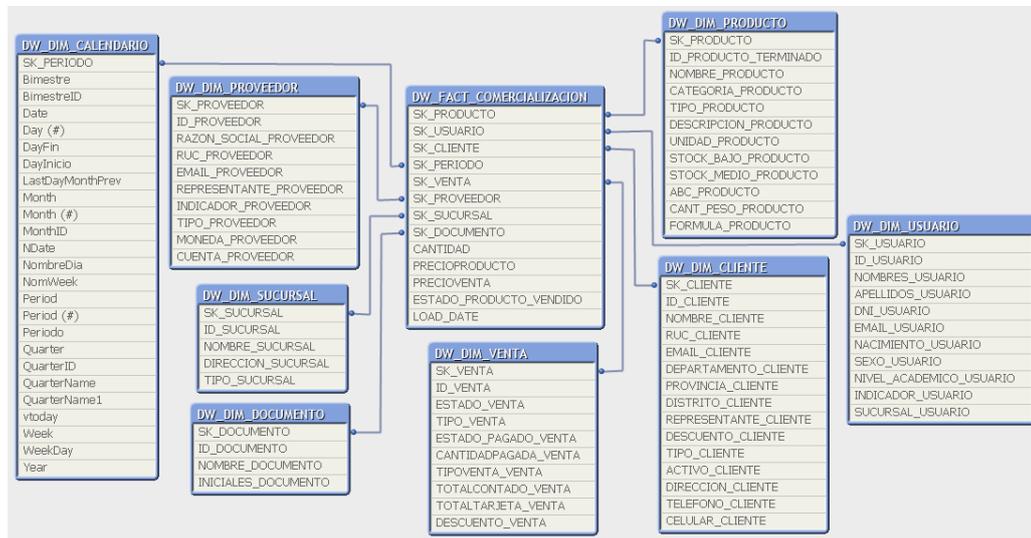


Figura 111 Modelo dimensional de comercialización en QlikView

Ventas y meta mensual: el reporte mostrado responde al requerimiento R-04 (Venta anual) y también al requerimiento R-05 (Meta para las ventas).

Interpretación - Se puede visualizar las ventas generales correspondientes al año 2017 y una meta mensual establecida a partir de la data histórica tomando en cuenta solo las ventas activas.



Figura 112 Ventas y meta mensual – Ventas generales

Interpretación - Se puede visualizar las ventas de la sucursal “San Martin” correspondientes al año 2017 y una meta mensual establecida a partir de la data histórica para la sucursal seleccionada tomando en cuenta solo las ventas activas.



Figura 113 Ventas y meta mensual – Por sucursal

Interpretación - Se puede visualizar a cuantos ascienden las ventas de “Quesillo al vacío Kg” correspondientes al año 2017 y una meta mensual establecida a partir de la data histórica tomando en cuenta solo las ventas activas.



Figura 114 Ventas y meta mensual – Por producto

Ventas por documento: el reporte mostrado responde al requerimiento R-06 (Ventas por documento).

Interpretación – Se puede visualizar las ventas generales correspondientes al 2017 distribuidas por tipo de documento de venta tomando en cuenta solo las ventas activas.



Figura 115 Ventas por documento – Ventas generales

Interpretación – Se puede visualizar las ventas correspondientes al “Queso Mantecoso Kg” del 2017 distribuidas por tipo de documento de venta tomando en cuenta solo las ventas activas.



Figura 116 Ventas por documento – Por producto

Interpretación – Se puede visualizar las ventas correspondientes al cliente “Milkal SAC” 2017 distribuidas por tipo de documento de venta tomando en cuenta solo las ventas activas.



Figura 117 Ventas por documento – Por cliente

Ventas por precio: el reporte mostrado responde al requerimiento R-07 (Ventas por precio).

Interpretación – Se puede visualizar las ventas generales del 2017 mostrando a cuánto ascienden las ventas con el precio de venta (Precio con descuento) y mostrando a cuanto hubiera ascendido las ventas con el precio del producto (Precio sin descuento) tomando en cuenta solo las ventas activas.



Figura 118 Ventas por precio – Ventas generales

Interpretación – Se puede visualizar las ventas de “Queso Paria Kg” del 2017 mostrando a cuánto ascienden las ventas con el precio de venta (Precio con descuento) y mostrando a cuanto hubiera ascendido las ventas con el precio del producto (Precio sin descuento) tomando en cuenta solo las ventas activas.



Figura 119 Ventas por precio – Por producto

Interpretación – Se puede visualizar las ventas por tipo de venta como “Paquetes” del 2017 mostrando a cuánto ascienden las ventas con el precio de venta (Precio con descuento) y mostrando a cuanto hubiera ascendido las ventas con el precio del producto (Precio sin descuento) tomando en cuenta solo las ventas activas.



Figura 120 Ventas por precio – Por tipo de venta

Producto con más demanda: el reporte mostrado responde al requerimiento R-08 (Productos con más demanda)

Interpretación – Se puede visualizar los productos con más demanda que representan el 50% de las ventas generales del 2017 tomando en cuenta solo las ventas activas.



Figura 121 Productos con más demanda - General

Interpretación – Se puede visualizar los productos con más demanda que representan el 50% de las ventas de la sucursal “ADP” del 2017 tomando en cuenta solo las ventas activas.



Figura 122 Productos con más demanda - Por Producto

Interpretación – Se puede visualizar los productos con más demanda que representan el 50% de las ventas con el cliente “Milkal SAC” del 2017 tomando en cuenta solo las ventas activas.



Figura 123 Productos con más demanda - Por cliente

Cuentas por cobrar: el reporte mostrado responde al requerimiento R-09 (Cuentas por cobrar)

Interpretación – Se puede visualizar las cuentas por cobrar, mostrando el total por año tomando en cuenta solo las ventas activas.



Figura 124 Cuentas por cobrar – Por años

Interpretación – Se puede visualizar las cuentas por cobrar, mostrando el total por mes según el año seleccionado tomando en cuenta solo las ventas activas.



Figura 125 Cuentas por cobrar – En meses por año

Interpretación – Se puede visualizar las cuentas por cobrar detallado por cliente, información que se filtraría según lo seleccionado en el gráfico y los filtros mostrados, y de igual manera, el grafico se podría filtrar según lo seleccionado en el detalle, toda esta información se muestra tomando en cuenta solo las ventas activas.



Figura 126 Cuentas por cobrar - Detalle por cliente

Aplicación BI para el área de Producción

Solo consume las dimensiones y hechos que correspondan a esta área.

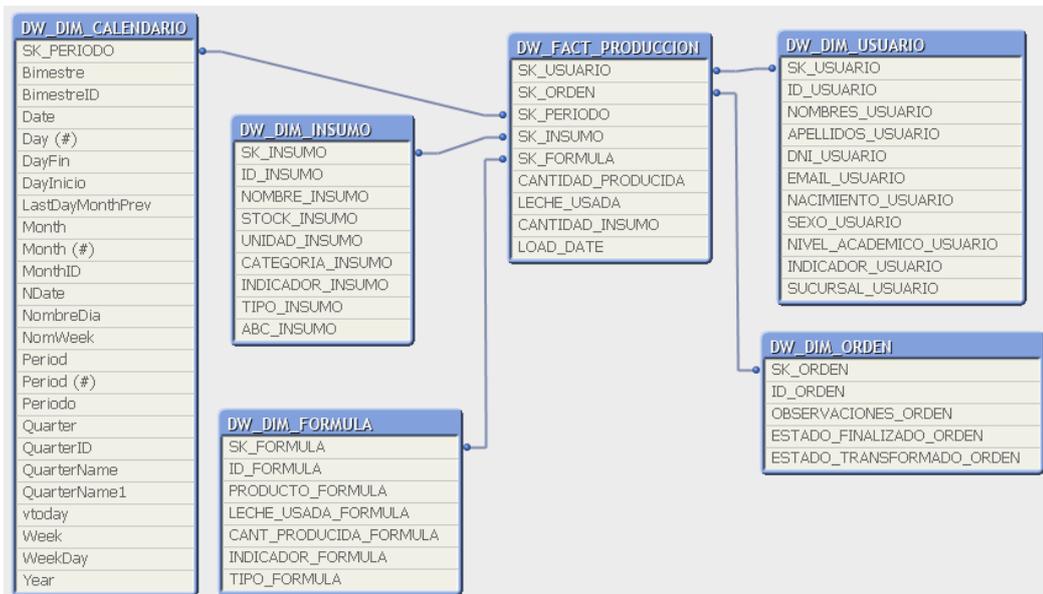


Figura 127 Modelo dimensional de Producción en QlikView

Consumo y proyección de leche: el reporte mostrado responde al requerimiento R-10 (Consumo de leche anual) y también al requerimiento R-11 (Proyección de consumo de leche).

Interpretación – Se puede visualizar el consumo de leche general y una proyección del consumo de la misma basada en la información histórica de las órdenes de producción.



Figura 128 Consumo y Proyección de leche

Consumo y proyección de insumo: el reporte mostrado responde al requerimiento R-12 (Uso de cada insumo) y también al requerimiento R-13 (Proyección de uso por insumo)

Interpretación – Se puede visualizar el consumo del insumo y una proyección del consumo de la misma basada en la información histórica.



Figura 129 Consumo y Proyección de insumo

Productos más producidos: el reporte mostrado responde al requerimiento R-14 (Productos más producidos)

Interpretación – Se puede visualizar los productos más producidos que representen el 50% de la producción según el mes seleccionado basado en la data histórica.



Figura 130 Productos más producidos

3.2.Tratamiento, análisis de datos y presentación de resultados

A continuación se presentan los resultados de la investigación en base a la información recogida mediante las técnicas e instrumentos de recolección de datos, las que se muestran mediante cuadros y gráficos estadísticos de acuerdo a la hipótesis de trabajo, para lo que se aplicó una medición de tiempos a los reportes presentados y una encuesta para medir el grado de aceptación, obteniendo los siguientes resultados:

3.2.1.Tiempo de trabajo por cada reporte generado

Tabla 67 Medición de tiempo - Pre test y Post test

Medición de tiempos				
Reporte		Código	Minutos al mes	
			Pre - Test	Post - Test
1	Ventas y meta mensual*	RE1	960.00	0.19
2	Ventas por documento	RE2	35.00	0.17
3	Ventas por precio	RE3	1440.00	0.18
4	Producto con más demanda**	RE4	360.00	0.18
5	Cuentas por cobrar	RE5	120.00	0.17
6	Consumo y proyección de leche*	RE6	140.00	0.18
7	Consumo y proyección de insumo**	RE7	720.00	0.17
8	Productos más producidos**	RE8	360.00	0.16

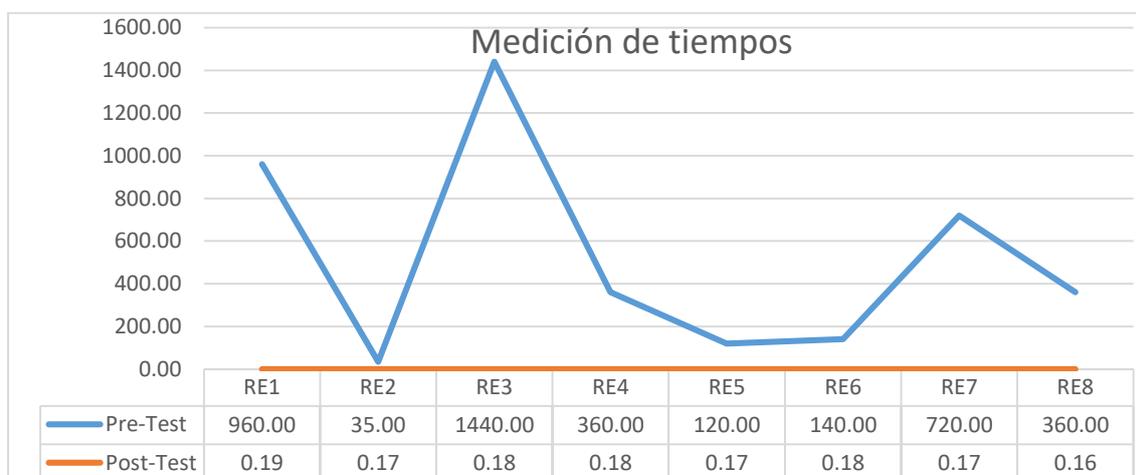


Figura 131 Gráfico de medición de tiempos

Nota:

* El reporte solo existe de manera parcial y se le asignó un tiempo estimado en la medición.

** El reporte no existe y la medición de tiempo se basó en un estimado.

Se puede observar que se ha logrado agilizar los 8 reportes presentados, donde podemos resaltar que se han reducido sus tiempos significativamente, esto beneficiará enormemente tanto al área de producción como al área de comercialización en el análisis de la información para la toma de decisiones.

3.2.2. Grado de satisfacción en la presentación de reportes

En el Anexo 1 podemos ver el modelo de la encuesta realizada

Tabla 68 Grado de Satisfacción - Pre Test

Nº	Pregunta	Pre Test	Post Test
1	El acceso a la información que necesita para la toma de decisiones es directo.	1.00	5.00
2	El tiempo de espera para recibir la información solicitada no excede los 60 minutos.	1.67	5.00
3	El formato en el que se recibe la información permite el análisis inmediato de la misma.	1.33	4.67
4	La interfaz es de fácil entendimiento	1.00	4.33
5	El procesamiento y análisis de reportes no demanda tiempo extra.	2.00	5.00
6	De forma general, los reportes recibidos le brindan información útil para la toma de decisiones.	2.33	4.67
7	El nivel de detalle de la información mostrada es el adecuado para el análisis.	2.00	5.00
8	El sistema actual le brinda información histórica para contrastar en el tiempo.	1.00	5.00
9	El proceso que sigue para obtener la información deseada es práctico y sencillo	1.33	5.00
10	Mejor control sobre las ventas y la producción	2.00	4.67
	Total	15.67	48.33

Se puede observar que se ha logrado aumentar el grado de satisfacción con respecto a los reportes obtenidos que necesitan para la toma de decisiones, esto se ve reflejado tanto en el área de producción como en el área de comercialización.

CAPÍTULO IV ANÁLISIS Y DISCUSIÓN DE RESULTADOS

4.1. Análisis de resultados

4.1.1. Indicador razón: Tiempo de trabajo por cada reporte generado

a) Contrastación de la hipótesis

Para la evaluación de la hipótesis planteada se va a utilizar la prueba T de Student para datos pareados o muestras relacionadas, es decir, esta prueba estadística exige dependencia entre ambas muestras, en las que hay dos momentos, un pre - test y un post - test. Con ello se da a entender que en el primer periodo, las observaciones servirán de testigo, para conocer los cambios que se suscitan después de aplicar el data warehouse. Con la prueba t se comparan las medias y las desviaciones estándar de grupo de datos y se determina si entre esos parámetros las diferencias son estadísticamente significativas o si sólo son diferencias aleatorias.

Se utilizará la siguiente fórmula para muestras relacionadas:

$$t = \frac{\bar{\partial}}{\sigma\partial/\sqrt{N}}$$

Donde:

t : Valor estadístico del procedimiento.

$\bar{\partial}$: Valor promedio o media aritmética de las diferencias entre los momentos después y antes.

$\sigma\partial$: Desviación estándar de las diferencias entre los momentos después y antes.

N : Tamaño de la muestra.

La media aritmética de las diferencias se obtiene de la siguiente manera:

$$\bar{\partial} = \frac{\sum \partial}{N}$$

La desviación estándar de las diferencias se obtiene como sigue:

$$\sigma\partial = \sqrt{\frac{\sum(\partial - \bar{\partial})^2}{N - 1}}$$

Se utilizará una tabla resumen en donde se mostrará los datos obtenidos de la medición de tiempos en el pre – test y el post – test para la obtención de reportes.

Tabla 69 Resumen de medición de tiempo - Pre test y Post test

N°	Reporte	Tiempo min.		$d=(X_D-$	$(d-\bar{d})$	$(d-\bar{d})^2$
		Pre-test (X_A)	Post-test (X_D)			
1	RE1	960.00	0.19	-959.81	-443.11	196346.47
2	RE2	35.00	0.17	-34.83	481.87	232198.70
3	RE3	1440.00	0.18	-1439.82	-923.12	852150.53
4	RE4	360.00	0.18	-359.82	156.88	24611.33
5	RE5	120.00	0.17	-119.83	396.87	157505.80
6	RE6	140.00	0.18	-139.82	376.88	142038.53
7	RE7	720.00	0.17	-719.83	-203.13	41261.80
8	RE8	360.00	0.16	-359.84	156.86	24605.06
N=	8	4135.00	1.40	-4133.60		1670718.23

Donde:

X_A : Representa a la evaluación de los tiempos en el pre-test.

X_D : Representa a la evaluación de los tiempos en el post-test.

$d= (X_D-X_A)$: Diferencias entre los momentos después y antes, es decir el post-test menos el pre-test.

b) Formulación de hipótesis estadística

- ✓ **Hipótesis Nula (H_0):** La implementación de un data warehouse no tendrá ningún impacto positivo en la toma de decisiones en el área de comercialización y producción de la Industria Alimentaria Huacariz.

$$H_0: \mu_D \geq \mu_A ; \mu_D - \mu_A \geq 0$$

- ✓ **Hipótesis Alternativa (H_A):** La implementación de un data warehouse tendrá un impacto positivo en la toma de decisiones en el área de comercialización y producción de la Industria Alimentaria Huacariz.

$$H_A: \mu_D < \mu_A ; \mu_D - \mu_A < 0$$

c) Nivel de significancia

Para esta investigación utilizaremos 5% de nivel de significancia, debido a que se trata de un proyecto de investigación, por lo cual se utilizará ese nivel de significancia.

$$\alpha = 0.05$$

d) Valor estadístico del procedimiento

$$\bar{d} = \frac{-4133.60}{8} = -516.70$$

$$\sigma_{\bar{d}} = \sqrt{\frac{1670718.23}{8 - 1}} = 488.54$$

$$t = \frac{-516.70}{488.54/\sqrt{8}} = -2.99$$

De acuerdo a la tabla T-Student el valor de la probabilidad p del valor estadístico del procedimiento es el siguiente:

$$p(t < -2.99) = 0.010093$$

e) Establecer región crítica

gl: Grados de libertad (N-1) $\Rightarrow gl = 7$

tt: Distribución T de Student $\Rightarrow tt_{0.05;7} = -1.895$

De acuerdo a lo obtenido:

$$t < tt = -2.99 < -1.895$$

$$p < \alpha = 0.010093 < 0.05$$

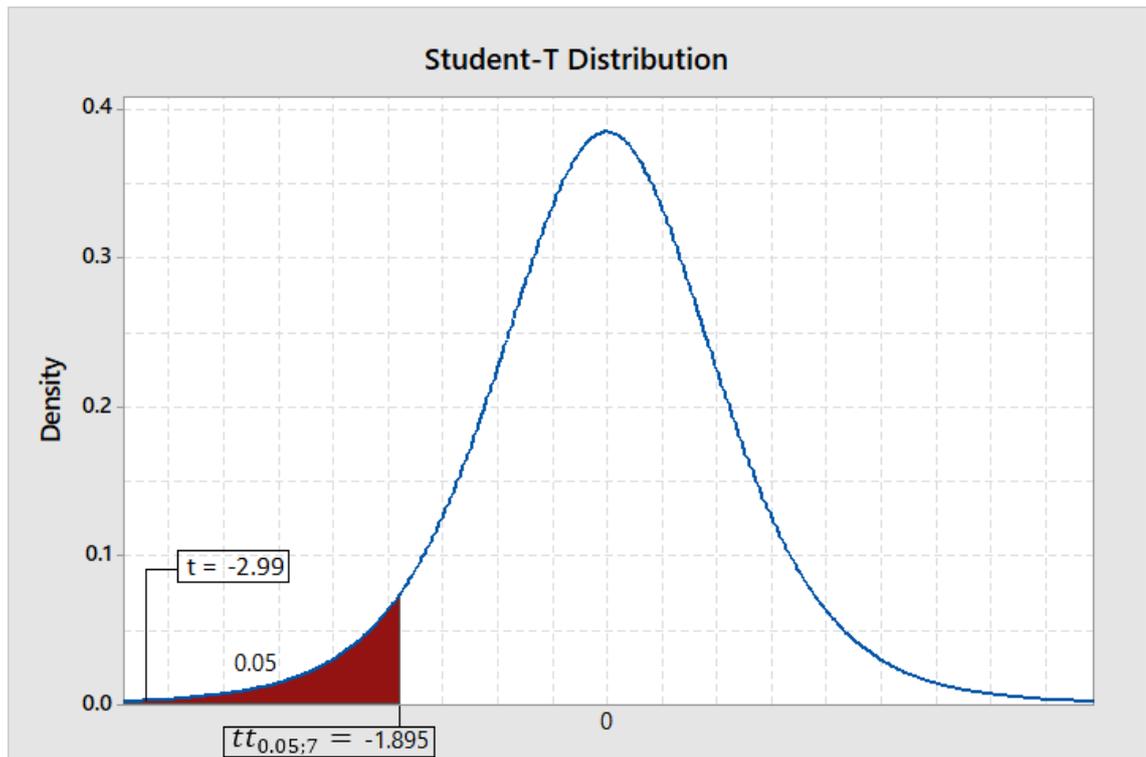


Figura 132 Distribución T-Student

Como son verdaderas las dos proposiciones, entonces rechazamos las hipótesis nula (H_0) y aceptamos la hipótesis alternativa (H_A).

f) Toma de decisión

Aceptamos la hipótesis alternativa y diremos que la implementación de un data warehouse tendrá un impacto positivo en la toma de decisiones en el área de comercialización y producción de la Industria Alimentaria Huacariz, con un nivel de confianza del 95% y un nivel de significancia del 5%.

4.1.2. Indicador ordinal: Grado de satisfacción de los usuarios en cuanto al Data Warehouse

Los resultados mostrados a continuación son los practicados a las personas que trabajan en las áreas de comercialización y producción de la Industria Alimentaria Huacariz, lo cual fue a un total de 3 personas.

Tabla 70 Grado de satisfacción - Pre Test

Pre test		Para nada satisfecho	Poco satisfecho	Indeciso	Satisfecho	Muy satisfecho	Total	Puntaje Promedio
Nº	Pregunta	1	2	3	4	5		
1	El acceso a la información que necesita para la toma de decisiones es directo.	3					3	1.00
2	El tiempo de espera para recibir la información solicitada no excede los 60 minutos.	1	2				3	1.67
3	El formato en el que se recibe la información permite el análisis inmediato de la misma.	2	1				3	1.33
4	La interfaz es de fácil entendimiento	3					3	1.00
5	El procesamiento y análisis de reportes no demanda tiempo extra.		3				3	2.00
6	De forma general, los reportes recibidos le brindan información útil para la toma de decisiones.		2	1			3	2.33
7	El nivel de detalle de la información mostrada es el adecuado para el análisis.		3				3	2.00
8	El sistema actual le brinda información histórica para contrastar en el tiempo.	3					3	1.00
9	El proceso que sigue para obtener la información deseada es práctico y sencillo	2	1				3	1.33
10	Mejor control sobre las ventas y la producción		3				3	2.00

Tabla 71 Grado de satisfacción - Post Test

Post test		Para nada satisfecho	Poco satisfecho	Indeciso	Satisfecho	Muy satisfecho	Total	Puntaje Promedio
Nº	Pregunta	1	2	3	4	5		
1	El acceso a la información que necesita para la toma de decisiones es directo.					3	3	5.00
2	El tiempo de espera para recibir la información solicitada no excede los 60 minutos.					3	3	5.00
3	El formato en el que se recibe la información permite el análisis inmediato de la misma.				1	2	3	4.67
4	La interfaz es de fácil entendimiento				2	1	3	4.33
5	El procesamiento y análisis de reportes no demanda tiempo extra.					3	3	5.00
6	De forma general, los reportes recibidos le brindan información útil para la toma de decisiones.				1	2	3	4.67
7	El nivel de detalle de la información mostrada es el adecuado para el análisis.					3	3	5.00
8	El sistema actual le brinda información histórica para contrastar en el tiempo.					3	3	5.00
9	El proceso que sigue para obtener la información deseada es práctico y sencillo					3	3	5.00
10	Mejor control sobre las ventas y la producción				1	2	3	4.67

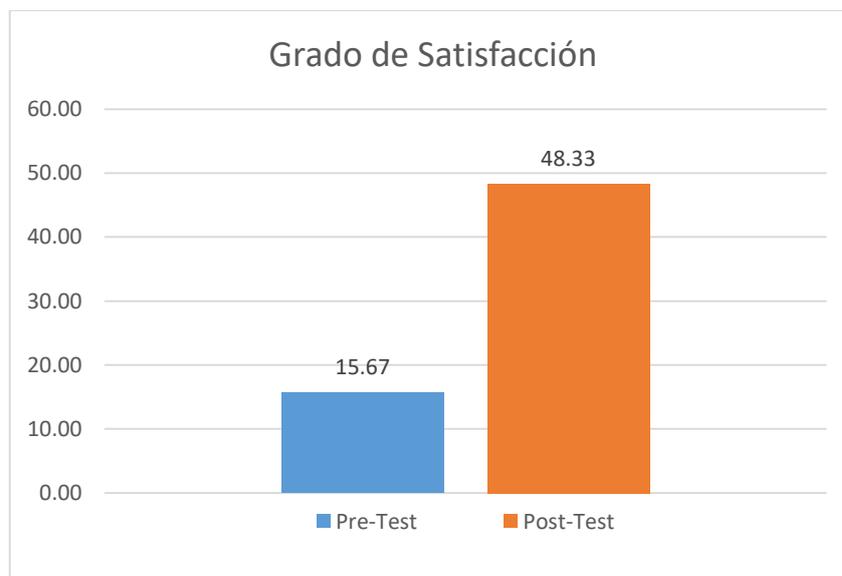


Figura 133 Gráfico comparativo del grado de satisfacción

Como podemos ver en las tablas presentadas del pre-test, del post-test y en el gráfico comparativo, los usuarios indican un gran crecimiento en el grado de satisfacción, mostrándose complacidos con los tiempos de presentación de reportes, calidad de datos, formato presentado y accesibilidad a los datos.

4.2.Discusión de resultados

Esta investigación tuvo como propósito identificar y analizar los requerimientos de información que intervendrían en los modelos dimensionales del data warehouse, lo cual se obtuvo tras aprender todo lo que se pudiera de los archivos que generaban los usuarios como resultado del procesamiento de datos para su respectivo análisis, permitiendo encontrar puntos clave como el nivel de granularidad y la información con la que necesita contar el data warehouse, luego de ello se entrevistó a los usuarios para aclarar cuestiones que no se comprendían y solicitar opiniones sobre los problemas y las limitaciones que presenta el sistema en cuanto a la información que brinda y el formato en el que son presentados.

Gracias a los requerimientos encontrados se pudo diseñar e implementar el data warehouse para las áreas de comercialización y producción siguiendo las pautas encontradas en la literatura de la metodología de Ralph Kimball donde indica las consideraciones necesarias que se debe tener durante el diseño, adicional a ello se tomó como referencia el trabajo de graduación de Claudia Estefanía Jaramillo Mejía desarrollado en el año 2015 titulado “Implementación de un sistema de inteligencia de negocios en el departamento de ventas de una empresa nacional del sector lácteo y desarrollo de reportes dinámicos utilizando herramientas de software libre” [8], pues al estar enfocado en el departamento de ventas y haber demostrado que su modelo logró satisfacer sus objetivos, sirvió como base para diseñar el modelo de las áreas de comercialización y producción.

Los reportes que permiten consumir el data warehouse fueron creados en una herramienta de business intelligence llamada QlikView, estos reportes no solo permiten mostrar la data en forma de tablas, sino también en forma

de gráficos, permitiendo así una mayor comprensión de la información por parte del usuario final, además de presentar información puntual e histórica, permite mostrar proyecciones de ventas, consumo de leche e insumos, hecho que en ninguno de los antecedentes mostrados se puede apreciar, lo cual brinda un aporte adicional a la información que se presenta en las áreas de comercialización y producción, todo esto generó una gran aceptación por parte de los usuarios, los cuales mostraron su grado de satisfacción a través de una encuesta basada en la escala de Likert teniendo en cuenta un pre test y un post test.

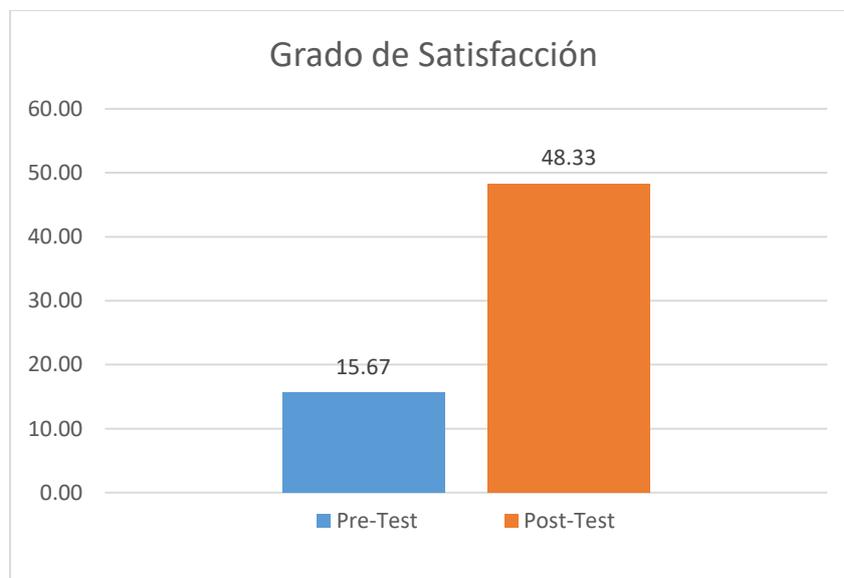


Figura 134 Grado de satisfacción

De acuerdo a los resultados encontrados podemos determinar que el impacto del data warehouse es positivo, ya que con la implementación se ha logrado la mejora del indicador de razón (tiempo de trabajo por cada reporte generado), traduciéndose esto como un ahorro económico para la empresa.

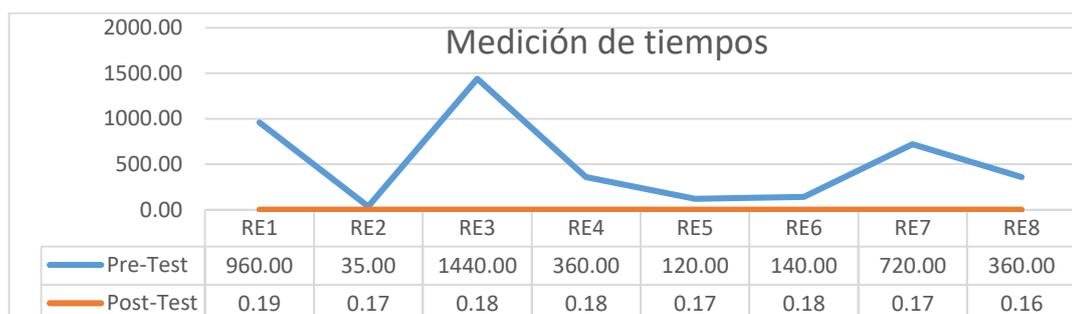


Figura 135 Gráfico de medición de tiempos

Finalmente con los indicadores económicos $VAN > 0$ que genera una rentabilidad de 28,450.71 nuevos soles en un periodo de 5 años, un TIR de 137% mayor a la tasa COK de %22.85 y un IR de 3.90, es decir, que por cada sol invertido se gana 2.90 soles. Demostrándose con esto que el proyecto es factible y generara rentabilidad a la empresa.

CAPÍTULO V CONCLUSIONES Y RECOMENDACIONES

5.1. Conclusiones

Se cumplió con el objetivo de implementar un Data Warehouse para optimizar la toma de decisiones en el área de comercialización y producción de la Industria Alimentaria Huacariz, para que de esta manera se pueda formular mejores estrategias y tomar decisiones más acertadas en las áreas estudiadas.

Se han logrado satisfacer los requerimientos de las áreas de comercialización y producción para el análisis de las ventas e insumos de las áreas respectivamente, obtenidos a través de reuniones con los encargados de tomar decisiones en cada área los cuales fueron descritos en la metodología empleada.

El diseño y carga del Data Warehouse con los datos proporcionados por el área de sistemas, si permitió obtener una salida de información, brindando así mejor análisis para apoyar la toma de decisiones en las áreas de comercialización y producción de la Industria Alimentaria Huacariz.

Los reportes fueron construidos con la información solicitada de cada área estudiada para la toma de decisiones, para estos reportes se utilizó la herramienta llamada QlikView, los cuales fueron aprobadas por las áreas respectivas.

La implementación del Data Warehouse para las áreas de comercialización y producción permite reducir el tiempo de espera por la información necesaria para la toma de decisiones en cuanto a las ventas e insumos en cada área respectivamente, además de ello se mejora los formatos en los que la información es presentada permitiendo un entendimiento más rápido de la misma, un punto más a rescatar, es que los usuarios tendrán acceso directo a la información que necesiten, dejando de lado la dependencia hacia el área de sistemas que venían mostrando.

5.2.Recomendaciones

Es recomendable realizar un Back Up de la información y del Data Warehouse, ante posibles dificultades, esto servirá para tener mayor seguridad.

La seguridad es un aspecto importante cuando de ventas e insumos hablamos, por tal motivo es importante cuidar la seguridad relacionada a esta información, por ello se debe tener un control en el acceso a los reportes presentados.

El éxito del Data Warehouse depende del uso constante de los usuarios de las áreas de comercialización y producción antes de tomar decisiones en sus respectivas áreas.

Se recomienda que se sanee la información del origen de datos de las áreas de comercialización y producción.

Entre otras opciones se recomienda revisar TM1 IBM ya que se perfila como una respuesta del business intelligence tradicional a los requerimientos que el mercado plantea, se presenta como una solución informática que minimiza los errores y multiplica la precisión en el análisis [37].

Como otra alternativa se recomienda revisar los sistemas de apoyo a la toma de decisiones (DDS) ya que son sistemas informáticos utilizados para servir de apoyo, más que automatizar el proceso de toma de decisiones [38].

REFERENCIAS BIBLIOGRÁFICAS

- [1] . E. Rizo Rizo, M. Tápanes Mora , J. P. Febles, V. Estrada Senti, E. Sánchez Pérez y R. Real Reyes, «Monografías de la Universidad de Matanzas Camilo Cienfuego,» [En línea]. Available: <http://monografias.umcc.cu/monos/2006/Informatica/Importancia%20de%20la%20utilizacion%20de%20un%20Data%20Warehouse.pdf>. [Último acceso: 01 Junio 2016].
- [2] J. Paz Flores, «IBM,» 30 Diciembre 2010. [En línea]. Available: <http://www.ibm.com/developerworks/ssa/local/data/dm-bi-pymes/>. [Último acceso: 02 Junio 2016].
- [3] A. ROJAS ZALDÍVAR, «REPOSITORIO ACADEMICO USMP,» 2014. [En línea]. Available: <http://www.repositorioacademico.usmp.edu.pe/handle/usmp/1061>. [Último acceso: 01 Junio 2016].
- [4] M. Ocas Terrones, «UPN Box Repositorio Institucional,» 2012. [En línea]. Available: <http://hdl.handle.net/11537/178>. [Último acceso: 15 Noviembre 2016].
- [5] «www.andina.com.pe,» Andina del Perú para el mundo, 30 07 2009. [En línea]. Available: <http://www.andina.com.pe/agencia/noticia-cajamarca-produce-unas-25-toneladas-semanales-queso-y-apunta-a-exportacion-245772.aspx>. [Último acceso: 28 10 2017].
- [6] F. S. Guillén Rodríguez, «UPN Box Repositorio Institucional,» 2012. [En línea]. Available: <http://hdl.handle.net/11537/123>. [Último acceso: 15 Noviembre 2016].
- [7] J. E. Córdova Yupanqui, «Repositorio Digital de Tesis PUCP,» 22 Agosto 2013. [En línea]. Available: <http://tesis.pucp.edu.pe/repositorio/handle/123456789/4709>. [Último acceso: 03 Junio 2016].
- [8] C. E. JARAMILLO MEJÍA, «Repositorio Digital de la Universidad Central del Ecuador,» 2015. [En línea]. Available: <http://www.dspace.uce.edu.ec/handle/25000/4195>. [Último acceso: 03 Junio 2016].
- [9] «Wikipedia,» Wikipedia, 7 Diciembre 2016. [En línea]. Available: https://es.wikipedia.org/wiki/Base_de_datos. [Último acceso: 12 Diciembre 2016].
- [10] «Tecnosoluciones,» 2002. [En línea]. Available: https://tecnosoluciones.com/librerias/Image/Carrusel_Contentidos/BasedeDatos3.jpg. [Último acceso: 12 Diciembre 2016].
- [11] «Wikipedia,» 27 Octubre 2016. [En línea]. Available: https://es.wikipedia.org/wiki/Sistema_de_informaci%C3%B3n. [Último acceso: 16 Noviembre 2016].
- [12] «Cibertareas,» 14 Agosto 2013. [En línea]. Available: <https://cibertareas.info/que-es-un-sistema-de-informacion-herramientas-informaticas-1.html>. [Último acceso: 16 Noviembre 2016].
- [13] M. Peralta, «Monografias.com,» 6 Septiembre 2001. [En línea]. Available: <http://www.monografias.com/trabajos7/sisinf/sisinf.shtml>. [Último acceso: 16 Noviembre 2016].
- [14] «sinnexus,» sinnexus, 2016. [En línea]. Available: http://www.sinnexus.com/business_intelligence/datawarehouse.aspx. [Último acceso: 20 06 2017].
- [15] Y. Torres, «Wikipedia,» 31 Enero 2013. [En línea]. Available: <http://datawarehouseunefa.blogspot.pe/>. [Último acceso: 26 Octubre 2016].
- [16] C. Herrera, «AdictosAlTrabajo.com,» 30 Octubre 2007. [En línea]. Available:

- <https://www.adictosaltrabajo.com/tutoriales/datawarehouse/>. [Último acceso: 26 Octubre 2016].
- [17] «Evaluando software,» 21 Septiembre 2016. [En línea]. Available: <http://www.evaluandosoftware.com/abc-del-data-mart/>. [Último acceso: 2 Noviembre 2016].
- [18] «Sinnexus,» Sinnexus, 2007. [En línea]. Available: http://www.sinnexus.com/business_intelligence/index.aspx. [Último acceso: 04 Junio 2016].
- [19] «Sinnexus,» Sinnexus, 2016. [En línea]. Available: http://www.sinnexus.com/business_intelligence/arquitectura.aspx. [Último acceso: 29 10 2017].
- [20] «Wikipedia,» 21 Septiembre 2016. [En línea]. Available: https://es.wikipedia.org/wiki/Modelado_dimensional. [Último acceso: 7 Noviembre 2016].
- [21] «Blogger,» 30 Enero 2014. [En línea]. Available: <http://inteligenciadenegociosval.blogspot.pe/2014/01/metodologia-de-kimball.html>. [Último acceso: 10 Noviembre 2016].
- [22] G. R. Rivadera, Mayo 2010. [En línea]. Available: <http://www.ucasal.edu.ar/htm/ingenieria/cuadernos/archivos/5-p56-rivadera-formateado.pdf>. [Último acceso: 14 Noviembre 2016].
- [23] «Wikipedia,» 10 Enero 2017. [En línea]. Available: https://es.wikipedia.org/wiki/Toma_de_decisiones. [Último acceso: 25 Enero 2017].
- [24] M. B. Ferrero y O. J. A. Chiotti, «SCRIBD,» [En línea]. Available: <https://es.scribd.com/document/75063150/Cap-11>. [Último acceso: 24 10 2017].
- [25] J. Pérez Porto y A. Gardey, «Definicion.de,» 2012. [En línea]. Available: <http://definicion.de/informacion/>. [Último acceso: 25 Enero 2017].
- [26] «Wikipedia,» 15 Enero 2017. [En línea]. Available: <https://es.wikipedia.org/wiki/Informaci%C3%B3n>. [Último acceso: 25 Enero 2017].
- [27] «IBM,» IBM, [En línea]. Available: https://www.ibm.com/support/knowledgecenter/es/SS9UM9_9.1.0/com.ibm.datatools dimensional.ui.doc/topics/c_dm_star_schemas.html. [Último acceso: 19 10 2017].
- [28] F. «Business Intelligence,» Blogger, 04 09 2011. [En línea]. Available: <http://biverano2011.blogspot.pe/2011/09/modelo-estrella-y-modelo-copo-de-nieve.html>. [Último acceso: 19 10 2017].
- [29] «Wikipedia,» Wikipedia, 10 11 2013. [En línea]. Available: https://es.wikipedia.org/wiki/Esquema_en_copo_de_nieve. [Último acceso: 19 10 2017].
- [30] «IBM,» IBM, [En línea]. Available: https://www.ibm.com/support/knowledgecenter/es/SS9UM9_9.1.1/com.ibm.datatools dimensional.ui.doc/topics/c_dm_snowflake_schemas.html. [Último acceso: 19 10 2017].
- [31] «Wikipedia,» Esquema en estrella, 05 08 2017. [En línea]. Available: https://es.wikipedia.org/wiki/Esquema_en_estrella. [Último acceso: 19 10 2017].
- [32] «Microsoft,» Microsoft, 14 03 2017. [En línea]. Available: <https://docs.microsoft.com/en-us/sql/integration-services/sql-server-integration-services>. [Último acceso: 24 09 2017].
- [33] «Learn MSBI tutorials,» 2017. [En línea]. Available: <http://www.learnmsbitutorials.net/etl-process-ssis-example.php>. [Último acceso: 25

- septiembre 2017].
- [34] C. Fernández, «Data Prix,» 22 07 2009. [En línea]. Available: <http://www.dataprix.com/arquitectura-data-warehouse-areas-datos-nuestro-almacen-corporativo>. [Último acceso: 15 08 2017].
 - [35] «Logicalis,» 19 03 2014. [En línea]. Available: <https://blog.es.logicalis.com/analytics/modelado-y-diseo-de-ods-y-dwh>. [Último acceso: 16 08 2017].
 - [36] M. A. Garcia y B. Harmsen, QlikView 11 para Desarrolladores, Packt Publishing Ltd, 2012.
 - [37] «<https://blog.es>,» Logicalis, 06 04 2014. [En línea]. Available: <https://blog.es.logicalis.com/analytics/como-optimizar-la-toma-de-decisiones-con-el-tm1-ibm>. [Último acceso: 28 11 2017].
 - [38] «<https://prezi.com>,» Prezi, 19 08 2010. [En línea]. Available: <https://prezi.com/mqlq7li5inbr/software-de-apoyo-de-toma-de-decisiones-y-toma-de-decisiones-en-grupo/>. [Último acceso: 28 11 2017].
 - [39] E. M. d. C. Palma Serrano, «Gestiopolis,» 14 Noviembre 2014. [En línea]. Available: <http://www.gestiopolis.com/inteligencia-de-negocios-business-intelligence/>. [Último acceso: 4 Noviembre 2016].
 - [40] R. Kimball y M. Ross, The Data Warehouse Toolkit, Third Edition, Indianápolis: John Wiley & Sons, Inc., 2013.
 - [41] R. Hernández Sampieri, C. Fernández Collado y M. d. P. Baptista Lucio, Metodología de la Investigación, Quinta Edición, 2010.

ANEXOS

ANEXO 1: MODELO DE ENCUESTA – GRADO DE SATISFACCIÓN

Encuesta - Grado de Satisfacción

1. El acceso a la información que necesita para la toma de decisiones es directo.
 - a) Muy satisfecho
 - b) Satisfecho
 - c) Indeciso
 - d) Poco satisfecho
 - e) Para nada satisfecho

2. El tiempo de espera para recibir la información solicitada no excede los 60 minutos.
 - a) Muy satisfecho
 - b) Satisfecho
 - c) Indeciso
 - d) Poco satisfecho
 - e) Para nada satisfecho

3. El formato en el que se recibe la información permite el análisis inmediato de la misma.
 - a) Muy satisfecho
 - b) Satisfecho
 - c) Indeciso
 - d) Poco satisfecho
 - e) Para nada satisfecho

4. La interfaz es de fácil entendimiento
 - a) Muy satisfecho
 - b) Satisfecho
 - c) Indeciso
 - d) Poco satisfecho
 - e) Para nada satisfecho

5. El procesamiento y análisis de reportes no demanda tiempo extra.
 - a) Muy satisfecho
 - b) Satisfecho
 - c) Indeciso
 - d) Poco satisfecho
 - e) Para nada satisfecho

6. De forma general, los reportes recibidos le brindan información útil para la toma de decisiones.
 - a) Muy satisfecho
 - b) Satisfecho
 - c) Indeciso
 - d) Poco satisfecho
 - e) Para nada satisfecho

7. El nivel de detalle de la información mostrada es el adecuado para el análisis.
 - a) Muy satisfecho
 - b) Satisfecho
 - c) Indeciso
 - d) Poco satisfecho
 - e) Para nada satisfecho

8. El sistema actual le brinda información histórica para contrastar en el tiempo.
 - a) Muy satisfecho
 - b) Satisfecho
 - c) Indeciso
 - d) Poco satisfecho
 - e) Para nada satisfecho

9. El proceso que sigue para obtener la información deseada es práctico y sencillo.
 - a) Muy satisfecho
 - b) Satisfecho
 - c) Indeciso
 - d) Poco satisfecho
 - e) Para nada satisfecho

10. Mejor control sobre las ventas y la producción
 - a) Muy satisfecho
 - b) Satisfecho
 - c) Indeciso
 - d) Poco satisfecho
 - e) Para nada satisfecho