

UNIVERSIDAD TÉCNICA DE AMBATO



FACULTAD DE INGENIERÍA EN SISTEMAS, ELECTRÓNICA E INDUSTRIAL

CENTRO DE ESTUDIOS DE POSGRADO

MAESTRÍA EN GESTIÓN DE BASES DE DATOS

TEMA:

LA TECNOLOGÍA DATAWAREHOUSE Y SU INCIDENCIA EN LA
GESTIÓN DE INFORMACIÓN GERENCIAL DE CURTIDURÍA
TUNGURAHUA S.A. DE LA CIUDAD DE AMBATO

Trabajo de Investigación

Previo a la obtención del Grado Académico de Magíster en Gestión de Bases de
Datos

Nombre del Autor: Ing. Carmen Elena Sánchez Barreno

Nombre del Director: Ing. Mg. Elsa Pilar Urrutia Urrutia.

Ambato - Ecuador

2012

Al Consejo de Posgrado de la UTA.

El tribunal receptor de la defensa del trabajo de investigación con el tema: “**LA TECNOLOGÍA DATAWAREHOUSE Y SU INCIDENCIA EN LA GESTIÓN DE INFORMACIÓN GERENCIAL DE CURTIDURÍA TUNGURAHUA S.A. DE LA CIUDAD DE AMBATO**” Presentada por: Ing. Carmen Elena Sánchez Barreno y Conformada por: Ing. Mg. David Omar Guevara Aulestia, Ing. Mg. Franklin Oswaldo Mayorga Mayorga, Ing. Mg. Edwin Hernando Buenaño Valencia, Miembros del Tribunal de Defensa, Ing. Mg. Elsa Pilar Urrutia Urrutia, Director del trabajo de investigación y presidido por: Ing. M.Sc., Oswaldo Paredes Ochoa, Presidente del Tribunal; Ing. Mg. Juan Garcés Chávez, Director del CEPOS-UTA, una vez escuchada la defensa oral el Tribunal aprueba y remite el trabajo de investigación para usoy custodia en las bibliotecas de la UTA.

Ing. M.Sc. Oswaldo Paredes Ochoa

Presidente del Tribunal de Defensa

Ing. Mg. Juan Garcés Chávez

Director CEPOS

Ing. Mg. Elsa Pilar Urrutia Urrutia
Directora del Trabajo de Investigación

Ing. Mg. David Omar Guevara Aulestia
Miembro del Tribunal

Ing. Mg. Franklin Oswaldo Mayorga Mayorga
Miembro del Tribunal

Ing. Mg. Edwin Hernando Buenaño Valencia
Miembro del Tribunal

AUTORÍA DE LA INVESTIGACIÓN

La responsabilidad de las opiniones, comentarios y críticas emitidas en el trabajo de investigación con el tema: **“LA TECNOLOGÍA DATAWAREHOUSE Y SU INCIDENCIA EN LA GESTIÓN DE INFORMACIÓN GERENCIAL DE CURTIDURÍA TUNGURAHUA S.A. DE LA CIUDAD DE AMBATO”**, nos corresponde exclusivamente a Ing. Carmen Elena Sánchez Barreno, Autor y Ing. Mg. Elsa Pilar Urrutia, Directora del Trabajo de Investigación; y el patrimonio intelectual del mismo a la Universidad Técnica de Ambato.

Ing. Carmen Elena Sánchez B
Autor

Ing. Mg. Elsa Pilar Urrutia Urrutia
Directora del Trabajo de Investigación

DERECHOS DE AUTOR

Autorizo a la Universidad Técnica de Ambato, para que haga de este trabajo de investigación o parte de él un documento disponible para su lectura, consulta y procesos de investigación, según normas de la Institución.

Cedo los Derechos de mi trabajo de investigación, con fines de difusión pública, además apruebo la reproducción de esta, dentro de las regulaciones de la Universidad.

Ing.Carmen Elena Sánchez Barreno

Dedicatoria

A Dios.

Por haberme permitido llegar hasta este punto y haberme dado salud para lograr mis objetivos, además de su infinita bondad y amor.

A mi esposo Juan Carlos

Por haberme apoyado en todo momento, por ser el pilar fundamental que inspira mi superación personal y además ser la razón de mi dedicación desde que forma parte de mi vida.

A mis padres Fernando y Carmen

Por los ejemplos de perseverancia y constancia que los caracteriza y que me han infundado siempre, por sus consejos, sus valores, por la motivación constante que me ha permitido ser una persona de bien, pero más que nada, por su amor.

Elena.

Agradecimiento

A todas las personas que apoyaron a la culminación de este proyecto, en especial a mi tutora Ing. Elsa Pilar Urrutia Urrutia, por su paciencia, conocimiento, y muy acertada dirección.

A la Universidad Técnica de Ambato y a la Facultad de Ingeniería en Sistemas, Electrónica e Industrial especialmente al noble equipo de docentes que a diario comparten sus conocimientos.

Elena.

ÍNDICE GENERAL

PRELIMINARES	PÁGINAS
Portada	i
Autoría de Tesis	iii
Derechos de Autor	iv
Dedicatoria	v
Agradecimiento	vi
Índice General	vii
Índice de Tablas	x
Índice de Figuras	xi
Resumen Ejecutivo	xv
Introducción	xvii
CAPÍTULO I	
EL PROBLEMA	
Planteamiento del problema	1
Contextualización	1
Árbol de Problemas	4
Análisis Crítico	5
Prognosis	6
Formulación del Problema	7
Interrogantes de la Investigación	7
Delimitación de Contenido	7
Delimitación Espacial	8
Delimitación Temporal	8
Unidades de Observación	8
Justificación	9
Objetivos	11
Objetivo General	11
Objetivos Específicos	11
CAPÍTULO II	
MARCO TEÓRICO	
Antecedentes Investigativos	12
Filosófica	13
Tecnológica	14
Administrativa	14
Legal	15
Categorías Fundamentales	16
Red de Inclusiones Conceptuales	17

Constelación de Ideas de la Variable Independiente	18
Constelación de Ideas de la Variable Dependiente	19
Categorías Variable Independiente	20
Categorías Variable Dependiente	36
Hipótesis	48

CAPÍTULO III

METODOLOGÍA

Modalidades básicas de investigación	49
Investigación de Campo	49
Investigación Documental-bibliográfica	49
Tipos o niveles de investigación	50
Exploratorio	50
Descriptivo	50
Población y Muestra	50
Operacionalización de Variables	52
Técnicas e Instrumentos	54
Validez y Confiabilidad	54
Plan de Recolección de Información	54
Plan de Procesamiento de Información	55

CAPÍTULO IV

ANÁLISIS E INTERPRETACION DE RESULTADOS

Análisis e interpretación de resultados	61
Verificación de la hipótesis	82
Método Lógico	85
Método Estadístico	86

CAPÍTULO V

CONCLUSIONES Y RECOMENDACIONES

Conclusiones	90
Recomendaciones	91

CAPÍTULO VI

LA PROPUESTA

Datos Informativos	92
Antecedentes de la propuesta	93
Justificación	93
Objetivos	95
Objetivo General	95
Objetivos Específicos	95
Análisis de Factibilidad	96
Fundamentación	98
Definiciones generales	98

METODOLOGÍA	113
ANÁLISIS	114
Evaluación de costes en la empresa	114
Evaluación de la infraestructura de la empresa	115
Planeamiento del proyecto	118
Definición de Requisitos	118
Análisis de las fuentes de datos	119
Auditoria de datos	119
Definición del plan	122
DISEÑO	123
Modelo dimensional	123
Arquitectura del Datawarehouse	129
Diseño Procesos ETL (Extracción, Transformación y Carga de datos)	129
Extracción	130
Transformación y Carga	138
IMPLEMENTACION	144
Construcción de los cubos de información	144
Optimización para mejorar desempeño de consultas	157
Construcción del DTS definitivo	165
Creación del Tarea programada para la ejecución del DTS	167
Reporting Services	168
IMPLANTACION	174
Configurando consulta de datos	174
Reportes	177
Plan de Capacitación a usuarios	184
CONCLUSIONES Y RECOMENDACIONES	188
Conclusiones	188
Recomendaciones	190
BIBLIOGRAFÍA	191
ANEXOS	194

ÍNDICE DE TABLAS

Tabla N° 3.1 Población	51
Tabla N°3.2 Operacionalización de variables independiente	52
Tabla N° 3.3 Operacionalización de variables dependiente	53
Tabla N° 3.4 Encuesta N ° 1	57
Tabla N° 3.5 Encuesta N ° 1	59
Tabla N° 3.6 Entrevista	60
Tabla N° 4.1 Tabulación de respuesta a encuesta 1 pregunta 1	61
Tabla N° 4.2 Tabulación de respuesta a encuesta 1 pregunta 2	62
Tabla N° 4.3 Tabulación de respuesta a encuesta 1 pregunta 3	63
Tabla N° 4.4 Tabulación de respuesta a encuesta 1 pregunta 4	64
Tabla N° 4.5 Tabulación de respuesta a encuesta 1 pregunta 5	65
Tabla N° 4.6 Tabulación de respuesta a encuesta 1 pregunta 6	66
Tabla N° 4.7 Tabulación de respuesta a encuesta 1 pregunta 7	67
Tabla N° 4.8 Tabulación de respuesta a encuesta 1 pregunta 8	68
Tabla N° 4.9 Tabulación de respuesta a encuesta 1 pregunta 9	69
Tabla N° 4.10 Tabulación de respuesta a encuesta 1 pregunta 10	70
Tabla N° 4.11 Tabulación de respuesta a encuesta 2 pregunta 1	71
Tabla N° 4.12 Tabulación de respuesta a encuesta 2 pregunta 2	72
Tabla N° 4.13 Tabulación de respuesta a encuesta 2 pregunta 3	73
Tabla N° 4.14 Tabulación de respuesta a encuesta 2 pregunta 4	74
Tabla N° 4.15 Tabulación de respuesta a encuesta 2 pregunta 5	75
Tabla N° 4.16 Tabulación de respuesta a encuesta 2 pregunta 6	76
Tabla N° 4.17 Tabulación de respuesta a encuesta 2 pregunta 7	77
Tabla N° 4.18 Tabulación de respuesta a encuesta 2 pregunta 8	78
Tabla N° 4.19 Tabulación de respuesta a encuesta 2 pregunta 9	79
Tabla N° 4.20 Tabulación de respuesta a encuesta 2 pregunta 10	80
Tabla N° 4.21 Tabulación de respuesta a encuesta 2 pregunta 11	81
Tabla N° 4.22: Comparativo de proceso actual vs esperado	83
Tabla N° 4.23 Comprobación de hipótesis: Aplicaciones para procesar información	87
Tabla N° 4.24 Comprobación de hipótesis: Herramienta de toma de decisiones	87
Tabla N° 4.25 Frecuencias Observadas	88
Tabla N° 4.26 Frecuencias Esperadas	88
Tabla N° 4.27 Cálculo de χ^2	89
Tabla N° 6.1 Detalle de tablas sistema de gestión	119
Tabla N° 6.2: Grupo de capacitación Producción	184
Tabla N° 6.3: Grupo de capacitación Ventas	185
Tabla N° 6.4: Grupo de capacitación Logística de despachos	185
Tabla N° 6.5: Grupo de capacitación Logística de abastecimiento	185
Tabla N° 6.6: Cronograma de Capacitación	187

ÍNDICE DE FIGURAS

Fig. 1.1: Árbol de problemas	4
Fig. 2.1: Categorías Fundamentales	17
Fig. 2.2: Subcategorías de la VI	18
Fig. 2.3: Subcategorías de la VD	19
Fig. 2.4: Sistema DSS	27
Fig. 2.5: Representación gráfica de un hecho y sus dimensiones	29
Fig. 2.6: Diagrama estrella del hecho	32
Fig. 2.7: Cubo de datos	33
Fig. 2.8: Vista gráfica, del paquete creado	35
Fig. 4.1: Tabulación de respuesta a encuesta 1 pregunta 1	61
Fig. 4.2: Tabulación de respuesta a encuesta 1 pregunta 2	62
Fig. 4.3: Tabulación de respuesta a encuesta 1 pregunta 3	63
Fig. 4.4: Tabulación de respuesta a encuesta 1 pregunta 4	64
Fig. 4.5: Tabulación de respuesta a encuesta 1 pregunta 5	65
Fig. 4.6: Tabulación de respuesta a encuesta 1 pregunta 6	66
Fig. 4.7: Tabulación de respuesta a encuesta 1 pregunta 7	67
Fig. 4.8: Tabulación de respuesta a encuesta 1 pregunta 8	68
Fig. 4.9: Tabulación de respuesta a encuesta 1 pregunta 9	69
Fig. 4.10: Tabulación de respuesta a encuesta 1 pregunta 10	70
Fig. 4.11: Tabulación de respuesta a encuesta 2 pregunta 1	71
Fig. 4.12: Tabulación de respuesta a encuesta 2 pregunta 2	72
Fig. 4.13: Tabulación de respuesta a encuesta 2 pregunta 3	73
Fig. 4.14: Tabulación de respuesta a encuesta 2 pregunta 4	74
Fig. 4.15: Tabulación de respuesta a encuesta 2 pregunta 5	75
Fig. 4.16: Tabulación de respuesta a encuesta 2 pregunta 6	76
Fig. 4.17: Tabulación de respuesta a encuesta 2 pregunta 7	77
Fig. 4.18: Tabulación de respuesta a encuesta 2 pregunta 8	78
Fig. 4.19: Tabulación de respuesta a encuesta 2 pregunta 9	79
Fig. 4.20: Tabulación de respuesta a encuesta 2 pregunta 10	80
Fig. 4.21: Tabulación de respuesta a encuesta 2 pregunta 11	81
Fig. 6.1: Elementos de un datawarehouse	100
Fig. 6.2: Procesos del datawarehouse	105
Fig. 6.3: Metodología Inmon	108
Fig. 6.4: Metodología Kimball	109
Fig. 6.5: Metodología del datawarehouse	113
Fig. 6.6: Organigrama Estructural Administración	116
Fig. 6.7: Organigrama Estructural Producción	117
Fig. 6.8: Datamart Formulación	124
Fig. 6.9: Datamart Paquetes	125
Fig. 6.10: Datamart Pedidos	126

Fig.6.11: Datamart Transferencias	127
Fig. 6.12: Datamart Ventas	128
Fig. 6.13: Arquitectura del datawarehouse	129
Fig. 6.14: Diagrama de procesos ETL	130
Fig. 6.15: Archivos planos de Oracle	131
Fig. 6.16: Creación de un paquete	131
Fig. 6.17: Creación de enlaces a archivos planos	132
Fig. 6.18: Conexión archivos planos	132
Fig. 6.19: Configuración de separadores de columna	133
Fig. 6.20: Delimitación de columnas	133
Fig. 6.21: Herramienta para configurar destino	134
Fig. 6.22 Configuración del servidor destino	134
Fig. 6.23: Diagrama Final de extracción	135
Fig. 6.24: Propiedades de transformación origen.	136
Fig. 6.25: Propiedades de transformación destino.	136
Fig. 6.26: Propiedades de transformación.	137
Fig. 6.27: Creación de BDD SATGE	138
Fig. 6.28: Creación de BDD DWH.	139
Fig. 6.29: Creación de BDD DW1	140
Fig. 6.30: Listado de procedimientos ETL	140
Fig. 6.31: Procedimiento almacenado para cargar Transferencias.	141
Fig. 6.32: Procedimiento almacenado para cargar Ventas	141
Fig. 6.33: Procedimiento almacenado para cargar dimensión fecha	142
Fig. 6.34: Vista para cargar dimensión material (modelo Copo de nieve)	143
Fig. 6.35: Diagrama final procesos ETL	143
Fig. 6.36: Cubos de información	144
Fig. 6.37: Base DW1 en Analysis Services.	145
Fig. 6.38: Conexión al DWH desde Analysis Manager	145
Fig. 6.39: Proveedor de Conexión al DWH desde Analysis Services	146
Fig. 6.40: Cadena de Conexión al DWH desde Analysis Services	146
Fig. 6.41: Opciones de creación de dimensiones	147
Fig. 6.42: Selección de la tabla o vista que conformara la dimensión	147
Fig. 6.43: Selección tipo de dimensión	148
Fig. 6.44: Niveles de una dimensión	148
Fig. 6.45: Opciones avanzadas de una dimensión	149
Fig. 6.46: Dimensión creada	149
Fig. 6.47: Dimensiones publicas	150
Fig. 6.48: Edición de dimensiones	150
Fig. 6.49: Pre visualización de datos	151
Fig. 6.50: Nuevo cubo	151
Fig. 6.51: Selección de la tabla de hechos	152
Fig. 6.52: Editor de cubos	153
Fig. 6.53: Editor cubos	153

Fig. 6.54: Editor cubos- Agregar dimensiones	154
Fig. 6.55: Selección de la dimensión ara el cubo	154
Fig. 6.56: Dimensión Agregada	155
Fig. 6.57: Agregar dimensiones privadas	155
Fig. 6.58: Cubo ventas final tipo estrella	156
Fig. 6.59: Procesamiento del cubo	156
Fig. 6.60: Método de procesamiento del cubo	157
Fig. 6.61: Diseño de agregaciones	160
Fig. 6.62: Rendimiento cubo de información	161
Fig. 6.63: Procesamiento de las agregaciones en el cubo	161
Fig. 6.64: Procesamiento del cubo	162
Fig. 6.65: Selección de los procedimientos almacenados a optimizar	163
Fig. 6.66: Revisión procedimientos almacenados optimizado	163
Fig. 6.67: Opciones para procesar las agregaciones	164
Fig.6.68: Procesamiento de la optimización basada en el uso	164
Fig. 6.69: Tarea de ejecución de limpieza de tablas	165
Fig. 6.70: Tarea de ejecución de paquetes	165
Fig. 6.71: Tarea de ejecución de procedimientos de carga	166
Fig. 6.72: Tarea de ejecución del cubo	166
Fig. 6.73: Diagrama final de ejecución	167
Fig. 6.74: Calendarización tarea programada de procesamiento	167
Fig. 6.75: Panel de monitoreo del tarea programada de procesamiento	168
Fig. 6.76: Creación de un proyecto de report services	170
Fig. 6.77: Definición del origen para el report services	170
Fig. 6.78: Definición de la consulta a reportearse	171
Fig. 6.79: Tipo de reporte	171
Fig. 6.80: Definición del diseño del reporte	172
Fig. 6.81: Definición del estilo del reporte	172
Fig. 6.82: Definición de la ubicación del reporte	173
Fig. 6.83: Configuración de Ms Excel para la obtención del reporte	174
Fig. 6.84: Elección del origen de datos	174
Fig. 6.85: Creación del origen	175
Fig. 6.86: Conexión con analysis server	175
Fig. 6.87: Selección del cubo	176
Fig. 6.88: Importar datos	176
Fig. 6.89: Reporte Histórico de pedidos	177
Fig. 6.90: Reporte Histórico de Administración de bodega pt	177
Fig. 6.91: Reporte Histórico de reprocesos	178
Fig. 6.92: Reporte Stock disponible en bodega	178
Fig. 6.93: Reporte Histórico de ventas por cliente	179
Fig. 6.94: Reporte Histórico de consumos de MP	179
Fig. 6.95: Reporte Histórico de ingresos	180
Fig. 6.96: Reporte Histórico de movimientos	180

Fig. 6.97: Reporte de costes de fórmulas	181
Fig. 6.98: Panel de Administración del Reporting Services	181
Fig. 6.99: Reporte dentro del Reporting Services	182
Fig. 6.100: Reporte Histórico de órdenes de producción	182
Fig. 6.101: Reporte Histórico de clientes	183

UNIVERSIDAD TÉCNICA DE AMBATO
FACULTAD DE INGENIERÍA EN SISTEMAS, ELECTRÓNICA E
INDUSTRIAL
MAESTRÍA EN GESTIÓN DE BASES DE DATOS

“LA TECNOLOGÍA DATAWAREHOUSE Y SU INCIDENCIA EN LA
GESTIÓN DE INFORMACIÓN GERENCIAL DE CURTIDURÍA
TUNGURAHUA S.A. DE LA CIUDAD DE AMBATO”

Autor: Ing. Carmen Elena Sánchez Barreno

Tutor: Ing. Mg. Elsa Pilar Urrutia Urrutia

Fecha: 19 de Septiembre del 2012

RESUMEN

Curtiduría Tungurahua S.A, es una empresa líder en la industria del procesamiento del cuero, aquí se ha identificado los problemas que se generan en la toma de decisiones sobre la gestión de la información determinando que estos generan pérdidas en la eficiencia de recursos tanto humano como económico, y su principal inconveniente, es que carece de una herramienta para la gestión rápida y oportuna de la información generada por los distintos centros de costos de la empresa.

Este tema es enfrentado con la presente investigación planteando una alternativa de solución basada en la tecnología datawarehouse, conocida herramienta para análisis de información, cuyo desarrollo arrojará datos confiables con los cuales la gerencia de producción podrá contar para la toma de decisiones acertadas y oportunas.

DESCRIPTORES

Inteligencia de negocios, sistema de soporte de decisiones, datawarehouse, procesamiento analítico en línea

TECHNICAL UNIVERSITY OF AMBATO
FACULTY OF ENGINEERING SYSTEMS, ELECTRONICS AND
INDUSTRIAL
MASTERS DEGREE IN DATABASES MANAGEMENT

"DATAWAREHOUSE TECHNOLOGY AND ITS IMPACT ON
MANAGEMENT INFORMATION MANAGEMENT IN CURTIDURÍA
TUNGURAHUA SA OF AMBATO CITY "

Author: Engineer Carmen Elena Sanchez Barreno

Tutor: Engineer Mg. Elsa Pilar Urrutia Urrutia

Date: September 19, 2012

SUMMARY

Curtiduria Tungurahua SA, is a leader in the leather processing industry here has identified the problems generated in making management decisions determining information they generate efficiency losses of human and economic resources and its main drawback is that it lacks a tool for quick and timely management of information generated by the various cost centers of the company.

This subject is confronted with this investigation proposing an alternative solution based on data warehouse technology, known tool for data analysis, the development threw reliable data with which management of production may have for making sound and timely decisions.

DESCRIPTORS

Business intelligence, decision support systems, data warehousing, online analytical processing

INTRODUCCIÓN

A pesar que las empresas dependen de condiciones externas para mantenerse entre las mejores del país, es muy importante para estas sociedades e instituciones, manejar de forma adecuada los recursos tecnológicos y de información para desarrollar sistemas automatizados que sean consecuentes a la mejora y rápida ejecución de sus procesos.

El plan de tesis está estructurado por capítulos:

El Capítulo I denominado EL PROBLEMA contiene: Contextualización, análisis crítico, prognosis, formulación del problema, interrogantes (sub problemas), delimitación del objetivo de investigación, justificación y objetivos (general y específico), con el fin de calificar el contexto de esta investigación.

El Capítulo II denominado MARCO TEORICO se estructura con: los antecedentes investigativos, las fundamentaciones filosóficas, tecnológica, legal, red de inclusiones conceptuales, constelaciones de ideas, hipótesis con sus respectivas variables, es decir la fundamentación teórica sobre esta tecnología y su aplicación,

El Capítulo III denominado METODOLOGIA contiene: modalidades de investigación, nivel o tipo de investigación, población y muestra, operacionalización de variables, plan de recolección de información.

El Capítulo IV denominado ANALISIS E INTERPRETACION DE RESULTADOS como su nombre lo indica conformado por: el análisis e interpretación de la información recolectada a través de distintas herramientas usadas en la presente investigación.

El Capítulo V denominado CONCLUSIONES Y RECOMENDACIONES cuyo contenido determina las conclusiones obtenidas fruto del trabajo investigativo y a su vez establece algunas recomendaciones del trabajo realizado.

El Capítulo VI denominado LA PROPUESTA cuyo contenido establece el proyecto de solución a la problemática investigada en base a la tecnología aplicada, su funcionamiento y pruebas de operación en beneficio de Curtiduría Tungurahua S.A.

Finalmente se encuentra la biografía, que ha servido de apoyo a la presente investigación, y anexos con información adicional, que facilitaran el entendimiento del presente trabajo.

CAPÍTULO I

EL PROBLEMA

Tema: “LA TECNOLOGÍA DATAWAREHOUSE Y SU INCIDENCIA EN LA GESTIÓN DE INFORMACIÓN GERENCIAL DE CURTIDURÍA TUNGURAHUA S.A. DE LA CIUDAD DE AMBATO.”

Planteamiento del Problema

Contextualización

Siguiendo a (Inmon, 1992), un Datawarehouse se define “**como un conjunto de datos orientados por tema, integrados, variables en el tiempo y no volátiles que se emplean como apoyo a la toma de decisiones administrativas**”. Se trata de una tecnología que permite desarrollar un análisis multidimensional de los datos, y que además, favorece la organización de la información de acuerdo con ciertos parámetros establecidos, sobre los que se lleva a cabo el análisis de la información. Su diferencia fundamental con las bases de datos tradicionales manejadas en las empresas, reside en que los datos no son

volátiles, con lo que estamos refiriéndonos a bases de datos de una dimensión superior a las anteriores que permiten registrar los datos de modo permanente.

Ante este nuevo entorno empresas del sector productivo especialmente de la industria del cuero enfrentan un periodo de grandes cambios que han sido propiciados por causas de diferente índole.

Curtiduría Tungurahua S.A. es una empresa con más de 70 años de experiencia, que proporciona estabilidad a más de 100 familias, sin embargo la experiencia y permanencia en el mercado no significan que sea la más eficiente y de mayor rentabilidad en sus procesos de producción.

El incremento de la competencia obliga a realizar un análisis minucioso de cómo obtener rentabilidad ante los bajos precios conservando la calidad; sin embargo Curtiduría Tungurahua S.A atraviesa una serie de problemas que dificultan la toma de ciertas decisiones gerenciales que ratifiquen a la empresa como la número uno a nivel nacional dentro de la industria de la producción del cuero.

El principal problema por el que atraviesa la empresa radican en la falta de información proporcionada por el proceso mismo de producción, información confiable por ejemplo determinar el costo de elaboración de un producto de manera oportuna y eficiente, es decir la empresa no tiene la certeza de las conveniencias sobre la elaboración o no de un producto, sobre la atención o no a un pedido de un cliente.

La empresa durante los últimos seis meses ha venido realizando un excesivo “reproceso” a sus productos finales para lograr la calidad a la cual están acostumbrados sus clientes, lo que causa un consumo excesivo de materia prima cuyo costo como producto final incrementa afectando directamente la rentabilidad

del mismo, inconveniente que arrastra consigo un serie de complicaciones como desabastecimientos de materia prima y suministros, lo que afecta a otros departamentos como el de compras, que tiene que incrementar sus costos de importación ya que deben pagar fletes más caros por la premura del pedido o afecta directamente al departamento de ventas que se ve obligado al incumplimiento de los pedidos con los clientes, o a realizar tratos menos beneficios para la salida del producto.

Otro aspecto a considerar como obstáculo al desarrollo de la empresa es que la poca información con la que cuenta es poco confiable ya que al sistema actual de producción es alimentado muchas veces de manera errónea, lo que causa una presentación de informes a destiempo y muchas veces alejado a la realidad empresarial.

Por otro lado la falta de capacitación y la renovación del personal han contribuido considerablemente a que los problemas anteriormente mencionados se agraven ratificándose el hecho de realizar una análisis minucioso para determinar una posible solución que permita disponer de información valiosa que ayude a Curtiduría Tungurahua S.A a fijar estrategias y tomar decisiones que favorezcan el desarrollo del negocio.

En función de lo expuesto, la información y el conocimiento del sector se convierten en dos factores claves para aumentar la competitividad de una empresa, y por tanto, el manejo inteligente de este activo constituye un factor crítico para la empresa actual, del cual va a depender su competitividad.

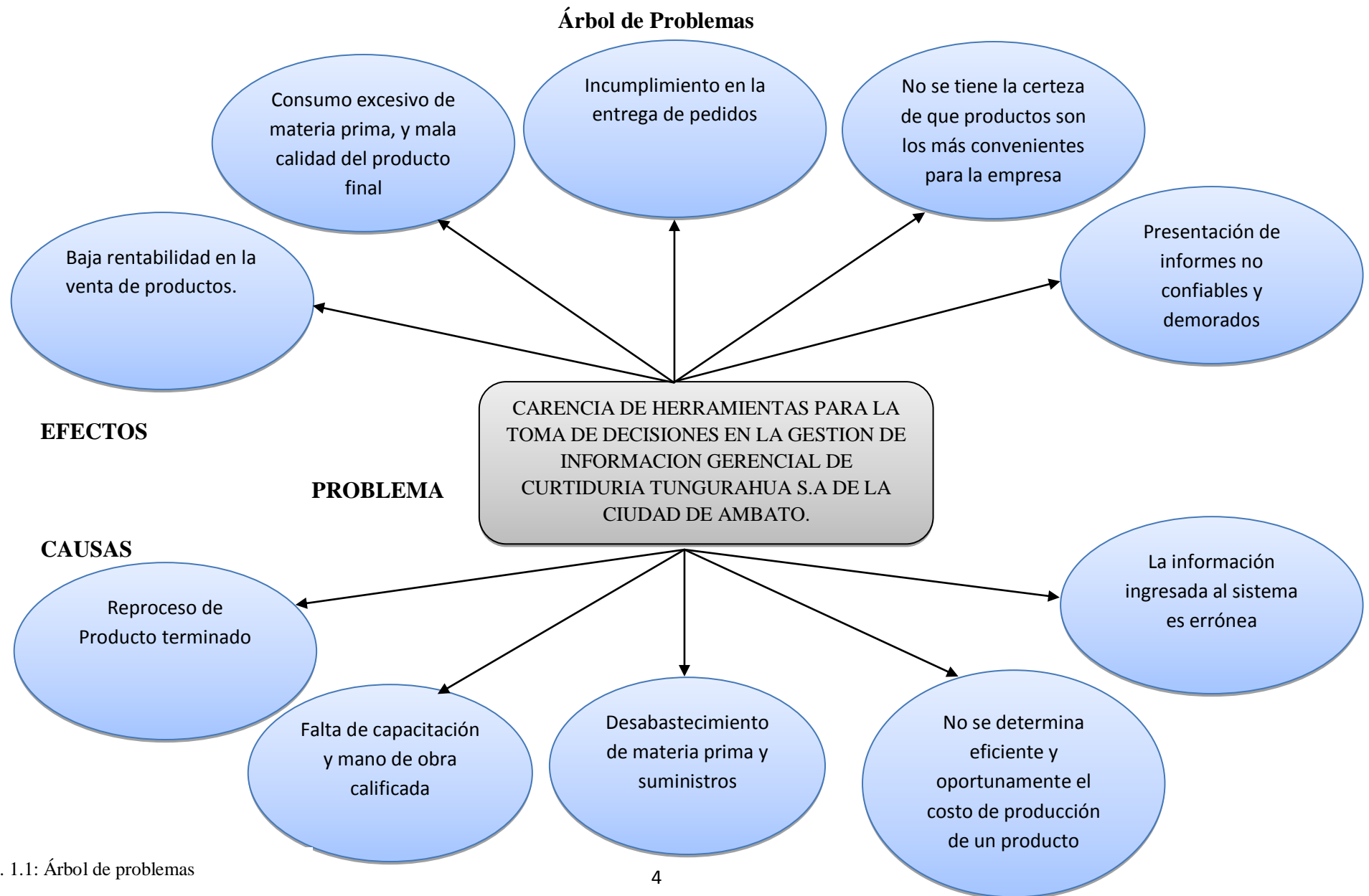


Fig. 1.1: Árbol de problemas

Elaborado por: Ing. Elena Sánchez

Análisis Crítico

Ante la incapacidad para apoyar la toma de decisiones gerenciales de los sistemas de información implantados actualmente en Curtiduría Tungurahua S.A conlleva a determinar una serie de consecuencias; el retraso en la venta y entrega de pedidos de productos a los clientes, el desabastecimiento de la materia prima y suministros necesarios para la producción obligando a los técnicos a cambiarlos temporalmente lo cual afecta los costos de elaboración del producto final, alterando los niveles de rentabilidad, además que se experimenta una pérdida de tiempo de producción.

Careciendo de las herramientas necesarias para almacenar y presentar de manera correcta la información de producción así como para obtener esta misma información de manera confiable y consolidada, se presenta la problemática de establecer el costo de producción así como su rentabilidad.

Por otro lado la personalización del producto para cada cliente obliga constantemente a cambiar el proceso de elaboración cambiando la programación de entregas de pedidos, obligando a realizar reprocesos consumiendo nuevamente materia prima.

Finalmente la información incorrecta ingresada al sistema ocasiona que los informes contengan información no confiable ni útil para la toma de decisiones de carácter estratégico en la empresa.

Prognosis

De mantenerse la carencia de un datawarehouse para la gestión de toma de decisiones en la administración de la producción en Curtiduría Tungurahua S.A, la empresa está en constante riesgo de que se generen pérdidas de recursos tanto en tiempo como en materia prima para la producción, provocando que la gestión administrativa no pueda desarrollarse de manera óptima y que las ventas se vean reducidas en su eficiencia.

Si se mantiene la problemática actual es decir la falta de información confiable para la toma de decisiones la empresa estará en riesgo de que su administración no sea la más apropiada, puesto que puntos como la rentabilidad de la empresa se ve directamente afectada pese al nivel de ventas que maneja Curtiduría Tungurahua S.A.

De no dar atención a la falta de seguimiento de los centros de producción y de no manejar correctamente los flujos de trabajo, en la fábrica se corre el riesgo de no contar a tiempo con los insumos, materia prima y suministros solicitados por cada centro de producción, retrasando así las planificaciones de producción y su respectiva toma de decisiones, y los problemas en los centros se agudizarán continuamente, lo que de alguna forma conllevan a la inestabilidad económica y administrativa de la fábrica.

Formulación del problema

¿Cómo incide un Datawarehouse para la toma de decisiones en la gestión de información gerencial de Curtiduría Tungurahua S.A de la ciudad de Ambato?

Interrogantes de la Investigación

- ¿Cuáles son los problemas que se generan en la toma de decisiones sobre la gestión de la información gerencial en Curtiduría Tungurahua S.A?
- ¿Cuáles son los pasos a seguir para estructurar un diseño de un Datawarehouse?
- ¿Desarrollo de un Datawarehouse la toma de decisiones sobre la gestión de la información gerencial en Curtiduría Tungurahua S.A es la alternativa más eficiente?

Delimitación del Objeto de la Investigación

Delimitación del contenido

Campo: Sistemas

Área: Gestión de Producción

Aspecto: Base de Datos

Delimitación Espacial

La presente investigación se desarrollara en los espacios físicos de Curtiduría Tungurahua S.A. de la ciudad de Ambato, ubicado en el parque Industrial Tercera Etapa, calle 8.

Delimitación Temporal

Esta investigación se enfocará en el período comprendido entre el 01 de enero y el 31 de diciembre del 2011.

Unidades de Observación

- Gerencia de Producción
- Jefatura de Planta
- Coordinadores de Áreas
- Jefatura de Ventas

Justificación

En cuanto a la **relevancia** del problema es necesario indicar que a nivel de empresas industriales y especialmente fabricas curtidoras, el proceso de producción es severamente complicado y cambiante, razón por lo cual se manejan procesos de producción de forma manual, y otras automatizadas pero que carecen del ligamiento adecuado para obtener la información rápida y oportuna.

Sin embargo, al formar parte de un mundo globalizado que tiende a la automatización de todos los procesos, y que la tecnología se encuentra en constante evolución, vemos con mucha complacencia como este tipo de fábricas cambian en su forma de pensar y de actual enfocando su esfuerzo para invertir en sistemas automatizados para la gestión de la producción.

La realización de este proyecto es **importante**, que pueden ayudar a las empresas a favorecer una estrategia de integración de la información, imprescindible para que se genere valor en la información resultante y se puedan apoyar decisiones claves que afecten al negocio con un mayor nivel de certidumbre, con la finalidad de superar los inconvenientes que presentan las bases de datos tradicionales que generalmente tienden a la fragmentación de la información.

Es **factible** realizar este proyecto, ya que se cuenta con la predisposición al cambio por parte de la junta de accionistas y gerencia, pero sobre todo con los medios necesarios para establecer un buen diseño de un datawarehouse con el propósito de mejorar la situación actual de la fábrica.

Se utilizara fuentes de información secundaria actualizada y especializada sobre el tema de investigación. El proyecto de investigación tendrá una **utilidad práctica** porque además se plantea una alternativa de solución al problema investigado.

Con la información oportuna y confiable, la gerencia especialmente de producción y departamento de ventas podrán de una manera real conocer su desempeño, con la finalidad de prevenir y corregir errores futuros en la planificación de la producción, advirtiendo un óptimo inventario de existencia de materia prima y suministros en las bodegas, y todo está a través de una **novedosa** solución.

Objetivos

Objetivo General

Determinar la incidencia de un datawarehouse para la toma de decisiones en la gestión de la información gerencial de Curtiduría Tungurahua S.A de la ciudad de Ambato.

Objetivos Específicos

- ✚ Identificar los problemas que se generan ante la falta de datawarehouse para la toma de decisiones sobre la información gerencial en Curtiduría Tungurahua S.A
- ✚ Cuantificar el estado actual de la información gerencial generada por los procesos de la empresa dentro del sistema transaccional.
- ✚ Desarrollar un datawarehouse como alternativa de solución al problema de la carencia de herramientas para la gestión de la información gerencial de Curtiduría Tungurahua S.A de la ciudad de Ambato.

CAPÍTULO II

MARCO TEORICO

Antecedentes Investigativos

En la Pontificia Universidad Católica existe una tesis cuyo tema es “Análisis, Diseño e Implementación de un Datawarehouse de soporte de decisiones para un Hospital del Sistema de Salud Público”, realizada por Villanueva Ojeda, Álvaro cuyo trabajo de tesis presenta una solución que los hospitales pueden implementar para satisfacer sus necesidades de gestión, análisis y toma de decisiones. Otorga un panorama de lo que está sucediendo en el hospital y presenta esta información en línea.

Concluye que tener todos los datos consistentes y ordenados en el Datawarehouse brinda una fuente confiable y estandarizada para el desarrollo de futuros Datamarts o para la ampliación del alcance de los existentes, facilitando el desarrollo de estos.

Otra tesis en la Universidad San Francisco cuyo tema es “Aplicación de un Datawarehouse orientado a la banca”, realizada por Carlos Oswaldo Medina Guerrero. Sus principales conclusiones expresan que las medidas del

éxito de un Datawarehouse son su uso y/o rendimiento de la inversión. Un Datawarehouse alcanza su valor más grande cuando se utiliza en exceso. Se dice que se conoce que se es acertado cuando los eufóricos usuarios exigen aún más tipos de datos y tener acceso a más detalle. Más allá de los usuarios que exigen más acceso y características, la empresa debe gozar de beneficios materiales, ej., las reducciones de costos operativos debido a cortes en los ciclos de informes, seguimiento al cliente.

Los beneficios materiales se pueden medir como un rendimiento de la inversión tradicional. Sin importar la medida usada, la terminación exitosa y el uso de un Datawarehouse depende más de la cultura (es decir, reactivo o proactivo) de la empresa que su profesionalismo técnico.

En cuanto a bibliografía especializada sobre el tema de investigación se menciona a: Building the Datawarehouse , del autor (Inmon, 1992); Fundamentals of Datawarehouse and Inteligencia de Negocios for Knowledge Management – Instructor Guide, de (IBM, 2002) y también tenemos Business Intelligence: Técnicas de Análisis para la Toma de Decisiones Importantes de los autores (Vitt, Luckevich, & Misner, 2003)

Filosófica

Para realizar el trabajo de grado el investigador asume los principios del paradigma crítico-propositivo porque hace un análisis crítico del problema determinando los requerimientos y a su vez presenta una propuesta de solución al problema investigado basado en la existencia de múltiples realidades socialmente construidas.

Este enfoque privilegia la interpretación, comprensión y explicación del estado actual de la problemática de manera global. La investigación está presentada con enfoques críticos de las principales coordinaciones de la empresa obtenidos a través de encuestas, es por tal motivo que propicia la participación de

la gerencia y jefaturas o coordinaciones en calidad de protagonistas

Tecnológica

Los avances tecnológicos subsanan el continuo desencuentro que hasta ha existido entre la tecnología y las reglas de un negocio. Ciertos avances tecnológicos permiten hoy obtener rendimientos óptimos con el uso de hardware, sistemas de información para la dirección, procesamiento analítico on-line (OLAP), MDA (Análisis Multidimensional), Datamart (DM) y Datawarehouse (DW).

En este trabajo se ha intentado destacar, a través de un caso concreto, que aplicar un Datawarehouse herramienta tecnológica que permite mejorar la gestión de la información almacenada en las bases de datos empresariales puede ser un camino adecuado para conseguir mejoras de desempeño empresarial.

Administrativa

El inicio de ésta empresa data del año 1938, en ese entonces se preparaba cueros barnizados a mano como especialidad de la curtiembre que poco a poco fue adquiriendo notoriedad por este cuero.

En los años 70 y 80 la Industria del Calzado está en auge, Curtiduría Tungurahua logro satisfacer las necesidades de la mayoría de estos fabricantes y camina a su ritmo durante muchos años, logrando un crecimiento sostenido que se fue notando con el pasar de los años.

En los últimos 10 años la empresa ha logrado renovar su infraestructura acorde a las necesidades y recursos que le han permitido liderar en la actividad de

curtidos por su diversidad y apoyo directo al productor del calzado con asistencia técnica especializada, y procurando acompañar tendencias, colores y productos.

En la actualidad ha logrado consolidar un grupo de clientes importantes de las más prestigiadas industrias, empresas, pequeñas fábricas y artesanos de calzado del país

Su estructura organizacional se divide en 4 niveles de gestión: Directivo, Ejecutivo; Coordinación, Operativo. Dentro del nivel directivo encontramos al Junta de accionistas y Presidencia; el nivel ejecutivo lo conforma la gerencia general; el de coordinación Gerencia de Producción, Jefaturas de Ventas, el Departamento Financiero, Departamento de Sistemas, RRHH; dentro de los Niveles de Operativos a todo el personal obrero.

La unidad de sistemas informáticos está bajo la dependencia de la gerencia, fungiendo las actividades de unidad técnica sobre la tecnología que abarca esta rama.

La empresa Curtiduría Tungurahua S.A está ubicada en el Parque industrial, tercera etapa, calle 8.

Legal

Para realizar el presente trabajo de grado se fundamenta en Art. 385 de la constitución de la República del Ecuador Sección Octava, (Ciencia, tecnología, innovación y saberes ancestrales) inciso 3, que nos indica que:

Dentro un marco de respeto al ambiente, la naturaleza, la vida, las culturas y la soberanía se puede desarrollar tecnologías e innovaciones que impulsen la

producción nacional, eleven la eficiencia y productividad, mejoren la calidad de vida y contribuyan a la realización del buen vivir.

Fuente: Constitución de la República de Ecuador (Contraloría General del Estado, 2010)

Ver Anexo 1.

Categorías Fundamentales

Red de Inclusiones Conceptuales

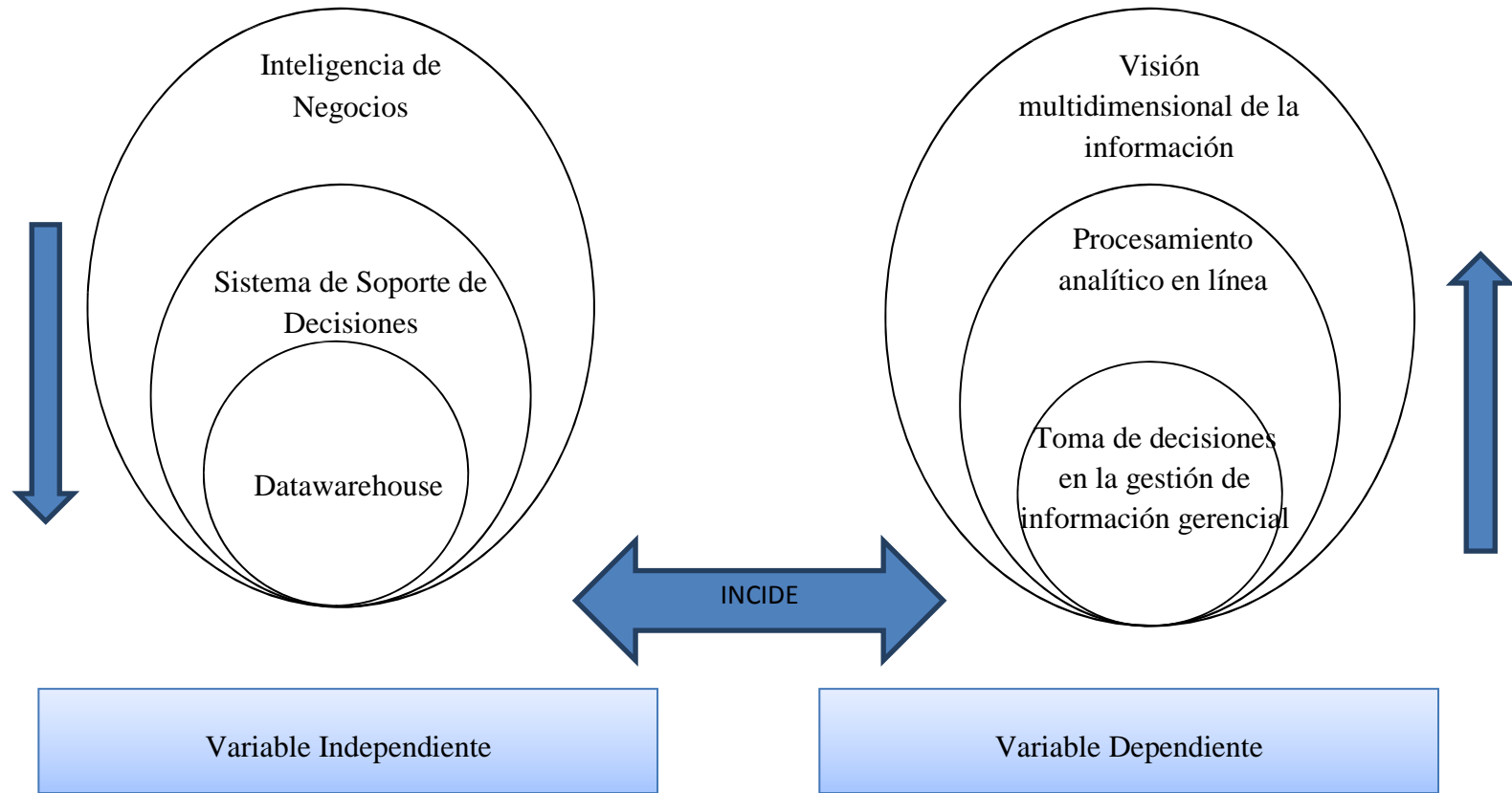


Fig. 2.1: Categorías Fundamentales

Elaborado por: Ing. Elena Sánchez

Constelación de Ideas de la Variable Independiente

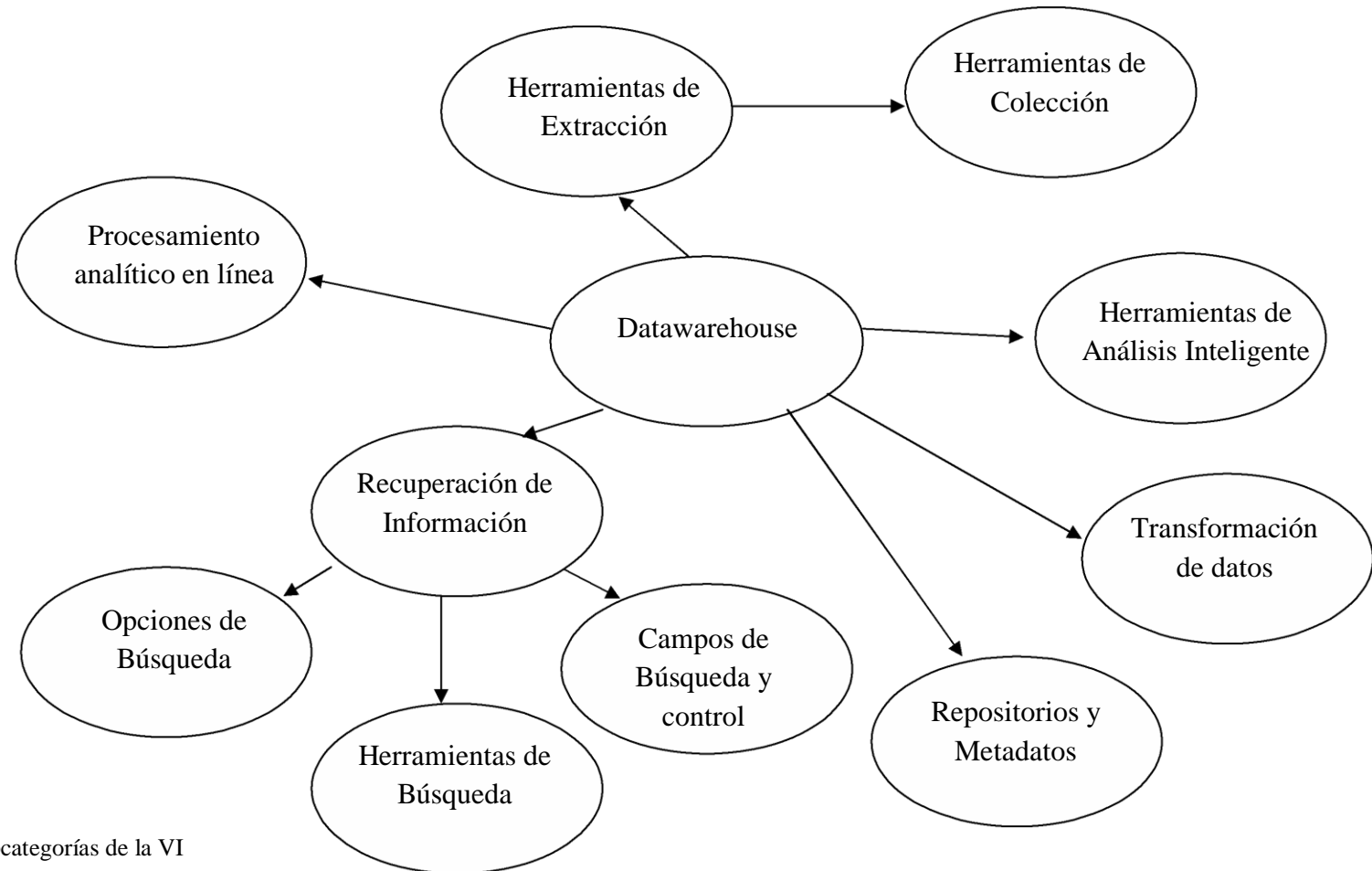


Fig. 2.2: Subcategorías de la VI

Elaborado por: Ing. Elena Sánchez

Constelación de Ideas de la Variable Dependiente

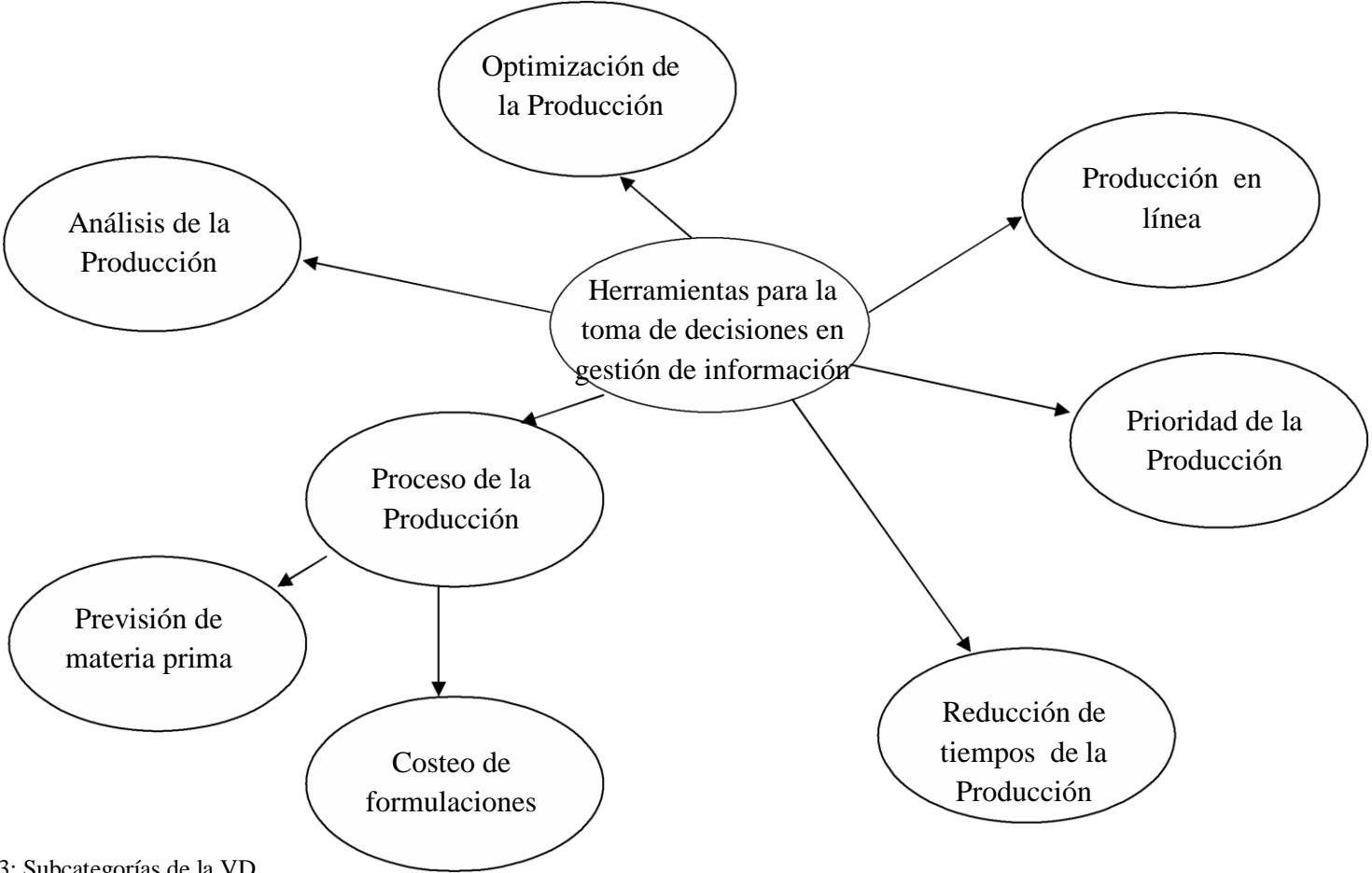


Fig. 2.3: Subcategorías de la VD

Elaborado por: Ing. Elena Sánchez

Categorías Fundamentales de la Variable Independiente




Inteligencia de Negocios

Inteligencia de negocios 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. (IT News, 2011).

Desde un punto de vista más pragmático, y asociándolo directamente con las tecnologías de la información, podemos definir Inteligencia de Negocios 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. (Business Intelligence, 2009).

Los principales productos de Inteligencia de Negocios que existen hoy en día son:

-  Cuadros de Mando Integrales (CMI)
-  Sistemas de Soporte a la Decisión (DSS)
-  Sistemas de Información Ejecutiva (EIS)

Por otro lado, los principales componentes de orígenes de datos en la Inteligencia de Negocios que existen en la actualidad son:

- ✚ Datamart
- ✚ Datawarehouse

Los sistemas y componentes del BI se diferencian de los sistemas operacionales en que están optimizados para preguntar y divulgar sobre datos. Esto significa típicamente que, en un datawarehouse, los datos están desnormalizados para apoyar consultas de alto rendimiento, mientras que en los sistemas operacionales suelen encontrarse normalizados para apoyar operaciones continuas de inserción, modificación y borrado de datos. En este sentido, los procesos ETL (extracción, transformación y carga), que nutren los sistemas BI, tienen que traducir de uno o varios sistemas operacionales normalizados e independientes a un único sistema desnormalizado, cuyos datos estén completamente integrados.

En definitiva, una solución BI completa permite:

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

Herramientas que ofrecen estas soluciones. (True Business,2010)

Oracle

Herramienta basada en la arquitectura Cliente/Servidor para la gestión de Bases de Datos Relacionales desarrollada por Oracle Corporation. Ofrece una interfaz intuitiva basada en el explorador, que es capaz de administrar las bases de

datos, crear tablas, vistas y otros objetos de bases de datos, importar, exportar y visualizar datos de tablas, ejecutar scripts de SQL y generar informes. Además, soporta transacciones, es estable, escalable y multiplataforma.

Cognos

Herramienta de Inteligencia de Negocios (BI) y Administración del Desempeño Financiero (FPM) es utilizado por más de 23.000 empresas y organizaciones líderes en todo el mundo. El software Cognos y los servicios Cognos de IBM pueden ayudarle a entregar su información en demanda.










Pentaho

Herramienta de Inteligencia de Negocios desarrollada bajo la filosofía del software libre para la gestión y toma de decisiones empresariales. Es una plataforma compuesta de diferentes programas que satisfacen los requisitos de BI. Ofreciendo soluciones para la gestión y análisis de la información, incluyendo el análisis multidimensional OLAP, presentación de informes, minería de datos y creación de cuadros de mando para el usuario.

JasperSoft

Conjunto de herramientas muy poderosas que permiten a las organizaciones generar información basada en sus datos de administración para la evaluación y toma diaria de decisiones, en forma dinámica y on-line. La plataforma de trabajo de JasperSoft permite integrar fácilmente las diversas fuentes de datos disponibles en la empresa, y por medio de técnicas de análisis multidimensional obtener indicadores que se presentan en tableros de control y reportes dinámicos, proveyendo de información vital a la alta gerencia.

Otras herramientas:

-  Dynamics
-  Ultimatus
-  Business Objects
-  Cubeware Sagent Solution Plattform
-  Microstrategy
-  Cognos
-  Applix TM1
-  Oracle 9I Application Server
-  Suite Bitman/Artus Bussiness Intelligence

Inteligencia de Negocios como solución tecnológica

Permite centralizar, depurar y afianzar los datos. Las tecnologías de BI permiten reunir, normalizar y centralizar toda la información de la empresa, mediante un almacén de datos, permitiendo así su explotación sin esfuerzo. De esta forma, los departamentos comercial, operativo y financiero basan las decisiones estratégicas en la misma información.

Descubrir información no evidente para las aplicaciones actuales. En el día a día de las aplicaciones de gestión se pueden esconder pautas de comportamiento, tendencias, evoluciones del mercado, cambios en el consumo o en la producción, que resulta prácticamente imposible reconocer sin el software adecuado. Es lo que se puede calificar como extraer información de los datos, y conocimiento de la información.

Optimizar el rendimiento de los sistemas. Las plataformas de BI se diseñan para perfeccionar al máximo las consultas de alto nivel, realizando las transformaciones oportunas a cada sistema (OLTP - OLAP), y liberando los servidores operacionales

Un Sistema de Soporte a la Decisión (DSS¹) es una herramienta de Inteligencia de Negocios enfocada al análisis de los datos de una organización.

En principio, puede parecer que el análisis de datos es un proceso sencillo, y fácil de conseguir mediante una aplicación hecha a medida o un ERP sofisticado. Sin embargo, no es así: estas aplicaciones suelen disponer de una serie de informes predefinidos en los que presentan la información de manera estática, pero no permiten profundizar en los datos, navegar entre ellos, manejarlos desde distintas perspectivas... etc.

(Business Intelligence, 2009)

Diferencia con otras herramientas de Inteligencia de Negocios

El principal objetivo de los Sistemas de Soporte a Decisiones es, a diferencia de otras herramientas como los Cuadros de Mando (CMI)² o los Sistemas de Información Ejecutiva (EIS)³, explotar al máximo la información residente en una base de datos corporativa (datawarehouse o datamart), mostrando informes muy dinámicos y con gran potencial de navegación, pero siempre con una interfaz gráfica amigable, vistosa y sencilla.

Otra diferencia fundamental radica en los usuarios a los que están destinadas las plataformas DSS: cualquier nivel gerencial dentro de una organización, tanto para situaciones estructuradas como no estructuradas. (En este sentido, por ejemplo, los CMI están más orientados a la alta dirección).

Por último, destacar que los DSS suelen requerir (aunque no es imprescindible) un motor OLAP⁴ subyacente, que facilite el análisis casi ilimitado de los datos para hallar las causas raíces de los problemas/pormenores de la compañía.

¹ DSS, Sistema de Soporte a la Decisión – Decision support system

² CMI, Cuadros de mando integrales

³ EIS, Executive Information System

⁴ OLAP, On-Line Analytical Processing

Tipos de Sistemas de Soporte a Decisiones

Sistemas de información gerencial (MIS)⁵

Los sistemas de información gerencial, también llamados Sistemas de Información Administrativa dan soporte a un espectro más amplio de tareas organizacionales, encontrándose a medio camino entre un DSS tradicional y una aplicación CRM/ERP implantada en la misma compañía.

Sistemas de información ejecutiva (EIS)

Los sistemas de información ejecutiva son el tipo de DSS que más se suele emplear en Inteligencia de Negocios, ya que proveen a los gerentes de un acceso sencillo a información interna y externa de su compañía, y que es relevante para sus factores clave de éxito.

Sistemas expertos basados en inteligencia artificial

Los sistemas expertos, también llamados sistemas basados en conocimiento, utilizan redes neuronales para simular el conocimiento de un experto y utilizarlo de forma efectiva para resolver un problema concreto. Este concepto está muy relacionado con la minería de datos.

⁵ MIS, Management Information Systems

Sistemas de apoyo a decisiones de grupo (GDSS)⁶

Un sistema de apoyo a decisiones en grupos es "un sistema basado en computadoras que apoya a grupos de personas que tienen una tarea (u objetivo) común, y que sirve como interfaz con un entorno compartido". El supuesto en que se basa el GDSS es que si se mejoran las comunicaciones se pueden mejorar las decisiones.

Sistema de Soporte de Decisiones (DSS)

Según (IBM, 2002), un Sistema de Soporte de Decisiones contiene todos los servicios y procesos, para seleccionar, manipular, y analizar información y presentar resultados. Debe de permitir acceso transparente a la data en varias partes del Datawarehouse y proveer una interfaz común para los diferentes grupos de usuarios. Un DSS también puede ser definido como un sistema computacional diseñado para apoyar en los procesos de la toma de decisiones en una organización. Un DSS es la ventana del usuario a los datos almacenados en el ambiente del Datawarehouse.

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

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

No requiere conocimientos técnicos. Un usuario no técnico puede crear nuevos gráficos e informes y navegar entre ellos, haciendo drag&drop o drill

⁶ GDSS, Group Decision Support Systems

through. Por tanto, para examinar la información disponible o crear nuevas métricas no es imprescindible buscar auxilio en el departamento de informática.

Rapidez en el tiempo de respuesta, ya que la base de datos subyacente suele ser un datawarehouse corporativo o un datamart, con modelos de datos en estrella o copo de nieve. Este tipo de bases de datos están optimizadas para el análisis de grandes volúmenes de información.

Integración entre todos los sistemas/departamentos de la compañía. El proceso de ETL⁷ previo a la implantación de un Sistema de Soporte a la Decisión garantiza la calidad y la integración de los datos entre las diferentes unidades de la empresa. Existe lo que se llama: integridad referencial absoluta.

Cada usuario dispone de información adecuada a su perfil. No se trata de que todo el mundo tenga acceso a toda la información, sino de que tenga acceso a la información que necesita para que su trabajo sea lo más eficiente posible.

Disponibilidad de información histórica. En estos sistemas está a la orden del día comparar los datos actuales con información de otros períodos históricos de la compañía, con el fin de analizar tendencias, fijar la evolución de parámetros de negocio... etc.



Fig. 2.4: Sistema DSS

Fuente: (IBM, 2002)

⁷ ETL, Extract, transform and load (extracción, transformación y carga)

Datawarehouse

Conceptos Fundamentales

Para empezar vamos a enunciar una serie de conceptos básicos que nos permitirán entender toda la teoría que seguirá en los restantes apartados de este documento:

Datawarehouse: Es la integración de datos consolidados, almacenados en un dispositivo de memoria no volátil, proveniente de múltiples y posiblemente diferentes fuentes de datos. Con el propósito del análisis y a partir de este tomar decisiones en función de mejorar la gestión del negocio. Contiene un conjunto de cubos de datos que permiten a través de técnicas de OLAP consolidar, ver y resumir los datos acorde a diferentes dimensiones de estos. (Chaudhuri & Dayal, 1997)

Datamarts: Es un subconjunto del Datawarehouse, usado normalmente para el análisis parcial de los datos. Ej.: El Datamart de los datos del departamento ventas y el Datamart de Inventarios. El objetivo de subdividir está dado por la complejidad computacional del análisis global de todas las dimensiones del Datawarehouse y por la necesidad de rapidez. (Microsoft SQL Server Data Warehousing Training, 2000)

Data Mining: Es el descubrimiento de conocimiento oculto en las bases de datos. Relaciones entre estos y tendencias que permiten una toma de decisiones acertada. Incluye Asociación, Caracterización, Clasificación, Análisis de Series Cronológicas, etc. (Chaudhuri & Dayal, 1997)

OLTP (Online Transaction Processing): Se les llama así a las aplicaciones orientadas principalmente a la inserción, actualización y eliminación de datos, diseñada casi siempre usando el modelo Relacional. Estos sistemas están optimizados para realizar estas operaciones en un tiempo corto. (Microsoft Books Online, 2000)

OLAP (Online Analytical Processing): Son los sistemas que se usan para analizar los datos que las OLTP introducen en la Base de Datos. A diferencia de los primeros

estos casi siempre usan el modelo multidimensional para organizar los datos en la Base de Datos ya que brindan mejores resultados a la hora del análisis de estos. (Microsoft Books Online, 2000)

El Modelo de Hechos Dimensionales

El análisis de los datos en un tiempo finito, ha traído consigo estudios sobre la mejor forma de almacenar y representar estos datos para que puedan ser consultados de una forma más rápida. El uso del Modelo Multidimensional es una de las aproximaciones más acertadas y seguidas por los especialistas en estos días. Este se basa en el estudio de los eventos del negocio analizados desde sus distintas dimensiones. Así:

Definición 1: Llamamos evento o Hecho a una operación que se realiza en el negocio en un tiempo determinado. Son objeto de análisis para la toma de decisiones. Se Representan en una caja con su nombre y las medidas que lo caracterizan. (Wrembel & Concilia, 2007)

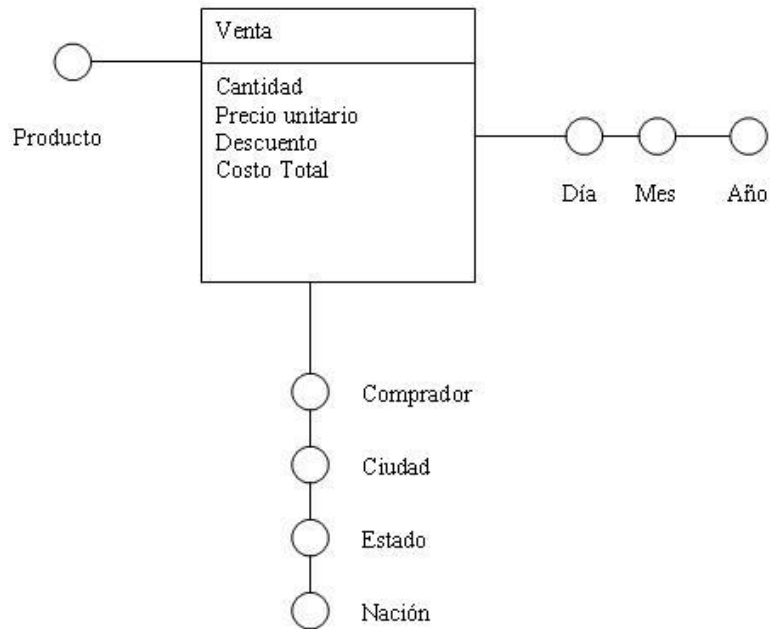


Fig. 2.5: Representación gráfica de un hecho y sus dimensiones

Fuente: (Wrembel & Concilia, 2007)

Los Hechos están estrechamente relacionados con el tiempo. Los eventos que son estáticos no tiene objetivo de análisis para este modelo, aunque son muy pocos los hechos que no ocurren con determinada periodicidad en un negocio. Los hechos están caracterizados por medidas numéricas como se muestra en el ejemplo de la figura 1: la cantidad, el precio unitario, el descuento, etc., son las medidas del Hecho (VENTA).

Nota: Fíjese que el producto que se vende, su costo y la fecha de la venta no son características de esta como lo podrían ser en cualquier diseño relacional. En este caso, esos serían dimensiones de ese Hecho, por las que, puede ser analizado más adelante.

Definición 2: Una Medida es una propiedad de un Hecho (casi siempre numérica), que es usada para su análisis. (Wrembel & Concilia, 2007)

Nota: Un hecho puede no poseer ninguna medida. En ese caso se dice que el Hecho es vacío y solo se usa para contar la aparición de este en el tiempo.

Definición3: Una Dimensión es una característica de un hecho que permite su análisis posterior, en el proceso de toma de decisiones (Wrembel & Concilia, 2007)

Nota: Un hecho debe estar relacionado al menos con una dimensión: "El tiempo".

Es un interés del negocio tomar decisiones sobre los hechos que ocurren en este, pero para esto se necesita su análisis. Por ejemplo Las ventas en la semana antes del 14 de Febrero, puede ser un objeto de análisis para un negocio comercial. Para esto se necesita tener el Hecho Ventas analizado en la dimensión Tiempo. En este caso en los Días:

$7 \leq d \leq 14$. Si se quisiera saber que productos fueron los más vendidos en esos días entonces tendríamos que adicionar una nueva dimensión de análisis, Producto. Así adicionando dimensiones a nuestro estudio se pudieran llegar a conclusiones sobre si

el siguiente año en esa época debería comprarse más objetos de un producto o menos de otro. Elemento este muy importante para la futura estrategia de la empresa.

Definición 4: En una empresa pueden existir varios hechos que sean analizados por dimensiones iguales. En este caso se les llama a estas dimensiones: Dimensiones Compartidas. (Wrembel & Concilia, 2007). Un ejemplo de esto es El Hecho Ventas puede ser analizado en las dimensiones Tiempo y Producto. Lo mismo ocurre con el Hecho Compras.

Las dimensiones deben ser atómicas y las relaciones entre estas crean jerarquías que permiten un análisis jerárquico de los datos. Un ejemplo de esto es el Tiempo, que es dividido en tres dimensiones. Día, Mes y Año. Cada uno es una dimensión distinta, pero relacionadas jerárquicamente en una relación de 1 a muchos, que permite el análisis del Hecho, por días, meses o años, o la combinación de ellos. Esto da al traste con las acostumbradas (OLTP)⁸ que manejan el Tiempo como una propiedad de una entidad, y lo tratan como un todo. Por lo que, como podemos inferir de aquí: en muchos casos hará falta convertir las bases de datos de estos sistemas a la nueva filosofía. SQL Server tiene facilidades para esto llamadas DTS (Data Transformation Services) que permite leer datos desde cualquier SGBD⁹ que posea un driver ODBC¹⁰ o implemente la nueva tecnología OLE DB¹¹ de Microsoft.

Diagrama en Estrella

Uno de los tipos de consultas más usadas en las OLAP es la llamada Estrella. Su nombre lo adquiere debido a que su implementación en un ambiente relacional (MOLAP Multidimensional Online Analytical Processing) está dado por varias tablas que almacenan las jerarquías dimensionales y una tabla que contiene el hecho con una relación 1: m con estas tablas de dimensiones. Veamos un ejemplo gráfico:

⁸ OLTP, Procesamiento de Transacciones En Línea (OnLine Transaction Processing)

⁹ SGBD, Sistema de gestión de bases de datos (Database Management System)

¹⁰ ODBC, *Conectividad para abrir una base de datos* (Open Database Connectivity)

¹¹ OLE DB, Enlace e incrustación de objetos para base de datos (Object Linking and Embedding for Databases)

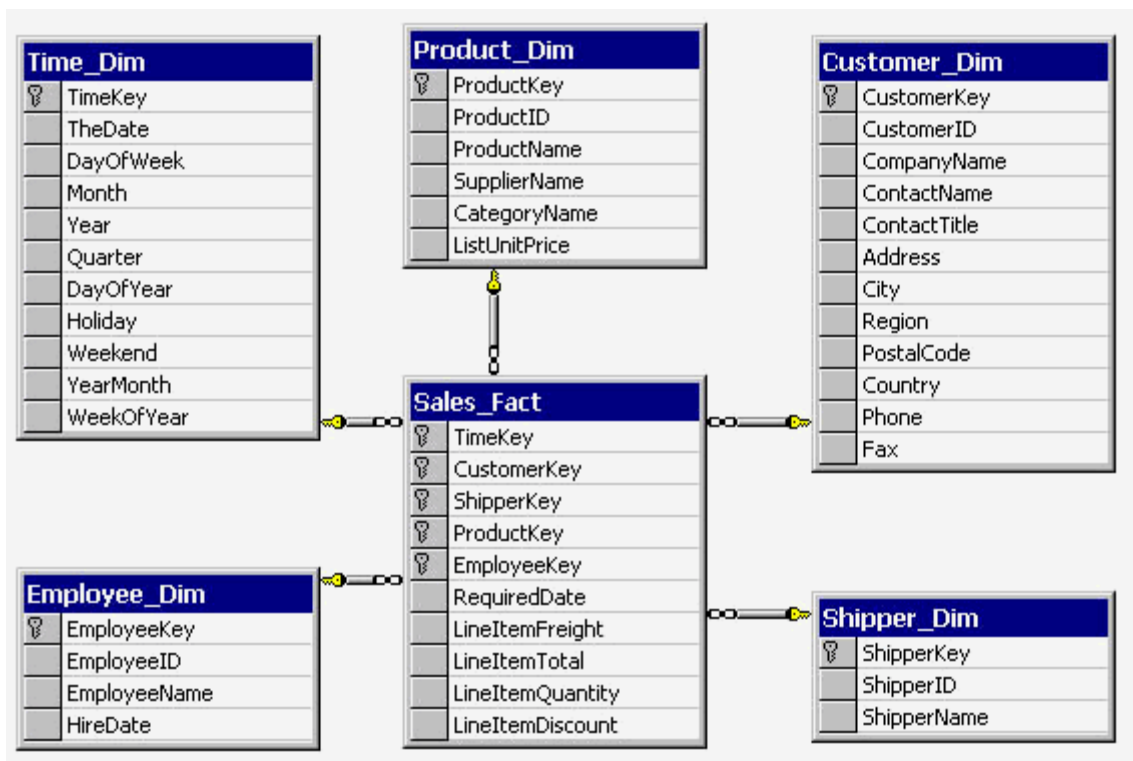


Fig. 2.6: Diagrama estrella del hecho

Fuente: (Wrembel & Concilia, 2007)

Para cerrar con broche dorado este tema es necesario hacer alusión a los llamados Cubos de datos: Estos no son más que el conjunto formado por todas las tablas Dimensión y la tabla Hecho que al final dan una vista en forma de Cubo cuyas celdas están compuestas por las medidas de la tabla Hecho. Esta es la base de las aplicaciones OLAP. El cubo de datos es lo que hace que los reportes sean obtenidos con un bajo tiempo de respuesta y que el análisis de los datos pueda ser tan diverso, pues cada cara del cubo se refiere a un análisis distinto de las medidas almacenadas. Veamos el ejemplo gráfico del cubo:

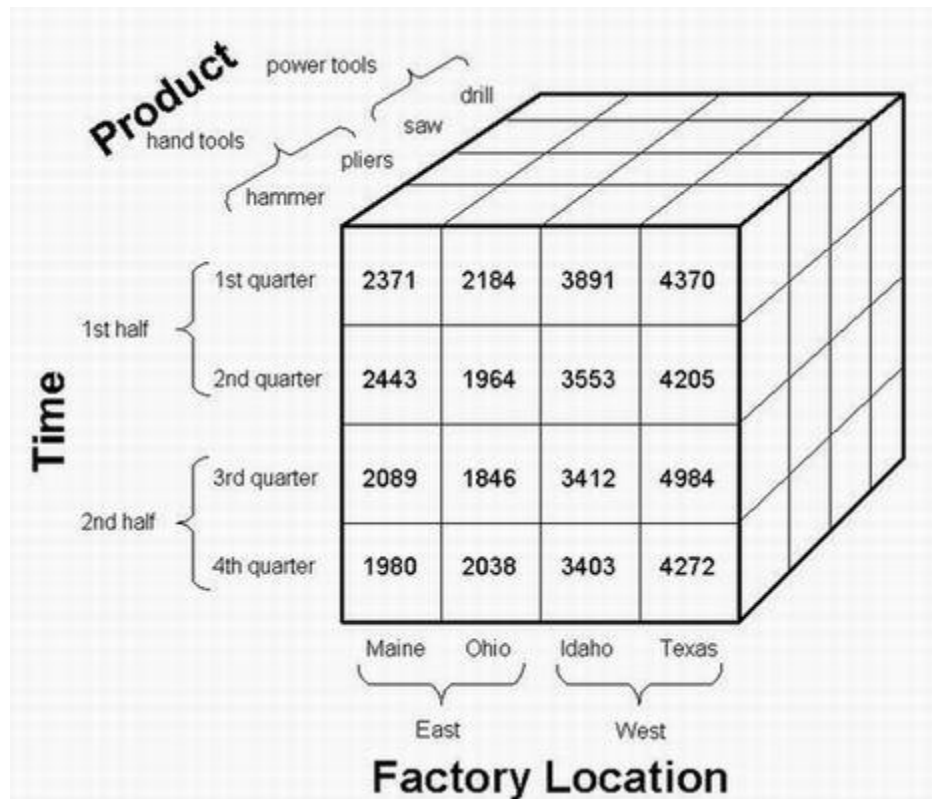


Fig. 2.7: Cubo de datos

Fuente: (Wrembel & Concilia, 2007)

Como podemos ver en el ejemplo la cantidad de producción puede ser analizada por producto, teniendo en cuenta la dimensión Producto, Por Tiempo, por Localización de las Industrias o en su conjunto por todas ellas a la vez o cualquier combinación de estas. Esto le da al analista o al sistema experto una amplia gama de posibilidades de las cuales puede tomar ventaja. En nuestro caso de estudio de las ventas. El cubo de datos formado por la Tabla Sales_Fac en conjunto con las restantes tablas de Dimensión nos permite analizar las ventas por Empleado, por Consumidor, por Tiempo, Etc.

Extracción y Transformación de los Datos

SQL puede conectarse a cualquiera de los formatos creados por Microsoft Office, archivos texto y a bases de datos como FOXPRO y ACCESS muy usadas por múltiples sistemas de escritorio que resuelven problemas importantes en muchas empresas. Además puede conectarse a cualquier SGBD que posea un driver ODBC compatible con Microsoft o implemente la tecnología OLE DB. El poder obtener datos de páginas de Excel con datos resumidos o exportarlas a este, para que los usuarios más avanzados puedan crear gráficos y otros. Es algo que convierte a SQL en la herramienta perfecta para crear sistemas Datawarehouse. Para realizar estas tareas Microsoft cuenta con una rama dentro del árbol de administración del SQL Enterprise Manager, que contiene un área de diseño con los elementos que se necesitan para diseñar un paquete de transformación de datos (DTS Package). (Datawarehouse Extracción y Transformación, 2008)

Para crear un paquete solo se hace clic derecho sobre el área vacía de esa rama y se selecciona la opción New Package, inmediatamente se abre una ventana con un área de diseño que permite empezar a diseñar el paquete en cuestión. Hay tres elementos fundamentales a la hora de diseñar un paquete:

Las conexiones: que permiten conectarse a cualquier fuente de datos como las que relacionamos anteriormente.

Las tareas: que permiten transformar los datos de cada conexión antes de que sean copiados a otra conexión. Estas tareas pueden ejecutar Scripts de ActiveX, pueden transferir otros archivos desde o hacia un sitio FTP, pueden enviar un mensaje a un operador de SQL. Además de poder filtrar datos, crear nuevas tablas en la BD resultante, etc.

Los flujos de trabajo: que permiten definir hacia donde irán los datos luego de que se le apliquen las respectivas acciones que los transformarán a la nueva forma deseada.

Así usando estos tres elementos fundamentales se crea una especie de script gráfico que sigue una secuencia lógica para dar como resultado la transformación de los datos almacenados de un formato a otro como SQL Server desde donde podrán ser analizados con el Datawarehouse.

Aunque desde el Enterprise Manager de SQL Server se puede ejecutar los paquetes que se crean haciendo clic derecho sobre ellos y seleccionando Execute o en un Schedule. Existen varias herramientas adicionales que pueden usarse desde la ventana de comandos del sistema operativo. De ellas, la más interesante para el uso de scripts es dtsrun. Para obtener la ayuda completa de esta podemos ejecutar dtsrun /? En una pc (Computadora Personal) con SQL Server instalado.

Ahora podrá ver el nombre del paquete y si desea ejecutarlo solo debe hacer clic derecho sobre este y seleccionar Execute. Así las cuentas y sus saldos serán insertados en la nueva base de datos de contabilidad de la empresa.

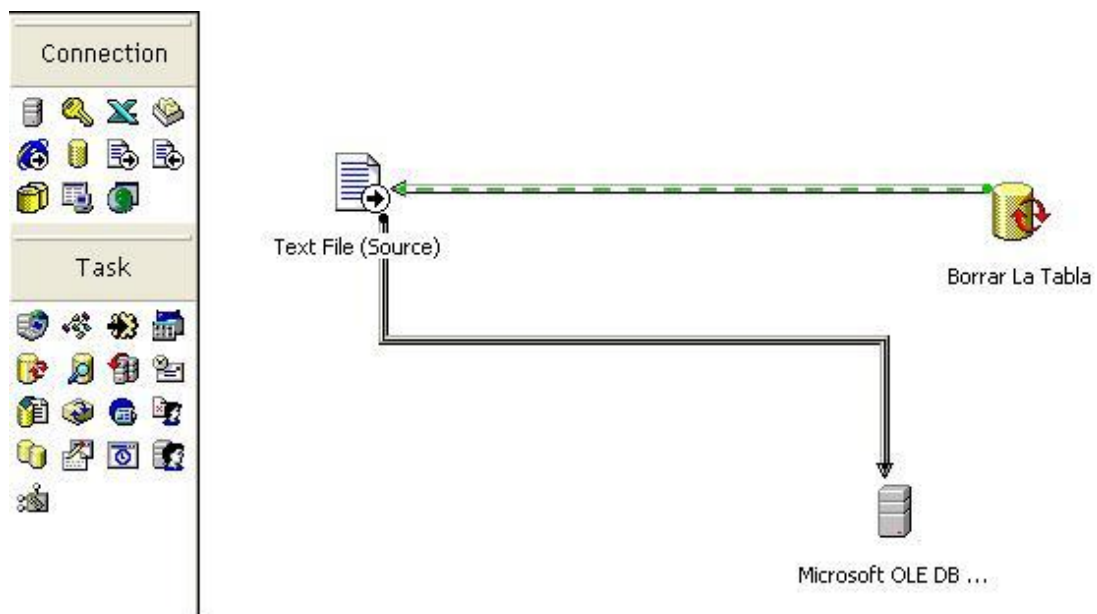


Fig. 2.8: Vista gráfica, del paquete creado

Fuente: (Datawarehouse Extracción y Trasformación, 2008)

Categorías Fundamentales Variable Dependiente

TOMA DE DECISIONES

En todos los aspectos de la vida nos tenemos que enfrentar todos los días a una toma de decisiones ya sean grandes o pequeños problemas que tengamos que solucionar. La toma de decisiones se da muy significativo dentro de la vida empresarial y por este factor se distinguen a las personas sobresalientes de las que no lo son tanto. (Vaquiro, 2012)

Para que se dé una toma de decisiones en los sistemas es necesario tener unos objetivos como son:

Tomar decisiones acertadas y basarse en la objetividad de los datos más que en los deseos y esperanzas para darles una interpretación adecuada.

El sistema de gestión de calidad mejora la información que se obtiene y mejora los orígenes para su obtención, se pueden hacer estudios y análisis de futuro, mejorando la información a corto plazo.

Uno de los problemas que se presentan en la toma de una decisión es por parte de los analistas de los sistemas ya que los resultados deben ser en hechos reales. Los analistas deberán intentar mejorar los resultados para conseguir las metas y los objetivos marcados en el plan de la organización, todo se da con el intercambio de información ya sea negativa o positivamente debe fluir por la organización

Los analistas son los encargados de señalar los defectos y proponer una pronta solución sin perjudicar la organización. (Vaquiro, 2012)

La información es la materia prima y fundamental en la toma de decisiones de los analistas en una organización ya que a mayor calidad de información, mejor es la calidad en la toma de decisiones

Se pueden seguir criterios analíticos exactos y si tiene una información perfecta ya que la información, vale tanto como el beneficio, o ausencia de pérdidas que se obtengan en base a esa información.

Todas las decisiones que se tomen tienen unos beneficios claves primordiales mediante los hechos.

El analista debe tener conocimiento de la organización y sobre todo tener clave la opinión del consumidor, estableciendo comunicación para que la información pueda ser recogida, analizada, resumida y estructurada.

Aumentar la habilidad para demostrar la efectividad de decisiones tomadas en los antecedentes de los registros basados en los hechos, para tomar decisiones acertadas basándose en el análisis.

Los analistas de los sistemas deberán aumentar la calidad de las decisiones, para que puedan identificar a tiempo las desviaciones en los objetivos a los cuales se quiere llegar.

La toma de decisiones mediante los hechos nos conducen a aplicar el principio de aproximación que se basa en los datos e información asegurando de que estos sean precisos y confiables, ya que con unos datos precisos, es posible aplicar unos métodos de evaluación adecuados para tomar decisiones adecuadas.

Todo esto hace que los datos sean accesibles para cualquiera que los necesite, ahorrándose así tiempo en resolver los problemas y ahonda en el conocimiento de las necesidades de los procesos de la organización, permitiendo lograr más fácilmente los objetivos de la organización.

El análisis de los datos y la información utilizando métodos válidos

Ha de emplearse una buena metodología y procedimiento para analizar la información, y toma de decisiones. (Toma de Decisiones, 2010)

Tomar decisiones y realizar acciones basadas en el análisis de los hechos, equilibradas con la experiencia y la intuición.

Es primordial a la hora de tomar decisiones analizar los hechos fríamente, apoyándose en experiencias pasadas para realizar predicciones y tomar las decisiones más acertadas. Las decisiones, han de ser tomadas de modo que los datos y experiencias pasadas, no han de entrar en contradicción con las futuras decisiones.

La toma de decisiones estratégicas se presenta a nivel corporativo, a nivel de unidad de negocio y a nivel funcional. Al término de esta regresión se encuentran los planes operativos para poner en práctica la estrategia. Esto hace que la administración estratégica sea más especializada.

La buena toma de decisiones permite vivir mejor. Nos otorga algo de control sobre nuestras vidas. De hecho, muchas de las frustraciones que sufrimos con nosotros mismos se deben a no poder usar la propia mente para entender el problema de decisión, y el coraje para actuar en consecuencia. Una mala decisión puede obligarnos a tomar otra mala decisión, como dijo Harry Truman: "Toda mala decisión que tomo va seguida de otra mala decisión".

Un buen analista de sistemas debe tomar muchas decisiones todos los días. Algunas de ellas son decisiones de rutina o intrascendentes mientras que otras tienen una repercusión drástica en las operaciones de la organización a la cuál investigan o trabajan. Algunas de estas decisiones podrían involucrar la ganancia o pérdida de los objetivos cumplimiento o incumplimiento de la misión y las metas de la organización.

Un modelo de decisiones más simple que tiene solo dos alternativas se denomina Maniqueísmo.

El Maniqueísmo es el concepto de dualidad que divide todo lo que forma parte del universo en dos alternativas distintas o dos polos opuestos, como por

ejemplo el bien y el mal, blanco y negro, día y noche, mente (o alma) y cuerpo, etc. Este concepto de dualidad fue un modelo suficiente de la realidad para aquella época para que el mundo fuera manejable y calculable. Sin embargo, hoy en día sabemos con certeza que todos los sistemas cambian y todos tienen un amplio espectro continuo. No existen los opuestos en la naturaleza. Debemos ver el mundo a través de los ojos de nuestra mente vivida; de lo contrario, no comprendemos bien las ideas complejas. (Toma de Decisiones, 2010)

La toma de decisiones y la experiencia es un elemento clave puesto que las decisiones deben tomarse sobre una realidad altamente compleja debido al enorme número de variables que entran en juego.

Las cinco características más importantes en la toma de decisiones son

Efectos futuros:

Tiene que ver con la medida en que los compromisos relacionados con la decisión afectarán el futuro. Una decisión que tiene una influencia a largo plazo, puede ser considerada una decisión de alto nivel, mientras que una decisión con efectos a corto plazo puede ser tomada a un nivel muy inferior.

Reversibilidad:

Se refiere a la velocidad con que una decisión puede revertirse y la dificultad que implica hacer este cambio. Si revertir es difícil, se recomienda tomar la decisión a un nivel alto; pero si revertir es fácil, se requiere tomar la decisión a un nivel bajo.

Impacto:

Esta característica se refiere a la medida en que otras áreas o actividades se ven afectadas. Si el impacto es extensivo, es indicado tomar la decisión a un nivel alto; un impacto único se asocia con una decisión tomada a un nivel bajo.

Calidad:

Este factor se refiere a las relaciones laborales, valores éticos, consideraciones legales, principios básicos de conducta, imagen de la compañía, etc. Si muchos de estos factores están involucrados, se requiere tomar la decisión a un nivel alto; si solo algunos factores son relevantes, se recomienda tomar la decisión a un nivel bajo.

Periodicidad:

Este elemento responde a la pregunta de si una decisión se toma frecuente o excepcionalmente. Una decisión excepcional es una decisión de alto nivel, mientras que una decisión que se toma frecuentemente es una decisión de nivel bajo.

Pasos en el proceso de la toma de decisiones

Determinar la necesidad de una decisión:

El proceso de toma de decisiones comienza con el reconocimiento de la necesidad de tomar una decisión, el mismo lo genera un problema o una disparidad entre cierto estado deseado y la condición real del momento.

Identificar los criterios de decisión:

Una vez determinada la necesidad de tomar una decisión, se deben identificar los criterios que sean importantes para la misma. Vamos a considerar un ejemplo.

Asignar peso a los criterios:

Los criterios enumerados en el paso previo no tienen mayor importancia. Es necesario ponderar cada uno de ellos y priorizar su importancia en la decisión.

Desarrollar todas las alternativas:

Es la base de la toma de decisiones y no es más que desplegar las alternativas. El tomador de la decisión tiene que confeccionar una lista de todas las alternativas posibles y que podrían utilizarse para resolver el problema.

Evaluar las alternativas:

Una vez identificadas las alternativas, el analista de las decisiones tiene que evaluar de manera crítica cada una de ellas. Las ventajas y desventajas de cada alternativa resultan evidentes cuando son comparadas.

Seleccionar la mejor alternativa (Toma de decisiones):

Una vez seleccionada la mejor alternativa se llega al final del proceso de la toma de decisiones, en el proceso racional. Esta selección es bastante simple. El tomador de decisiones tiene que escoger la alternativa que tuvo la calificación más alta en el paso número cinco. La toma de decisiones debe ser totalmente objetiva y lógica a la hora de tomarlas, tiene que tener una meta clara y todas las acciones en el proceso de toma de decisiones llevan de manera consistente a la selección de aquellas alternativas que maximizarán la meta.

Tipos De Decisiones

Decisión Programada:

Son programadas en la medida que son repetitivas y rutinarias, así mismo en la medida que se ha desarrollado un método definitivo para poder manejarlas. Al estar el problema bien estructurado, el mando no tiene necesidad de pasar por el trabajo y gasto de realizar un proceso completo de decisión.

Estas decisiones programadas cuentan con unas guías o procedimientos (pasos secuenciales para resolver un problema) , unas reglas que garanticen consistencias en

las disciplinas y con un alto nivel de justicia , aparte de una política, que son las directrices para canalizar el pensamiento del mando en una dirección concreta.

Decisión no Programada:

"La reestructuración de una organización" o "cerrar una división no rentable", son ejemplos de decisiones no programadas, También "la creación de una estrategia de mercado para un nuevo producto". (Vaquiro, 2012)

PROCESAMIENTO ANALITICO EN LINEA

El procesamiento analítico en línea (OLAP, *Online Analytical Processing*) es una tecnología que se usa para organizar grandes bases de datos empresariales y admitir la inteligencia empresarial. Las bases de datos OLAP se dividen en uno o más cubos. Cada cubo lo organiza y diseña un administrador de cubos para que se adapte a la forma en la que recupera y analiza los datos con el fin de que sea más fácil crear y usar los informes de las tablas dinámicas y los gráficos dinámicos que necesite.

Las bases de datos del procesamiento analítico en línea (OLAP) facilitan las consultas de la inteligencia empresarial. OLAP es una tecnología de base de datos que se ha perfeccionado para efectuar consultas e informes, en vez de procesar transacciones. El origen de datos de OLAP son las bases de datos de procesamiento transaccional en línea (OLTP, *Online Transactional Processing*) que se suelen almacenar en almacenes de datos. Los datos OLAP se derivan de estos datos históricos y se han agregado a estructuras que permiten realizar análisis complejos. Además, se organizan jerárquicamente y se almacenan en cubos en vez de en tablas. Se trata de una tecnología compleja que usa estructuras multidimensionales para proporcionar un acceso rápido a los datos con el fin de efectuar un análisis. Esta organización permite que un informe de tabla dinámica o gráfico dinámico muestre

resúmenes de alto nivel, como totales de ventas en todo un país o región, así mismo los detalles de los sitios en los que las ventas son especialmente altas o bajas.

Las bases de datos OLAP están diseñadas para agilizar la recuperación de datos. Puesto que es el servidor OLAP y no Microsoft Office Excel el que calcula los valores resumidos, hay que enviar menos datos a Excel cuando se cree o se cambie un informe. Esta posibilidad permite trabajar con cantidades mucho mayores de datos de origen que si los datos estuvieran organizados en una base de datos tradicional, en la que Excel recupera todos los registros y después calcula los valores resumidos.

Las bases de datos OLAP contienen dos tipos básicos de datos: medidas, que son datos numéricos, las cantidades y los promedios que usa para tomar decisiones empresariales razonadas, y dimensiones, que son las categorías que usa para organizar estas medidas. Las bases de datos OLAP le ayudan a organizar los datos mediante muchos niveles de detalle, con las mismas categorías que le son familiares para analizar los datos. (Microsoft Office OLAP, 2010)

En las siguientes secciones se describe cada uno de estos componentes con más detalle:

Cubo Estructura de datos que agrega las medidas mediante los niveles y jerarquías de cada una de las dimensiones que desee analizar. Los cubos combinan varias dimensiones, como tiempo, geografía y líneas de productos, con datos resumidos, como cifras de ventas o inventario. Los cubos no son "cubos" en el sentido estrictamente matemático porque no tienen que tener lados iguales. Sin embargo, son una metáfora adecuada para un concepto complejo.

Medida Conjunto de valores basados en una columna de la tabla de hechos del cubo y que suelen ser valores numéricos. Las medidas son los valores centrales del cubo que se pre procesan, agregan y analizan. Entre los ejemplos normales cabe destacar las ventas, los beneficios, los ingresos y los costos.

Elemento Elemento de una jerarquía que representa una o varias repeticiones de datos. Un miembro puede ser único o no único. Por ejemplo, 2007 y 2008 representan miembros únicos en el nivel de años de una dimensión de tiempo, mientras que enero representa un miembro no único en el nivel de meses ya que puede haber varios eneros en la dimensión de tiempo si ésta contiene datos de varios años.

Elemento calculado Elemento de una dimensión cuyo valor se calcula en tiempo de ejecución mediante una expresión. Los valores de los elementos calculados se pueden derivar de otros valores de elementos. Por ejemplo, un elemento calculado, Beneficios, se puede determinar restando el valor del elemento, Costos, del valor del elemento, Ventas.

Dimensión Conjunto de una o varias jerarquías de niveles de un cubo que comprende un usuario y utiliza como base para el análisis de datos. Por ejemplo, una dimensión geográfica puede incluir niveles de país o región, estado o provincia y ciudad. O bien, una dimensión de tiempo puede incluir una jerarquía con niveles de año, trimestre, mes y día. En un informe de tabla o gráfico dinámicos, cada jerarquía se convierte en un conjunto de campos que puede expandir o contraer para revelar niveles inferiores o superiores.

Jerarquía Estructura de árbol lógica que organiza los elementos de una dimensión tal que cada elemento tiene un elemento principal y cero o más elementos secundarios. Un elemento secundario es un elemento en el nivel inmediatamente inferior de una jerarquía que está directamente relacionada con el elemento actual. Por ejemplo, en una jerarquía de tiempo que contiene los niveles Trimestre, Mes y Día, enero es un elemento secundario de Trim1. Un elemento principal es un elemento en el nivel inmediatamente superior de una jerarquía que está directamente relacionado con el miembro actual. El valor principal suele ser una consolidación de los valores de todos sus elementos secundarios. Por ejemplo, en una jerarquía de tiempo que contiene los niveles Trimestre, Mes y Día, Trim1 es el elemento principal de enero.

Nivel En cada jerarquía, los datos se pueden organizar en niveles inferiores y superiores de detalle, como los niveles Año, Trimestre, Mes y Día de una jerarquía de tiempo.

Características OLAP de Excel

Recuperar datos OLAP Puede conectarse a los orígenes de datos OLAP como lo haría a otros orígenes de datos externos. No puede trabajar con bases de datos creadas con Microsoft SQL Server OLAP Services versión 7.0, Microsoft SQL Server Analysis Services versión 2000 ni Microsoft SQL Server Analysis Services versión 2005, los productos de servidor de Microsoft OLAP. Excel también puede trabajar con productos de otros proveedores compatibles con OLE-DB para OLAP.

Puede mostrar datos OLAP solo como informe de tabla dinámica o informe de gráfico dinámico o en una función de hoja de cálculo convertida a partir de un informe de tabla dinámica, pero no como intervalo de datos externo. Puede guardar los informes de las tablas dinámicas y de los gráficos dinámicos OLAP en plantillas de informes. Además, puede crear archivos de conexión de datos de Office (ODC, *Office Data Connection*) (.odc) para conectarse a bases de datos OLAP con el fin de efectuar consultas OLAP. Al abrir un archivo ODC, Excel muestra un informe de tabla dinámica, que ya está listo para que lo diseñe.

Diferencias de características entre datos de origen OLAP y los que no son OLAP

Si trabaja con informes de tablas dinámicas y gráficos dinámicos de datos de origen OLAP y de otros tipos de datos de origen, observará algunas diferencias en las características.

Recuperación de datos Un servidor OLAP devuelve nuevos datos a Excel cada vez que se cambia el diseño del informe. Con otros tipos de datos de origen externo, se consultan todos los datos la vez, o se pueden establecer opciones para que la consulta se realice solamente cuando se muestren elementos de campo de filtros de

informes distintos. También están disponibles otras opciones para actualizar el informe.

En informes basados en datos de origen OLAP, no está disponible la configuración del campo de filtro de informes, la consulta en segundo plano ni la optimización de la memoria.

Tipos de campos Los campos de dimensión de orígenes de datos OLAP se pueden usar solo como campos de fila (series), columna (categoría) o página. Los campos de medida se pueden usar exclusivamente como campos de valores. En otros tipos de orígenes de datos, se pueden usar todos los campos de cualquier parte de un informe.

Acceso a datos de detalle En datos de origen OLAP, el servidor determina qué niveles de detalle están disponibles y calcula los valores de resumen, por tanto, puede que no estén disponibles los registros de detalle que forman los valores de resumen. No obstante, el servidor puede proporcionar campos de propiedad que sí se pueden mostrar. Otros tipos de datos de origen no tienen campos de propiedad, pero se puede mostrar el detalle subyacente de los valores de los campos de datos y elementos, así como los elementos que no contienen datos.

Criterio de ordenación inicial En datos de origen OLAP, los elementos aparecen primero en el orden en que los devuelve el servidor OLAP. Después puede ordenar o reorganizar los elementos manualmente. En otros tipos de datos de origen, los elementos de un nuevo informe aparecen ordenados primero en orden ascendente por nombre de elemento.

Cálculos Los servidores OLAP proporcionan valores resumidos directamente para un informe, de modo que no se pueden cambiar las funciones de resumen de los campos de valores. Para otros tipos de datos de origen, se puede cambiar la función de resumen de un campo de valores y utilizar varias funciones de resumen para el mismo campo de valores. No se pueden crear campos calculados ni elementos calculados en informes con datos de origen OLAP.

Subtotales En los informes con datos de origen OLAP, no puede cambiar la función de resumen de los subtotales. Con otros tipos de datos de origen, se pueden cambiar funciones de resumen de subtotales y mostrar u ocultar subtotales para todos los campos de fila y columna.

VISION MULTIDIMENSIONAL DE LA INFORMACION

En la actualidad, el análisis de información sobre las diferentes transacciones y movimientos de una empresa, requiere de herramientas poderosas que le sean de utilidad en construcción y elaboración de informes que le permitan ver la realidad y tomar decisiones basados en hechos reales. (Dimas & Taborda, 2007)

De acuerdo a las diferentes necesidades de modelamiento de eventos y hechos reales el hombre se ha preocupado por desarrollar diferentes herramientas que le permitan obtener información y analizar el comportamiento de a partir de la modelación de datos.

Los cubos son elementos claves en OLAP, una tecnología que provee rápido acceso a datos en un almacén de datos. Los cubos proveen un mecanismo para buscar datos con rapidez y tiempo de respuesta uniforme independientemente de la cantidad de datos en el cubo o la complejidad del procedimiento de búsqueda. (Martinez Orol, 2010)

Los cubos son subconjuntos de datos de un almacén de datos, organizado y sumariado dentro de una estructura multidimensional. Los datos se sumarian de acuerdo a factores de negocio seleccionados, proveyendo el mecanismo para la rápida y uniforme tiempo de respuesta de las complejas consultas.

La definición del cubo, es el primero de tres pasos en la creación de un cubo.

Los otros pasos son, el especificar la estrategia de sumarización diseñando las agregaciones (elementos pre calculados de datos), y la carga del cubo para procesarlo. (Martinez Orol, 2010).

Hipótesis

“Un Datawarehouse incide positivamente en la toma de decisiones para Gestión de Información Gerencial en Curtiduría Tungurahua S.A. de la Ciudad de Ambato.”

Variable Independiente:

Datawarehouse

Variable Dependiente:

Toma de decisiones en la gestión de la Información gerencial

CAPÍTULO III
METODOLOGÍA
Modalidades Básicas de Investigación

Investigación de Campo

La investigación es de campo ya que se realizó encuestas al personal involucrado sobre los hechos de la problemática actual de la empresa, lo permite determinar claramente los requerimientos para la propuesta así como delimitar el alcance de la misma y realizar un trabajo con mayor objetividad y eficiencia al desarrollar la propuesta.

Investigación documental – bibliográfica

La investigación es documental – bibliográfica ya que con el propósito de fortalecer la investigación se recurre a diferentes fuentes obtenidas de libros especializados, publicaciones, artículos, internet, y otros, mismos que son citados claramente en la bibliografía de este trabajo, así como las fuentes son citadas, durante el desarrollo del presente trabajo donde corresponda.

Tipos o Niveles de Investigación

Exploratorio

La investigación pasa por el nivel de investigación exploratorio porque para la empresa el tema es poco estudiado y casi desconocido, es decir tienen un nivel superficial de conocimiento, la investigación trata de recoger información que permita probar la hipótesis y más que todo permite al investigador conocer mejor el objeto de estudio.

Descriptivo

Es descriptivo porque se busca informar los resultados obtenidos de la investigación entre la comparación de dos variables, tomando en cuenta que para esto se describe actividades exactas de los procesos actuales de la empresa lo que permite analizar generalidades significativas que contribuyen al desarrollo de la propuesta de manera eficiente.

Población y Muestra

La población con la cual se trabajará corresponde al gerente de producción, jefaturas por áreas, los coordinadores de producción, además del personal administrativo de información de la empresa Curtiduría Tungurahua S.A. por tal razón no existe muestra debido a que el universo es muy reducido.

PERSONAL	FRECUENCIA
Gerente de producción	1
Jefatura de planta	1
Jefatura de Ventas	1
Coordinadores de producción	3
Equipo de logística de despachos y ventas	4
Jefatura de logística de compras	1
TOTAL	11

Tabla. 3.1: Población

Elaborado por: Ing. Elena Sánchez

Operacionalización de Variables

Operacionalización de la Variable Independiente

Variable Independiente: Datawarehouse

CONCEPTUALIZACION	DIMENSIONES	INDICADORES	ITEMS BASICOS	TÉCNICAS E INSTRUMENTOS
<p>Un Datawarehouse es una base de datos corporativa 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.</p>	<p>Integrar y depurar información</p>	<p>Organizar, acceder y mostrar: La información de la base de datos del sistema de producción.</p>	<p>¿Cómo reúne la información necesaria para los reportes de producción?</p>	<p>Encuesta Cuestionario</p>
<p>La creación de un datawarehouse representa en la mayoría de las ocasiones el primer paso, desde el punto de vista técnico, para implantar una solución completa y fiable de Inteligencia de Negocios.</p>	<p>Desorganización en la información, lentitud en los reportes</p>	<p>En una sola consulta: Producto, estado de la producción, suministro, materia prima, tiempos.</p>	<p>¿Cómo determina el producto que ofrece mayor rentabilidad y la rotación de materia prima?</p>	<p>Encuesta Cuestionario</p>

Tabla 3.2: Operacionalización de la variable independiente

Elaborado por: Ing. Elena Sánchez

Operacionalización de la Variable Dependiente

Variable Dependiente: Toma de decisiones en la gestión de la información gerencial

CONCEPTUALIZACION	DIMENSIONES	INDICADORES	ITEMS BASICOS	TÉCNICAS E INSTRUMENTOS
Mediante la gestión de la producción se intenta ordenar el flujo de materiales en las empresas productoras o industriales. En una empresa productora hay materias primas que recorren las instalaciones de la fábrica, para ser procesadas y dando como resultado un producto final.	Administra el flujo de la producción	el la forma detallada en tiempos cortos, localización y estado de la producción	Recuperación y acceso a la información en forma detallada en tiempos cortos, localización y estado de la producción	¿Determine el proceso que realiza para la obtención de información sobre el estado de la producción? Encuesta Cuestionario
Pues bien, la gestión de la producción intenta ordenar el flujo de todos los materiales en la fábrica: cuándo hay que fabricar y en qué cantidades.	Prever materia prima y suministros	materia y bienes muebles, equipos materiales de oficina, etc.	Suministros de trabajo muebles, materiales de oficina, etc.	¿Cómo agiliza el proceso de adquisición de suministros o materia prima en cada departamento? Encuesta Cuestionario

Tabla 3.3: Operacionalización de la variable dependiente

Elaborado por: Ing. Elena Sánchez

Técnicas e Instrumentos

Encuesta: Dirigido al personal de producción y ventas, logística de compras, incluso un segundo universo definido por personal de producción para recabar información del manejo de datos en la fábrica.

Entrevista: dirigido a la gerencia de producción cuyo instrumento será guía de la entrevista para recaudar información del proceso de la gestión de producción respecto a las variables.

Plan de Recolección de la Información

Los instrumentos que se utilizaron para la recolección y registro de la información fueron por medio de entrevistas, encuestas, etc.

El uso de Internet buscando alternativas de información virtual, ya que en la actualidad es trascendental para cualquier investigación, complementada siempre por los documentos como tesis, libros, etc.

Plan para el Procesamiento de la Información

Los datos recogidos se transformaron siguiendo los siguientes procedimientos:

- Revisión de la información.
- Tabulación de la información.
- Estudio estadístico de los datos para la presentación de los resultados.

Una vez aplicados y analizada los instrumentos, se procedió a la tabulación de datos.

A continuación se efectuó la estructuración de conclusiones y recomendaciones que organizadas en una propuesta lógica y factible, permitirán participar proactivamente en la solución o minimización de la problemática planteada.

Finalmente, como parte medular de la investigación , se estructura una propuesta pertinente al tema de investigación que nos compete, enfocada a optimizar de los recursos existentes, fiabilidad, confiabilidad y desempeño sin perder de vista los objetivos trazados.

Encuesta N° 1. Dirigida a 11 personas involucradas directamente con el movimiento general de la empresa desde la producción hasta la venta del producto.

CURTIDURÍA TUNGUARAHUA S.A

Lugar de aplicación de Encuesta: _____

Objetivo de la Encuesta.- Determinar la incidencia de la carencia de herramientas para la gestión de información.

Sus respuestas le permitirán al investigador

✚ Desarrollar un trabajo real y efectivo.

✚ Agradecemos su colaboración

Sírvase marcar en el paréntesis con una X de acuerdo a su criterio

N.	PREGUNTAS	RESPUESTAS	COD.
1	¿Existen en la empresa aplicaciones o software informático que procese información rápidamente?	<ul style="list-style-type: none"> • Si • No 	1. () 2. ()
2	¿Existen políticas y procedimientos definidos que regulan sus actividades de producción y que involucren el tema tecnológico en los procesos?	<ul style="list-style-type: none"> • Si • No 	1. () 2. ()
3	¿Posee tecnología apropiada para mejorar el rendimiento de sus procesos particulares?	<ul style="list-style-type: none"> • Si • No 	1. () 2. ()
4	¿Cómo calificaría los sistemas informáticos que están implantados como apoyo a sus actividades?	<ul style="list-style-type: none"> • Aceptables • No aceptables 	1. () 2. ()
5	¿Cómo calificaría las línea de comunicación entre los distintos centros de costo para la logística de producción, compras y / o ventas?	<ul style="list-style-type: none"> • Aceptables • No aceptables 	1. () 2. ()
6	¿Cómo usuario se siente satisfecho con la utilización de las aplicaciones tecnológicas actuales de la empresa?	<ul style="list-style-type: none"> • Si • No 	1. () 2. ()
7	¿Ha tenido problemas de logística de producción causados por el sistema de información actual?	<ul style="list-style-type: none"> • Si • No 	1. () 2. ()
8	¿Ha tenido problemas de logística de reabastecimiento causados por el sistema de información actual?	<ul style="list-style-type: none"> • Si • No 	1. () 2. ()
9	¿Ha tenido problemas de logística de ventas y despachos causados por el sistema de información actual?	<ul style="list-style-type: none"> • Si • No 	1. () 2. ()

10	¿Utilizar un herramienta informática más eficiente le ayudaría en la toma de decisiones apropiadas que mejores sus procesos?	<ul style="list-style-type: none"> • Si • No 	1. () 2. ()
----	--	--	------------------

Tabla 3. 4: Encuesta N°. 1

Elaborado por: Ing. Elena Sánchez

Encuesta N° 2. Dirigida a 7 personas involucradas más a fondo con el movimiento de producción para determinar de manera más específica en que proceso se produce la problemática de la gestión de información.

CURTIDURÍA TUNGURAHUA S.A

Lugar de aplicación de Encuesta: _____

Objetivo de la Encuesta.- Determinar la incidencia de la carencia de herramientas para la gestión de información dentro de su proceso.

Sus respuestas le permitirán al investigador

✚ Desarrollar un trabajo real y efectivo.

✚ Agradecemos su colaboración

Sírvase marcar en el paréntesis con una X de acuerdo a su criterio

N.	PREGUNTAS	RESPUESTAS	COD.
1	¿Dispone usted información detallada de sus Órdenes de Producción?	<ul style="list-style-type: none"> • Siempre • A veces • Rara vez 	1. () 2. () 3. ()
2	¿Con que frecuencia existe errores en el detalle de su orden de Producción?	<ul style="list-style-type: none"> • Siempre • Frecuentemente • Rara vez 	1. () 2. () 3. ()
3	¿Está usted satisfecho con los reportes de Producción de la empresa?	<ul style="list-style-type: none"> • Plenamente • Aceptablemente • Rara vez 	1. () 2. () 3. ()
4	¿Qué tan frecuentes son los reprocesos de productos?	<ul style="list-style-type: none"> • Siempre • Frecuentemente • Rara vez 	1. () 2. () 3. ()
5	¿Dispone usted información detallada de los productos y sus cantidades aplicadas a un artículo en particular?	<ul style="list-style-type: none"> • Mucho • Poco • Nada 	1. () 2. () 3. ()
6	¿Conoce usted que formula de producción es la más económica a aplicar?	<ul style="list-style-type: none"> • Siempre • Frecuentemente • Rara vez 	1. () 2. () 3. ()
7	¿Con que frecuencia existe errores en el detalle de su inventario de materia prima?	<ul style="list-style-type: none"> • Frecuentemente • Rara vez • Nunca 	1. () 2. () 3. ()

8	¿Con que frecuencia se retrasa la entrega de su Orden de Producción?	<ul style="list-style-type: none"> • Frecuentemente • Ocasionalmente • Nunca 	1. () 2. () 3. ()
9	¿Conoce el volumen de venta por cliente?	<ul style="list-style-type: none"> • Si • A veces • No 	1. () 2. () 3. ()
10	¿Conoce el volumen de venta por zona?	<ul style="list-style-type: none"> • Siempre • Frecuentemente • Rara vez 	1. () 2. () 3. ()
11	¿Conoce usted el stock real inmediato en bodega para ofertar a sus clientes?	<ul style="list-style-type: none"> • Siempre • Frecuentemente • Rara vez 	1. () 2. () 3. ()

Tabla 3.5: Encuesta N°. 2

Elaborado por: Ing. Elena Sánchez

Entrevistas dirigida a la gerencia de producción de la empresa CURTIDURIA TUNGURAHUA S.A.

N° Pregunta	Pregunta	Respuesta
1	¿Confía en los reportes que le facilitan sus subordinados, por qué?	
2	Como elabora sus reportes para sus jefaturas.	
3	¿Cuánto tarda en verificar o enviar información requerida o solicitada?	
4	¿Se ha visto perjudicado por informes que le hayan o haya presentado?	
5	¿Por qué comete errores en presentar resultados de algún requerimiento de información?	
6	¿Cómo toma sus decisiones?	

Tabla 3.6: Entrevista

Elaborado por: Ing. Elena Sánchez

CAPITULO IV
ANÁLISIS E INTERPRETACIÓN DE RESULTADOS

Encuesta aplicada a funcionarios de los distintos departamentos de la empresa CURTIDURIA TUNGURAHUA S.A

Encuesta N° 1

Pregunta N° 1. ¿Existen en la empresa aplicaciones o software informático que procese información rápidamente?

OPCIONES	FRECUENCIA	RESULTADOS %
Si	3	27%
No	8	73%
TOTAL	11	100%

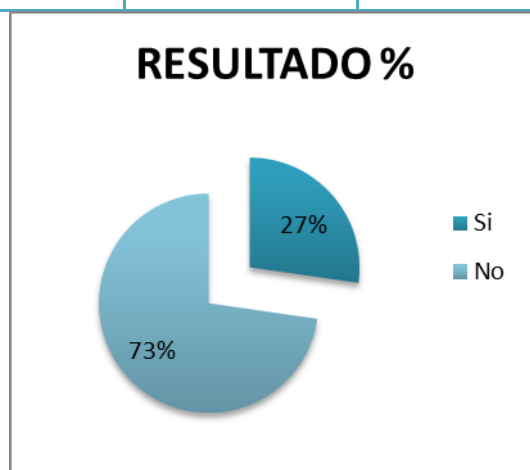


Tabla 4.1: Tabulación de respuesta a encuesta 1 pregunta 1

Fig. 4.1: Tabulación de respuesta a encuesta 1 pregunta 1

Elaborado por: Ing. Elena Sánchez

Análisis e Interpretación de Resultados

Del universo encuestado el 27% opina que las aplicaciones para procesar información en empresa funcionan adecuadamente, frente a un 73% que indica que estas aplicaciones no cumplen el procesamiento de información de manera eficiente, lo que nos permite interpretar que el software existe no es la herramienta adecuada para el procesamiento de información.

Pregunta N° 2. ¿Existen políticas y procedimientos definidos que regulan sus actividades de producción y que involucren el tema tecnológico en los procesos?

OPCIONES	FRECUENCIA	RESULTADOS %
Si	6	55%
No	5	45%
TOTAL	11	100%

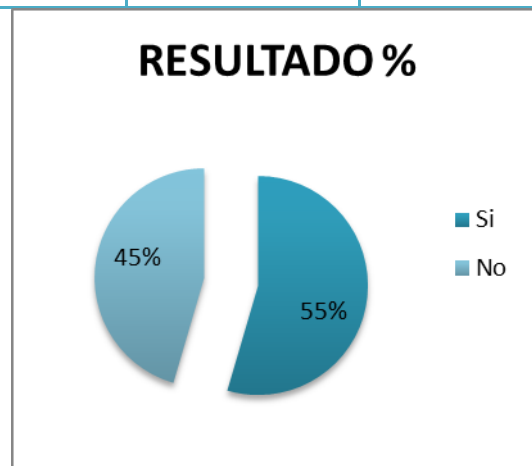


Tabla 4.2: Tabulación de respuesta a encuesta 1 pregunta 2

Fig. 4.2: Tabulación de respuesta a encuesta 1 pregunta 2

Elaborado por: Ing. Elena Sánchez

Análisis e Interpretación de Resultados

45 % de las personas que respondieron a esta pregunta piensan que no existen políticas ni procedimientos que involucren el tema tecnológico definidas en la empresa que aclaren los lineamientos a seguir frente a un 55% que piensa que si existen definidas políticas y procedimientos sin embargo estas no se aplican.

Pregunta N° 3. ¿Posee tecnología apropiada para mejorar el rendimiento de sus procesos particulares?

OPCIONES	FRECUENCIA	RESULTADOS %
Si	3	27%
No	8	73%
TOTAL	11	100%

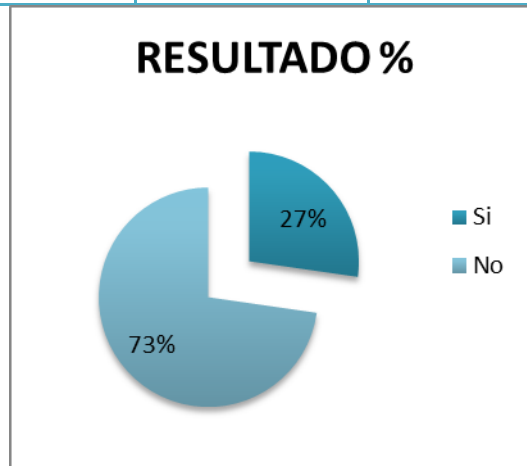


Tabla 4.3: Tabulación de respuesta a encuesta 1 pregunta 3

Fig. 4.3: Tabulación de respuesta a encuesta 1 pregunta 3

Elaborado por: Ing. Elena Sánchez

Análisis e Interpretación de Resultados

Las personas involucradas con esta preguntan nos indican que posee la tecnología apropiada refiriéndose a equipos de cómputo y un sistema de gestión para realizar sus procesos en un 27%, mientras que un 73% hace referencia que este tema debe ser mejorado.

Lo que nos permite concluir que el sistema de gestión actual no satisface con las necesidades de los usuarios a nivel de jefaturas.

Pregunta N° 4. ¿Cómo calificaría los sistemas informáticos que están implantados como apoyo a sus actividades?

OPCIONES	FRECUENCIA	RESULTADOS %
Aceptables	4	36%
No aceptables	7	64%
TOTAL	11	100%



Tabla 4.4: Tabulación de respuesta a encuesta 1 pregunta 4
 Fig. 4.4: Tabulación de respuesta a encuesta 1 pregunta 4
 Elaborado por: Ing. Elena Sánchez

Análisis e Interpretación de Resultados

Esta pregunta nos arroja que un 64% de los sistemas informáticos están dentro de un nivel de no aceptación o no cumplimiento eficiente para ser un apoyo a los procesos, mientras que un 36% opina que si es aceptable el nivel de apoyo para sus procesos. Con este resultado podemos ver que existe un nivel alto de insatisfacción con el actual sistema de gestión de la información.

Pregunta N° 5. ¿Cómo calificaría las línea de comunicación entre los distintos centros de costo para la logística de producción, compras y / o ventas?

OPCIONES	FRECUENCIA	RESULTADOS %
Aceptables	6	55%
No aceptables	5	45%
TOTAL	11	100%



Tabla 4.5: Tabulación de respuesta a encuesta 1 pregunta 5

Fig. 4.5: Tabulación de respuesta a encuesta 1 pregunta 5

Elaborado por: Ing. Elena Sánchez

Análisis e Interpretación de Resultados

Los canales de comunicación entre centros de costos para el perfecto funcionamiento de los sistemas y procesos en cada una de las áreas se encuentran en 55% aceptables según califica el universo encuestado, mientras que el 45% de estos, considera que debe ser mejorado.

Considerando que el sistema de gestión actual no permite tener claramente definido los canales de comunicación, nos permite decir que mejorar esta herramienta beneficiará a la empresa.

Pregunta N° 6. ¿Cómo usuario se siente satisfecho con la utilización de las aplicaciones tecnológicas actuales de la empresa?

OPCIONES	FRECUENCIA	RESULTADOS %
Si	2	18%
No	9	82%
TOTAL	11	100%



Tabla 4.6: Tabulación de respuesta a encuesta 1 pregunta 6

Fig. 4.65: Tabulación de respuesta a encuesta 1 pregunta 6

Elaborado por: Ing. Elena Sánchez

Análisis e Interpretación de Resultados

Observando la pregunta apenas el 18 % indica que está satisfecho con las aplicaciones tecnológicas de la empresa, ratificando lo que nos indican en la pregunta N° 4 de esta misma encuesta, lo que nos deja un alto porcentaje el 82 % del personal que indica que no está satisfecho y que nos permite interpretar que se debería implementar un solución que mejore este resultado.

Pregunta N° 7. ¿Ha tenido problemas de logística de producción causados por el sistema de información actual?

OPCIONES	FRECUENCIA	RESULTADOS %
Si	10	91%
No	1	9%
TOTAL	11	100%



Tabla 4.7: Tabulación de respuesta a encuesta 1 pregunta 7

Fig. 4.7: Tabulación de respuesta a encuesta 1 pregunta 7

Elaborado por: Ing. Elena Sánchez

Análisis e Interpretación de Resultados

El 91% de las personas nos dicen que al menos una vez han tenido algún inconveniente en los procesos de producción por causa de la gestión de la información en los sistemas de procesamiento actual, frente a un mínimo 9% de personas que nos indica que no han tenido inconvenientes.

Claramente podemos interpretar que la herramienta de gestión actual posee inconvenientes para la presentación de información en línea.

Pregunta N° 8. ¿Ha tenido problemas de logística de reabastecimiento causados por el sistema de información actual?

OPCIONES	FRECUENCIA	RESULTADOS %
Si	10	91%
No	1	9%
TOTAL	11	100%



Tabla 4.8: Tabulación de respuesta a encuesta 1 pregunta 8
 Fig. 4.8: Tabulación de respuesta a encuesta 1 pregunta 8
 Elaborado por: Ing. Elena Sánchez

Análisis e Interpretación de Resultados

Frente a esta pregunta los encuestados nos transmiten que un 91% de ellos han tenido inconvenientes debido al mal funcionamiento del actual sistema de información, mientras que un 9% indica que no se ha presentado inconvenientes con la logística de reabastecimiento de materia prima. Lo que nos enfrenta a la realidad de que el sistema actual de gestión de información no es el adecuado.

Pregunta N° 9. ¿Ha tenido problemas de logística de ventas y despachos causados por el sistema de información actual?

OPCIONES	FRECUENCIA	RESULTADOS %
Si	10	91%
No	1	9%
TOTAL	11	100%



Tabla 4.9: Tabulación de respuesta a encuesta 1 pregunta 9

Fig. 4.9: Tabulación de respuesta a encuesta 1 pregunta 9

Elaborado por: Ing. Elena Sánchez

Análisis e Interpretación de Resultados

El 91% de las personas opina que se han presentado problemas en la logística de despacho y ventas perjudicando su proceso de alguna forma, mientras que el 9% de las personas opina que no ha tenido problemas en este aspecto.

Interpretando estos resultados podemos definir que efectivamente el sistema de gestión de información actual no satisface requerimientos de usuarios a este nivel de jerarquía.

Pregunta N° 10. ¿Utilizar un herramienta informática más eficiente le ayudaría en la toma de decisiones apropiadas que mejores sus procesos?

OPCIONES	FRECUENCIA	RESULTADOS %
Si	11	100%
No	0	0%
TOTAL	11	100%



Tabla 4.10: Tabulación de respuesta a encuesta 1 pregunta 10
Fig. 4.10: Tabulación de respuesta a encuesta 1 pregunta 10
Elaborado por: Ing. Elena Sánchez

Análisis e Interpretación de Resultados

El 100 % del universo encuestado opina que una herramienta informática de mayor eficiencia les ayudaría a tomar decisiones mayormente acertadas para el correcto funcionamiento de sus procesos.

Encuesta N° 2

Pregunta N° 1. ¿Dispone usted información detallada de sus órdenes de producción?

OPCIONES	FRECUENCIA	RESULTADOS %
Siempre	6	86%
A veces	1	14%
Rara vez	0	0%
TOTAL	7	100%



Tabla 4.11: Tabulación de respuesta a encuesta 2 pregunta 1

Fig. 4.11: Tabulación de respuesta a encuesta 2 pregunta 1

Elaborado por: Ing. Elena Sánchez

Análisis e Interpretación de Resultados

Mediante las encuestas aplicados a 7 miembros del área de producción involucrados directamente con la pregunta se obtiene que el 84% de los mismos siempre disponen de la información detallada de las ordenes de producción que deben ejecutar diariamente, mientras que el 16 % dice que al detalle la información la posee a veces.

Lo que se nos refleja que no siempre es confiable la información sobre consumos de materia prima por lo tanto el coste de producción y rentabilidad tiene un margen de error.

Pregunta N° 2. ¿Con que frecuencia existe errores en el detalle de su orden de producción?

OPCIONES	FRECUENCIA	RESULTADOS %
Siempre	0	0%
Frecuentemente	7	100%
Rara vez	0	0%
TOTAL	7	100%



Tabla 4.12: Tabulación de respuesta a encuesta 2 pregunta 2

Fig. 4.12: Tabulación de respuesta a encuesta 2 pregunta 2

Elaborado por: Ing. Elena Sánchez

Análisis e Interpretación de Resultados

Las encuestas dirigidas a 7 miembros del área de producción nos indica que el 100% de las veces los errores en las ordenes de producción son frecuentes.

Como consecuencia de esto deducimos que la información alimentada al sistema no siempre es apropiada; y además al aplicarse la encuesta casi todo el personal indica que el tiempo de respuesta del sistema actual no es muy eficiente.

Pregunta N° 3. ¿Está usted satisfecho con los reportes de Producción de la empresa?

OPCIONES	FRECUENCIA	RESULTADOS %
Plenamente	0	0%
Aceptablemente	5	71%
Rara vez	2	29%
TOTAL	7	100%



Tabla 4.13: Tabulación de respuesta a encuesta 2 pregunta 3

Fig. 4.13: Tabulación de respuesta a encuesta 2 pregunta 3

Elaborado por: Ing. Elena Sánchez

Análisis e Interpretación de Resultados

Las encuestas dirigidas a 7 miembros entre el área de producción y ventas a nivel de jefaturas se obtuvieron que 71% de las veces los reportes que presentan sus subordinados son confiables y aceptables y un 29% de las veces no son aceptables.

Mayormente por tiempos de presentación más no confiabilidad en la información ya que muchos de los procesos hay que realizarlos manualmente, lo que nos indica que el problema se presenta en que no poseen un sistema adecuado de disponibilidad de información.

Pregunta N° 4. ¿Qué tan frecuentes son los reprocesos de productos?

OPCIONES	FRECUENCIA	RESULTADOS %
Siempre	0	0%
Frecuentemente	5	71%
Rara vez	2	29%
TOTAL	7	100%

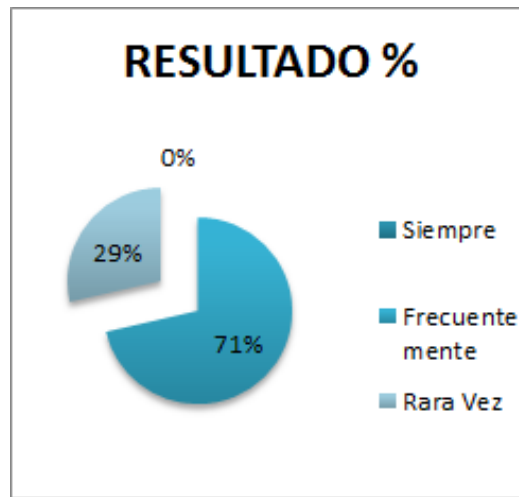


Tabla 4.14: Tabulación de respuesta a encuesta 2 pregunta 4
 Fig. 4.14: Tabulación de respuesta a encuesta 2 pregunta 4
 Elaborado por: Ing. Elena Sánchez

Análisis e Interpretación de resultados

La encuesta aplicada a nivel del personal de producción específicamente nos indica que el apenas el 29 % de las veces no realiza un reproceso de productos, frente a un 71% que opina que se produce un reproceso de producto lo que afecta el costeo de productos, deduciendo que la falta de información en línea causa incluso que se infle el costo de producción.

Pregunta N° 5. ¿Dispone usted información detallada de los productos y sus cantidades aplicadas a un artículo en particular?

OPCIONES	FRECUENCIA	RESULTADOS %
Siempre	0	0%
Frecuentemente	5	71%
Rara vez	2	29%
TOTAL	7	100%

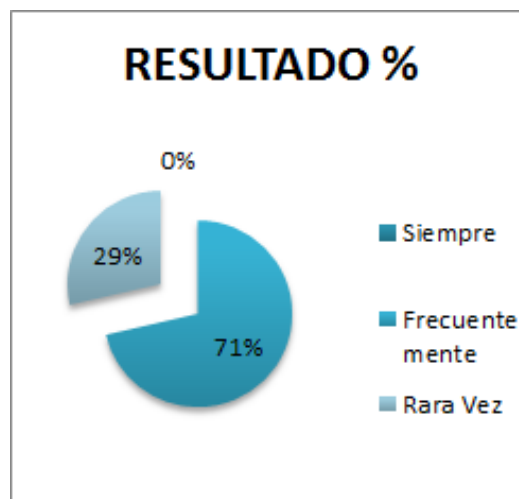


Tabla 4.15: Tabulación de respuesta a encuesta 2 pregunta 5

Fig. 4.15: Tabulación de respuesta a encuesta 2 pregunta 5

Elaborado por: Ing. Elena Sánchez

Análisis e Interpretación de Resultados

La encuesta nos arroja que el 71% de los encuestados conocen la formulación de cada producto a realizarse, lo que se entiende porque los reportes de producción no son perfectamente confiables, sin embargo nos indican que el tiempo de empleo para obtener información es deficiente, notándose también que el 29% de los encuestados nos indica que rara vez dispone de esta información en línea y lo realiza empíricamente.

Pregunta N° 6. ¿Conoce usted que formula de producción es la más económica a aplicar?

OPCIONES	FRECUENCIA	RESULTADOS %
Siempre	1	14%
Frecuentemente	0	0%
Rara vez	6	86%
TOTAL	7	100%

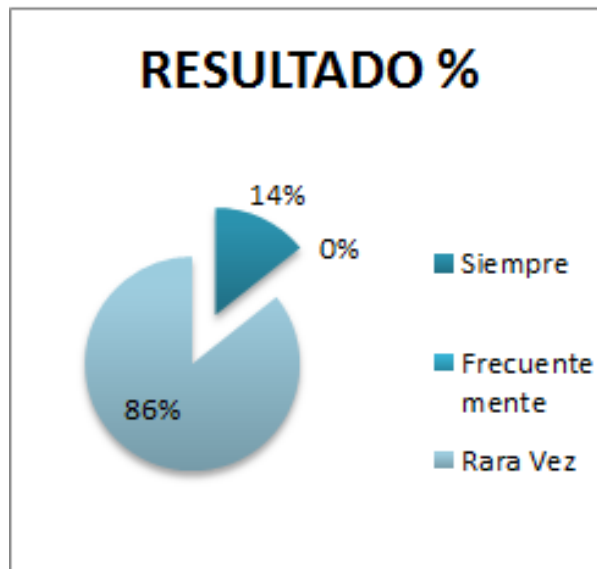


Tabla 4.16: Tabulación de respuesta a encuesta 2 pregunta 6
 Fig. 4.16: Tabulación de respuesta a encuesta 2 pregunta 6
 Elaborado por: Ing. Elena Sánchez

Análisis e Interpretación de Resultados

Los encuestados a través de esta pregunta nos indican que apenas el 14% de ellos conocen que formulas es la más adecuada en cuanto en costo y calidad se debe aplicar, sin embargo un alarmante 86% restante no conoce a ciencia cierta lo cual provoca que exista un margen de pérdida de recursos por el desconocimiento de información.

Pregunta N° 7. ¿Con que frecuencia existe errores en el detalle de su inventario de materia prima?

OPCIONES	FRECUENCIA	RESULTADOS %
Frecuentemente	5	71%
Rara vez	2	29%
Nunca	0	0%
TOTAL	7	100%

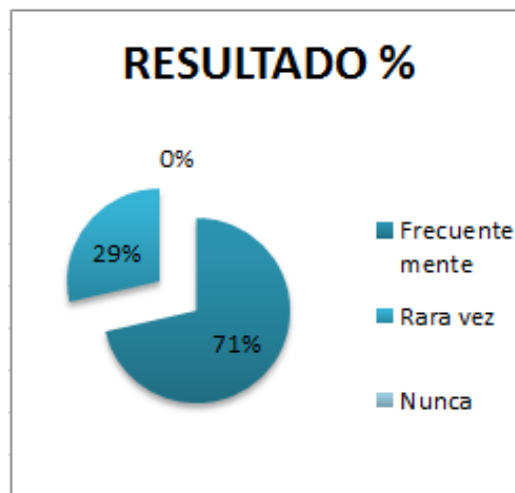


Tabla 4.17: Tabulación de respuesta a encuesta 2 pregunta 7

Fig. 4.17: Tabulación de respuesta a encuesta 2 pregunta 7

Elaborado por: Ing. Elena Sánchez

Análisis e Interpretación de Resultados

Los encuestados para esta pregunta respondieron que el 29% de las veces el muestreo de sus inventarios son casi acertados, mientras que el 71% aduce que frecuentemente existe algún error.

Esto nos permite identificar que ya sea por la lentitud del sistema o por errores operativos el inventario no es el adecuado al 100%.

Pregunta N° 8. ¿Con que frecuencia se retrasa la entrega de su Orden de Producción?

OPCIONES	FRECUENCIA	RESULTADOS %
Frecuentemente	4	57%
Ocasionalmente	3	43%
Nunca	0	0%
TOTAL	7	100%

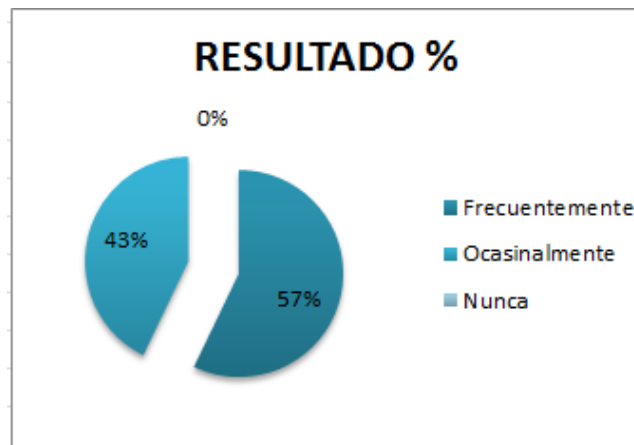


Tabla 4.18: Tabulación de respuesta a encuesta 2 pregunta 8

Fig. 4.18: Tabulación de respuesta a encuesta 2 pregunta 8

Elaborado por: Ing. Elena Sánchez

Análisis e Interpretación de Resultados

Podemos observar que un 57% de los encuestados nos comunican que frecuentemente los pedidos se retrasan, y un 43% ocasionalmente, de lo que podemos interpretar que el 100% de los pedidos salen retrasados, ocasionando que no se cumpla con los clientes provocando distintos problemas de imagen y económicos a la empresa.

Pregunta N° 9. ¿Conoce usted el volumen de venta por cliente?

OPCIONES	FRECUENCIA	RESULTADOS %
Si	1	14%
A veces	3	43%
No	3	43%
TOTAL	7	100%

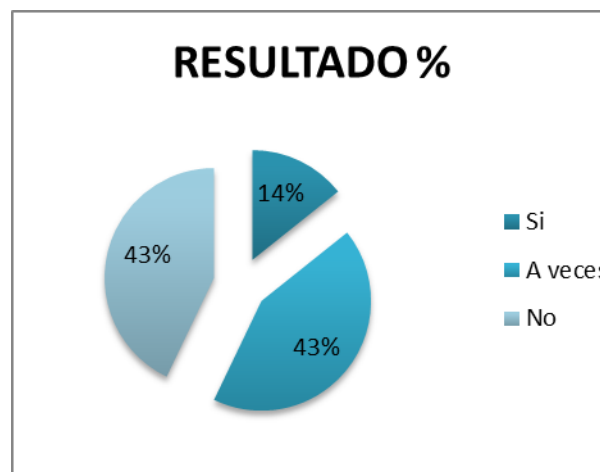


Tabla 4.19: Tabulación de respuesta a encuesta 2 pregunta 9
Fig. 4.19: Tabulación de respuesta a encuesta 2 pregunta 9
Elaborado por: Ing. Elena Sánchez

Análisis e Interpretación de Resultados

Podemos observar que del personal encuestado conoce su volumen de ventas apenas un 14% lo que implica que el 96% restante lo conoce a veces o simplemente lo desconoce.

Esta información la podemos considerar como inaceptable ya que implica una incidencia en las transacciones comerciales de la empresa, ya que al ser de tipo analítico para la jefatura de ventas, cuyo resultado influye directamente y de manera crítica para el área de cobranza.

Pregunta N° 10. ¿Conoce el volumen de venta por zona?

OPCIONES	FRECUENCIA	RESULTADOS %
Siempre	0	0%
Frecuentemente	3	43%
Rara vez	4	57%
TOTAL	7	100%

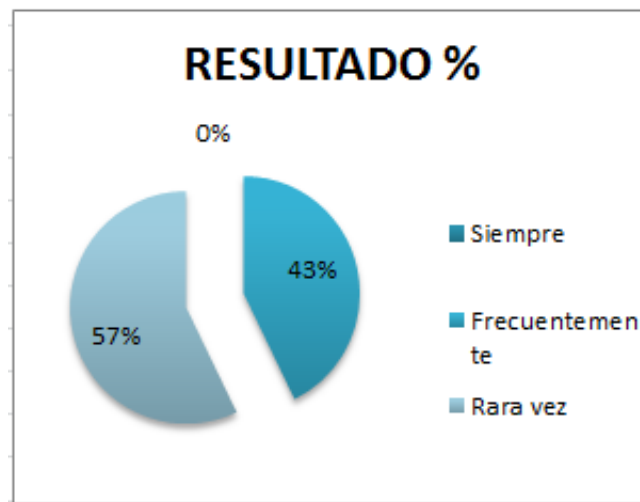


Tabla 4.20: Tabulación de respuesta a encuesta 2 pregunta 10

Fig. 4.20: Tabulación de respuesta a encuesta 2 pregunta 10

Elaborado por: Ing. Elena Sánchez

Análisis e Interpretación de Resultados

De un universo de 7 encuestados del área de ventas el 57 % tiene conocimiento frecuente de la zona y volumen de ventas y un 43% conoce rara vez esta información, lo cual nos permite afirmar que la información que la empresa maneja sobre el tema es a medias asertiva.

Pregunta N° 11. ¿Conoce usted el stock real inmediato en bodega para ofertar a sus clientes?

OPCIONES	FRECUENCIA	RESULTADOS %
Siempre	0	0%
Frecuentemente	3	43%
Rara vez	4	57%
TOTAL	7	100%

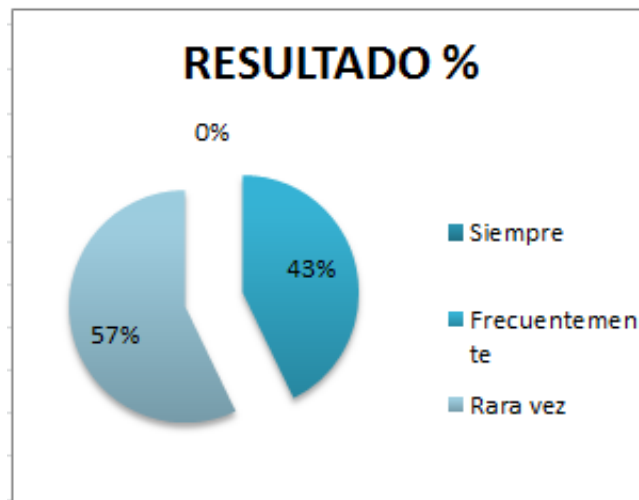


Tabla 4.21: Tabulación de respuesta a encuesta 2 pregunta 11
 Fig. 4.21: Tabulación de respuesta a encuesta 2 pregunta 11
 Elaborado por: Ing. Elena Sánchez

Análisis e Interpretación de Resultados

De los encuestados del área de ventas y producción el 57 % sabe con frecuentemente que inventario posee para ofertar a sus clientes, debido a que no existe una herramienta que proporcione esta información inmediata existe un 43% que no lo conoce a exactitud.

Verificación de la Hipótesis

En primer lugar se procede a plantear las hipótesis para ser verificadas.

Formulación de la hipótesis

✚ H_0 = Hipótesis Nula

✚ H_1 = Hipótesis Alterna

H_0 = Un Datawarehouse no incide positivamente en la toma de decisiones para Gestión de Información Gerencial en Curtiduría Tungurahua S.A. de la Ciudad de Ambato.

H_1 = Un Datawarehouse incide positivamente en la toma de decisiones para Gestión de Información Gerencial en Curtiduría Tungurahua S.A. de la Ciudad de Ambato.

Esta investigación utiliza los siguientes métodos para la verificación de la hipótesis:

Método Lógico Deductivo

Para la verificación de la hipótesis mediante este método utilizaremos la interpretación de las encuestas y entrevistas aplicadas al personal para definir un estado actual de los procesos de la empresa, y como estos se gestionan en su sistema transaccional, lo que además permitirá definir los requerimientos que deberán ser tomados en cuenta en la propuesta.

Estado actual de la información con la que toma decisiones la empresa Curtiduría Tungurahua S.A

PROCESO		ACTUAL	ESPERADO
Pedidos de Materia Prima	De acuerdo a la información obtenida durante el proceso, el personal nos indica que no siempre se conoce el inventario existente lo cual no permite una óptima realización de este proceso	Lo ideal sería tener información de inventarios definidos por centros de costo incluso, que sean reales para logística de reabastecimiento no cometa errores y no desencadene problemas como retraso de pedidos de clientes, o genere inventarios obsoletos entre otros.	
Producción, venta y reproceso de productos	El personal participativo nos indica que no siempre pueden acceder a información sobre su producción, venta y reprocesos ya que el sistema actual es muy lento en sacar los reportes que prefieren hacerlos en excel lo que está susceptible a errores y luego esta información cuando tienen tiempo verifica con el sistema de	Se debería tener un reporte en excel que se actualice automáticamente cuando el usuario lo requiera, y que este pueda decidir si desea datos de producción, ventas o reprocesos, o su vez tres reportes que indiquen esta información pero que sea actualizables al momento necesario.	

	gestión.	
Bodega de producto terminado	El coordinador de esta bodega nos indicaba que el sistema de gestión actual tarda mucho en actualizar el inventario de pt que tiene en bodega para ofertar a los clientes inmediatamente, prefería dirigirse físicamente a los estantes y realizar la toma física de esta información.	Una herramienta de preferencia excel que permita tener un listado actualizable cuando se lo requiera de la existencia de pt.
Consumos	La gerencia de producción, e incluso aquí la coordinación financiera realiza un análisis de consumos mensuales de materia prima para tener una panorama tanto de gasto como reabastecimiento y no sufrir imprevistos, proceso que toma desde el cierre de mes hasta tres días realizarlo recopilando información de varias personas.	Un reporte de preferencia en excel que permita tener estos consumos de alguna manera ligados por centro de costos, para poder realizar las tareas descritas en el proceso en menos tiempo.
Ingresos de materia prima a los centros de	Desde el punto de vista de las coordinaciones de	Tener un reporte de preferencia en excel que

costos	producción, muchas de las veces son cuestionados sobre los consumos de materia prima, para su justificación, necesitan verificar que los ingresos de MP ¹² que recibieron a sus centros de costos sean los correctos, para esto realizan procesos manuales de concatenación de información del sistema en excel.	permita actualizar las entregas de MP a cada centro de costo en las fechas que el usuario requiera.
Préstamos Devoluciones	- Existen movimientos transaccionales que indican los préstamos de materia prima y devoluciones de producto terminado pero estos reportes tardan mucho tiempo en ejecutarse en el actual sistema transaccional.	Los mismos reportes pero que su rendimiento sea óptimo entiéndase esto por tiempos más cortos de procesamiento.

Tabla 4.22: Comparativo de proceso actual vs esperado
Elaborado por: Ing. Elena Sánchez

Conociendo el estado actual de la empresa se puede lógicamente deducir que la hipótesis H₁ que indica que el desarrollo de un datawarehouse incide

¹² MP, Materia Prima

positivamente en la toma de decisiones para Gestión de Información Gerencial en Curtiduría Tungurahua S.A. de la Ciudad de Ambato es la aceptada, ya que mejoraría notablemente los procesos actuales de la empresa, por lo tanto la hipótesis queda verificada a través del método lógico.

Método Estadístico

La Ji Cuadrada simbólicamente definida como X^2 es una prueba estadística para evaluar hipótesis acerca de la relación entre dos variables categóricas

Elección de la prueba estadística

Para la verificación de la hipótesis se escogió la prueba Ji cuadrada, cuya fórmula es la siguiente:

$$x^2 = \sum \left(\frac{(f_0 - f_e)^2}{f_e} \right)$$

Simbología:

x^2 = Ji Cuadrada

f_0 = Frecuencia observada

f_e = Frecuencia esperada

Para realizar la matriz de tabulación cruzada se toma en cuenta dos preguntas de la encuesta realizada al personal de Curtiduría Tungurahua S.A, como se indica a continuación:

Pregunta N° 1. ¿Existen en la empresa aplicaciones o software informático que procese información rápidamente?

OPCIONES	FRECUENCIA	RESULTADOS %
Si	3	27%
No	8	83%
TOTAL	11	100%

Tabla 4.23.: Comprobación de hipótesis: Aplicaciones para procesar información

Pregunta N° 10. ¿Utilizar un herramienta informática más eficiente le ayudaría en la toma de decisiones apropiadas que mejores sus procesos?

OPCIONES	FRECUENCIA	RESULTADOS %
Si	11	100%
No	0	0%
TOTAL	11	100%

Tabla 4.24: Comprobación de hipótesis: Herramienta de toma de decisiones

✚ Definición del nivel de significación

El nivel de significación escogido para la investigación es del 5%

$$\alpha = 0.05$$

✚ Grado de Libertad

$$\text{Grado de libertad} = (\text{reglones} - 1)(\text{columna} - 1)$$

$$Gl = (r - 1)(c - 1)$$

$$Gl = (2 - 1)(2 - 1)$$

$$Gl = 1$$

✚ Frecuencia Observadas

Parámetros	Alternativas		Total
	SI	NO	
Existencia de herramientas informáticas para el procesamiento de información.	3	8	11
Mejoramiento de las herramientas de procesamiento para la toma de decisiones.	11	0	11
Total	14	8	22

Tabla 4.25.: Frecuencias Observadas

✚ Frecuencias Esperadas

Para calcular la frecuencia esperada se utiliza la siguiente fórmula:

$$fe = \frac{(Total\ o\ marginal\ de\ renglon)(Total\ o\ marginal\ de\ columna)}{N}$$

Parámetros	Alternativas	
	SI	NO
Existencia de herramientas informáticas para el procesamiento de información.	11	11
Mejoramiento de las herramientas de procesamiento para la toma de decisiones.	11	11

Tabla 4.26.: Frecuencias Esperadas

✚ Cálculo de Ji Cuadrada

$x^2 = \sum \left(\frac{(f_o - f_e)^2}{f_e} \right)$	f_o	f_e	$f_o - f_e$	$(f_o - f_e)^2$	$\frac{(f_o - f_e)^2}{f_e}$
Existencia de herramientas /SI	3	11	-8	64	5.8
Existencia de Herramientas /NO	8	11	-3	9	0.8
Mejoramiento de las herramientas /SI	11	11	0	0	0
Mejoramiento de las herramientas /NO	0	11	-11	121	11
					$x^2 = 17.6$

Tabla 4.27.: Cálculo de x^2

Decisión:

El valor de $x_t^2 = 10,9861$ (Ver Anexo 2) $< x_c^2 = 17.6$ (Tabla 4.26)

Por consiguiente se acepta la hipótesis alterna, es decir, que la aplicación de una nueva herramienta como un Datawarehouse incidirá positivamente en la toma de decisiones para Gestión de Información Gerencial en Curtiduría Tungurahua S.A. de la Ciudad de Ambato y se rechaza la hipótesis nula.

CAPITULO V

CONCLUSIONES Y RECOMENDACIONES

Conclusiones

- ✚ Del análisis de esta investigación se concluye que los usuarios a nivel de coordinaciones, jefaturas, y gerencia tienen problemas al no disponer de información en línea que les permita tomar decisiones, y si la poseen esta llega a destiempo retardando procesos que significativamente es una pérdida de recursos para Curtiduría Tungurahua S.A

- ✚ Se concluye que los usuarios no disponen de herramientas que optimicen sobre todo el tiempo de procesamiento de la información, se escudan en este problema para no seguir algunas políticas y procedimientos que la empresa tiene definidas, ocasionando una desorganización en algunos procesos como las ordenes de producción, y despachos, reabastecimiento de materia prima.

- ✚ Se concluye que es necesario el diseño de un datawarehouse para disponer de información en línea cuando se requiera tomar decisiones acertadas que no perjudiquen directamente a 150 empleados de la empresa e indirectamente a cientos de personas que forman parte de la familia Curtiduría Tungurahua S.A.

Recomendaciones

- ✚ Se recomienda al departamento de RRHH¹³ coordinar se brinde al personal una inducción sobre el uso de las herramientas tecnológicas que la empresa posee actualmente ya que por inexperiencia, e incluso desconocimiento también se están produciendo estos problemas de no tener información en línea, así como solicitar el mejoramiento de estas tecnologías para la satisfacción de los usuarios.
- ✚ Se recomienda a las coordinaciones mejorar los canales de comunicación entre los usuarios ya que la falta de comunicación causan ciertos desconocimientos sobre procesos que causa pérdidas de recursos para la empresa.
- ✚ Se recomienda a los directivos de Curtiduría Tungurahua aprobar el diseño de un Datawarehouse ya que es la tecnología más apropiada para implementarse en Curtiduría Tungurahua S.A esta tecnología puede satisfacer las necesidades de gestión, análisis y toma de decisiones dentro de la misma, otorgando un panorama de lo que está sucediendo en la empresa y presentando esta información en línea.

¹³ RRHH, Recursos Humanos

CAPÍTULO VI

LA PROPUESTA

Diseño y construcción de un datawarehouse que sea un apoyo gerencial en la toma de decisiones sobre los procesos de producción y ventas de Curtiduría Tungurahua S.A de la ciudad de Ambato.

Datos Informativos

La siguiente propuesta se realiza para la empresa denominada “Curtiduría Tungurahua S.A”, ubicada en el Parque Industrial, Tercera Etapa, calle 8 de la ciudad de Ambato en la provincia de Tungurahua.

Curtiduría Tungurahua S.A, es una empresa líder en la industrialización y comercialización de cuero, posee alrededor de 150 empleados.

Antecedentes de la propuesta

Las entidades del sector privado deben tomar decisiones orientadas a satisfacer las demandas de sus clientes y por ello es muy importante buscar mejorar los sistemas de información ligados a estos procesos de decisión. El presente tema de tesis propone la construcción de un Datawarehouse que servirá de apoyo en el proceso de toma de decisiones de la gerencia y jefaturas de la empresa, decidirá en base a datos históricos y cuadros en línea.

La empresa posee un sistema de gestión de información pero se han convertido en datos almacenados, lo que esta investigación propone es transformar estos datos en información para poder generar reportes que sirva o ayude a las líneas directivas y gerenciales de la empresa a tomar decisiones.

Para implementar este proyecto de tesis se realizaran todos los pasos de un proyecto de Inteligencia de Negocios: diseño y construcción del Datawarehouse y los datamarts, creación y programación de los procesos ETL, creación de los cubos, creación de los informes, finalmente la publicación sobre una plataforma web.

Justificación

Los datos que se han obtenido de la presente investigación han evidenciado esta necesidad tecnológica para solucionar problemas de disponibilidad de información en línea que permita tomar decisiones.

Se estableció anteriormente que el giro industrial de este tipo de empresa curtidoras posee procesos de producción tan complicados y variables, por lo que se requiere una plataforma tecnológica para unir estos procesos manuales con la

automatización y lograr información útil y oportuna.

Un sistema de este tipo permitirá la integración de la información, estrategia que facilitará reducir carga operativa a las líneas directivas, es decir optimizaría uso del personal, mejor atención al cliente, mejorar la calidad del servicio y producto que la empresa oferte, brindando la capacidad al personal realizar sus procesos de forma eficiente, además permitirá conocer el estado actual de la empresa, identificar fallas en los procesos, incluso permitirá auditar los procesos, y otros beneficios más específicos como que la gerencia y departamento de ventas podrán de una manera real conocer su desempeño, con la finalidad de prevenir y corregir errores futuros en la planificación de la producción, advirtiendo un óptimo stock de materia prima y suministros en las bodegas

Objetivos

Objetivo General

Construir un datawarehouse que sea un apoyo gerencial en la toma de decisiones sobre los procesos de producción y ventas de Curtiduría Tungurahua S.A de la ciudad de Ambato.

Objetivos Específicos

✚ Determinar los procesos e información específica que el giro del negocio de la Curtiduría Tungurahua S.A necesite para el correcto funcionamiento del datawarehouse.

✚ Determinar el modelo de datos y diseño de procesos que soportara el datawarehouse.

✚ Obtener un sistema funcional acorde a las especificaciones requeridas por la empresa.

✚ Definir un plan de capacitación a usuario para el acertado manejo del datawarehouse.

Análisis de Factibilidad

Otro aspecto importante de analizar en este momento, tiene relación con la factibilidad del proyecto.

Factibilidad Técnica

La empresa posee a su alcance la tecnología necesaria para el desarrollo del sistema, cuenta con el hardware adecuado, un servidor cuyas características se detalla más adelante en la fase de análisis en evaluación de infraestructura, cabe aclarar que su rendimiento es más que suficiente para alojar el datawarehouse, en cuanto a software, este es tipo propietario, una ventaja importante ya que la empresa posee licencias de sql server 2008, y Microsoft excel que serán las herramientas más explotadas.

Factibilidad Operacional u Organizacional

Lo primero que se especifica es que el proyecto cuenta con la predisposición al cambio por parte de la junta de accionistas y gerencia, pero sobre todo con los medios necesarios para establecer un buen diseño de un datawarehouse con el propósito de mejorar la situación actual de la fábrica, es decir, ya se ha capturado los requerimientos de los usuarios y se ha conocido los sistemas de información de la organización.

Se debe entonces determinar qué grado de relación existe entre estos dos factores, y acordar caminos de solución efectivos, de modo que se asegure contar con fuentes de información consistentes con la naturaleza actual del negocio.

El proyecto de investigación tendrá una **utilidad práctica** ya que se plantea una alternativa de solución al problema investigado.

Factibilidad Económica

El proyecto cuenta ya con la aprobación de la gerencia general, ya que después haber citado que contamos con la tecnología apropiada y operativa, la implementación de la solución no requiere mayor inversión económica, se dispone de todos los recursos tecnológicos e incluso el humano, puesto que la persona que desarrolla la solución ventajosamente trabaja para la empresa, lo que facilita su desarrollo tanto porque conoce a fondo los procesos internos actuales y lo que se requiere como objetivos alcanzar de esta propuesta.

Sin embargo los beneficios que la empresa tendrá con el flujo de información en línea, es relativamente alto ya que la relación costo/beneficio es mínimo/alto.

Fundamentación

Para comprender el contexto en que se desarrollará el presente trabajo de tesis, es importante entender lo que es la Inteligencia de Negocios (Inteligencia de Negocios-BI) y todo lo que ésta implica, puesto que dichos conceptos van de la mano con el trabajo a desarrollarse.

Definiciones generales

Datawarehouse

Es un repositorio de datos de muy fácil acceso, alimentado de numerosas fuentes, transformadas en grupos de información sobre temas específicos de negocios, para permitir nuevas consultas, análisis, reporteador y decisiones. (Pérez & Jiménez, 2011)

Los objetivos fundamentales de un Datawarehouse

Hace que la información de la organización sea accesible: los contenidos del Datawarehouse son entendibles y navegables, y el acceso a ellos son caracterizado por el rápido desempeño. Estos requerimientos no tienen fronteras y tampoco límites fijos. Cuando hablamos de entendible significa, que los niveles de la información sean correctos y obvios. Y navegables significa el reconocer el destino en la pantalla y llegar a donde queramos con solo un clic. Rápido desempeño significa, cero tiempos de espera. Todo lo demás es un compromiso y por consiguiente algo que queremos mejorar.

Hacer que la información de la organización sea consistente: la información de una parte de la organización puede hacerse coincidir con la información de la otra

parte de la organización. Si dos medidas de la organización tienen el mismo nombre, entonces deben significar la misma cosa. Y a la inversa, si dos medidas no significan la misma cosa, entonces son etiquetados diferentes. Información consistente significa, información de alta calidad. Significa que toda la información es contabilizada y completada. Todo lo demás es un compromiso y por consiguiente algo que queremos mejorar.

Es la fundación de la toma de decisiones: el Datawarehouse tiene los datos correctos para soportar la toma de decisiones. Solo hay una salida verdadera del Datawarehouse: las decisiones que son hechas después de que el Datawarehouse haya presentado las evidencias. La original etiqueta que preside el Datawarehouse sigue siendo la mejor descripción de lo que queremos construir: un sistema de soporte a las decisiones. (Pérez & Jiménez, 2011)

Para determinar las posibles alternativas de solución se procede a citar los tipos de Datawarehouse que existen y ver el que es más adecuado para esta propuesta.

Tipos de Datawarehouse

Destacamos los siguientes tipos referenciado por el siguiente artículo (Todo BI, 2007), cuya información es traducida de (Hackathorn & White, 2007)

- ✚ **Native datawarehouse appliance**; donde tanto el hardware como el software están estrechamente integrados en una sola plataforma. No se pueden licenciar por separado y tampoco se pueden utilizar de forma individualizada. Algunos ejemplos de este tipo son: DATAlegro, Netezza, y Teradata.

- ✚ **Software datawarehouse appliance**; en este caso, Bases de datos relacionales, tanto Open Source como comerciales son optimizadas para su uso en entornos Datawarehouse.

Aquí, se puede utilizar en diferentes configuraciones de Hardware.
Algunos ejemplos son Greenplum, Sybase e Ingres.

- + **Packaged datawarehouse appliance**; donde software y hardware comercial es configurado para funcionar como una única plataforma y que es comercializado por un único vendedor. Además, se instala y mantiene en un único sistema. Ejemplos de este tipo son HP Neoview, IBM Balance Warehouse y Sun/Greenplum.
- + **Data management appliance**; que obtiene los datos de forma intensiva desde un servidor. Estos sistemas pueden implicar procesos operacionales, analíticos o de almacenamiento. Algunos ejemplos son ParAccel y Dataupia.

Los elementos básicos de un Datawarehouse

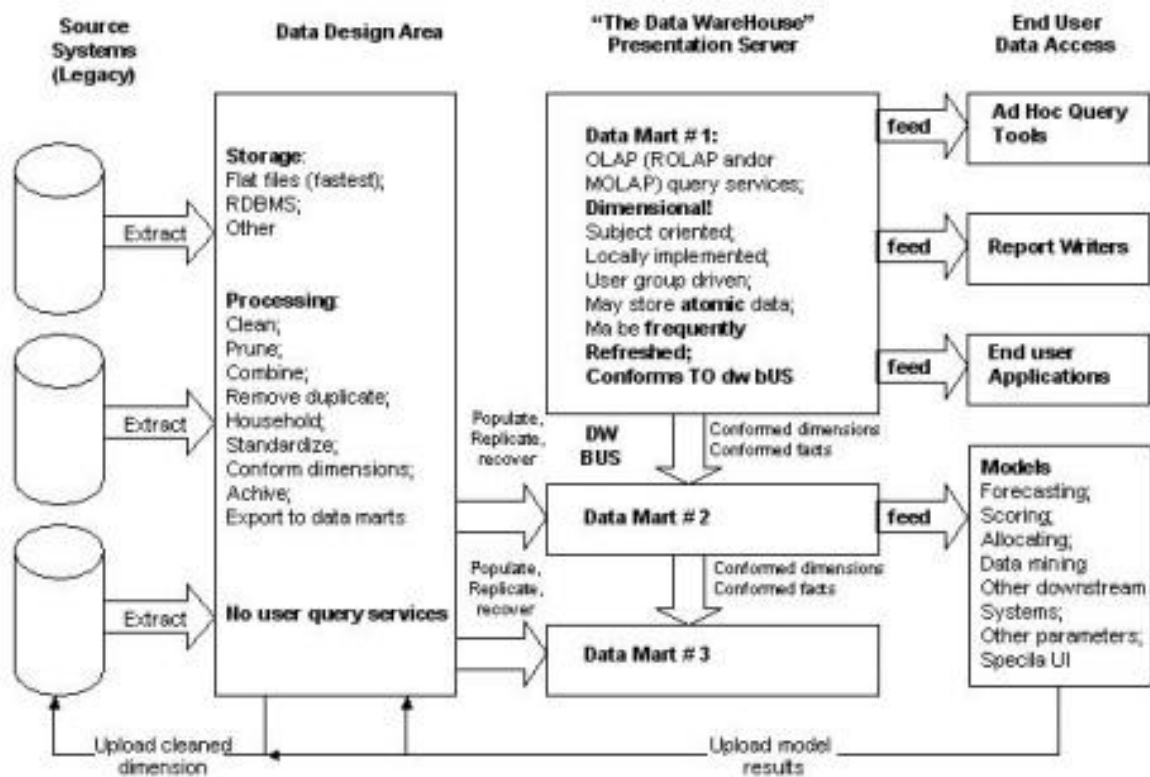


Fig. 6.1: Elementos de un datawarehouse
Fuente: (Pérez & Jiménez, 2011)

Sistema fuente: sistemas operacionales de registros donde sus funciones son capturar las transacciones del negocio. A los sistemas fuentes también se le conoce como Legacy System.

Área de tráfico de datos: es un área de almacenamiento y grupo de procesos, que limpian transforman, combinan, remover los duplicados, guardan, archivan y preparan los datos fuente para ser usados en el Datawarehouse.

Servidor de presentación: la maquina física objetivo en donde los datos del Datawarehouse son organizados y almacenados para consultas sql directos por los usuarios finales, reportes y otras aplicaciones.

Modelo dimensional: una disciplina específica para el modelado de datos que es una alternativa para los modelos de entidad – relación.

Procesos de negocios: un coherente grupo de actividades de negocio que dan sentido a los usuarios del negocio del Datawarehouse.

Datamart: es una versión especial de almacén de datos (datawarehouse). Son subconjuntos de datos con el propósito de ayudar a que un área específica dentro del negocio pueda tomar mejores decisiones. Los datos existentes en este contexto pueden ser agrupados, explorados y propagados de múltiples formas para que diversos grupos de usuarios realicen la explotación de los mismos de la forma más conveniente según sus necesidades.

El Datamart es un sistema orientado a la consulta, en el que se producen procesos batch de carga de datos (altas) con una frecuencia baja y conocida. Es consultado mediante herramientas OLAP (On line Analytical Processing - Procesamiento Analítico en Línea) que ofrecen una visión multidimensional de la información. Sobre estas bases de datos se pueden construir EIS (Executive Information Systems, Sistemas de Información para Directivos) y DSS (Decision Support Systems, Sistemas de Ayuda a la toma de Decisiones). Por otra parte, se conoce como Data Mining al proceso no trivial de análisis de grandes cantidades de datos con el objetivo de extraer información útil, por ejemplo para realizar clasificaciones o predicciones.

En síntesis, se puede decir que los datamarts son pequeños datawarehouse centrados en un tema o un área de negocio específico dentro de una organización.

Datawarehouse: búsquedas fuentes de datos de la empresa. Y es la unión de todos los datamarts que la constituyen.

Almacenamiento operacional de datos: es el punto de integración por los sistemas operacionales. Es el acceso al soporte de decisiones por los ejecutivos.

Olap: En lo que es consultas para la presentación de los datos, es importante explicar el concepto de OLAP (Online Analytical Processing). Según VITT (2002), OLAP proporciona un modelo de datos intuitivo y conceptual, para que los usuarios que no tengan experiencia como analistas puedan comprender y relacionar los datos mostrados. Este modelo es llamado análisis multidimensional, siendo habilitado para ver los datos a través de múltiples filtros, o dimensiones.

Los sistemas OLAP organizan los datos directamente como estructuras multidimensionales, incluyendo herramientas fáciles de usar por usuarios para

conseguir la información en múltiples y simultáneas vistas dimensionales. OLAP es también rápido para el usuario. Rápidos tiempos de respuesta permiten que los gerentes y analistas puedan preguntar y resolver más situaciones en un corto período de tiempo. Una dimensión es una vista de los datos categóricamente consistente.

Una característica de las dimensiones es la habilidad de hacer slice-and-dice. Slice (rebanada) y dice (cubo) hacen particiones de los datos en una base de datos multidimensional de acuerdo a los valores de ciertas dimensiones. Otra capacidad inherente en el diseño de OLAP es la rotación y anidamiento (Pivoting-and-Nesting) de las dimensiones. El pivoted permite rotar los datos desde las columnas hasta las filas. También es importante mencionar el concepto de drill, que en los sistemas OLAP tiene un significado muy específico.

Drill down es la acción de seleccionar un miembro para ver el siguiente nivel inferior de detalle en la jerarquía. Drill up es seleccionar un miembro para ver el siguiente nivel superior, esto es, una acción de bottom-up. La mecánica o funcionamiento de las interfaces OLAP, especialmente pointing-and-clicking (apuntar y seleccionar) para hacer drill-down dentro de las capas de interés se hace posible por la velocidad con que las consultas son resueltas.

Esta funcionalidad permite por completo a los gerentes y analistas un nuevo proceso para tratar con grandes cantidades de datos, un proceso conocido con el nombre de análisis ad hoc. En resumen, los sistemas OLAP organizan los datos por intersecciones multidimensionales. Esta organización, acompañada por una herramienta de interface para rotar y anidar dimensiones, permite a los usuarios visualizar rápidamente valores en detalle, patrones, variaciones y anomalías en los datos que estarían de otra manera ocultos por un análisis dimensional simple. A mayor número de dimensiones (dentro de los límites razonables), mayor es la profundidad del análisis.

MOLAP, ROLAP, HOLAP: Existen variaciones de OLAP según la cantidad de datos y la eficiencia requerida.

OLAP, no se recomienda para consultas complejas y que recorran muchas tablas. Una de estas variaciones es MOLAP (Multidimensional online analytical processing), según VITT (2002), los datos son colocados en estructuras especiales que se encuentran en un servidor central.

MOLAP ofrece el mayor rendimiento de recuperación de información. Por otra parte, existe la solución ROLAP (Relational online analytical processing), según (VITT 2002), permite tomar ventaja de uno de sus más grandes beneficios, el almacenamiento de inmensas cantidades de datos. El rendimiento de recuperación de la información para ROLAP frecuentemente no es tan rápido como otras opciones de almacenamiento.

ROLAP es recomendado para consultas pesadas que no se usan muy a menudo. Finalmente existe HOLAP (Hybrid online analytical processing), que es un híbrido entre MOLAP y ROLAP, y según VITT (2002).

HOLAP no es realmente un modo diferente de almacenamiento de datos. Más bien es la habilidad para diseminar los datos a través de bases de datos relacionales y multidimensionales con la finalidad de obtener lo mejor de ambos sistemas.

Aplicaciones para usuarios finales: una colección de herramientas que hacen los consultas sql, analizan y presentan la información objetivo para el soporte de las necesidades del negocio.

Herramientas de acceso a datos por usuarios finales: un cliente de Datawarehouse.

Ad Hoc Consulta sql Tool: un tipo específico de herramientas de acceso a datos por usuarios finales que invita al usuario a formar sus propias consultas sql manipulando directamente las tablas relacionales y sus uniones.

Modelado de aplicaciones: un sofisticado tipo de cliente de Datawarehouse con capacidades analíticas que transforma o digiere las salidas del Datawarehouse.

Los procesos básicos del Datawarehouse (ETL)

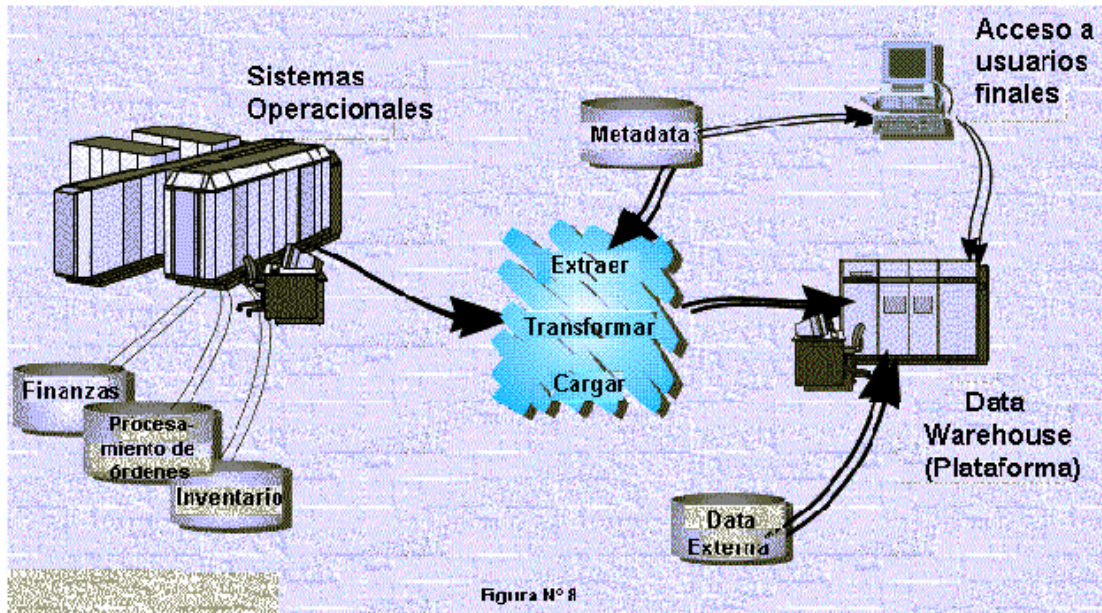


Figura N° 8

Fig. 6.2: Procesos del datawarehouse
Fuente: (D&A & Images, 2011)

Extracción: este es el primer paso de obtener la información hacia el ambiente del Datawarehouse.

Transformación: una vez que la información es extraída hacia el área de tráfico de datos, hay posibles paso de transformación como; limpieza de la información, tirar la basura que no nos sirve, seleccionar únicamente los campos necesarios para el Datawarehouse, combinar fuentes de datos, haciéndolas coincidir por los valores de las llaves, creando nuevas llaves para cada registro de una dimensión.

Carga: al final del proceso de transformación, los datos están en forma para ser cargados.

Evolución del Depósito

Según (INEI), no existe ninguna organización que haya triunfado en el desarrollo del datawarehouse de la empresa, en un sólo paso. Muchas, sin embargo, lo han logrado luego de un desarrollo paso a paso. Los pasos previos evolucionan conjuntamente con la materia que está siendo agregada.

Los datos en el datawarehouse no son volátiles y es un repositorio de datos de sólo lectura (en general). Sin embargo, pueden añadirse nuevos elementos sobre una base regular para que el contenido siga la evolución de los datos en la base de datos fuente, tanto en los contenidos como en el tiempo.

Uno de los desafíos de mantener un datawarehouse, es idear métodos para identificar datos nuevos o modificados en las bases de datos operacionales. Algunas maneras para identificar estos datos incluyen insertar fecha/tiempo en los registros de base de datos y entonces crear copias de registros actualizados y copiar información de los registros de transacción y/o base de datos diarios.

Estos elementos de datos nuevos y/o modificados son extraídos, integrados, transformados y agregados al datawarehouse en pasos periódicos programados. Como se añaden las nuevas ocurrencias de datos, los datos antiguos son eliminados. Por ejemplo, si los detalles de un sujeto particular se mantienen por 5 años, como se agregó la última semana, la semana anterior es eliminada.

Metodologías de un Datawarehouse

El desarrollo de un datawarehouse debe tener en cuenta las necesidades de los usuarios en cuanto a la presentación de informes y análisis. De otro modo, el almacén de datos se convertirá en un cajón de datos del que será difícil extraer información que los usuarios necesitan.

Metodología propuesta por Bill Inmon

Según (Inmon, 1992), tiene un enfoque a modo de explosión en el sentido de que en cierto modo no viene acompañada del ciclo de vida normal de las aplicaciones, sino que los requisitos irán acompañando al proyecto según vaya comprobándose su necesidad.

Esta visión de Inmon puede traer consigo mucho riesgo a la compañía, ya que invierte grandes esfuerzos en el desarrollo del DW y no es hasta la aparición de los DM¹⁴ cuando se empieza a explotar la inversión y obtener beneficios. (Rodríguez Sanz, 2010)

Esta estrategia se contempla en el marco de que es imposible conocer cuáles son las necesidades concretas de información de una empresa, el ambiente dinámico en que se mueve la organización, el cambio de estructura que conlleva el desarrollo de la nueva plataforma y los consiguientes cambios a los sistemas transaccionales que su introducción implica.

Esto hace muy probable que después de la gran inversión en tiempo y recursos en el desarrollo del DWH¹⁵, se haga patente la necesidad de cambios fundamentales que traen consigo altos costos de desarrollo para la organización, poniendo en evidente peligro el éxito de todo el proyecto en sí y que podían ser evitados con una pronta detección en una temprana puesta en explotación de un primer avance del DWH.

Esta metodología para la construcción de un sistema de este tipo es frecuente a la hora de diseñar un sistema de información, utilizando las herramientas habituales como el esquema Entidad/Relación pero al tener un 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”.

Esta es otra de las restricciones que trabajan en contra de la metodología de Inmon ya que implica un consumo de tiempo mayor, teniendo como consecuencia

¹⁴ DM, Datamart

¹⁵ DWH, Datawarehouse

que muchas empresas se inclinan por usar metodologías con las que obtengan resultados tangibles en un espacio menor de tiempo.

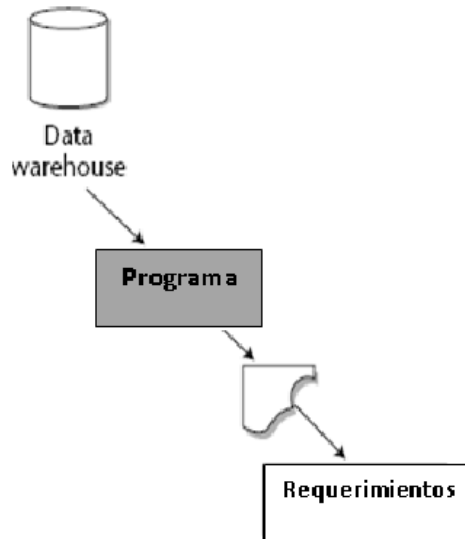


Fig. 6.3: Metodología Inmon
Fuente: (Rodríguez Sanz, 2010)

Metodología propuesta por Raph Kimball

En el año 1998 dicha metodología se recoge como proceso a seguir en el desarrollo de un DWH con el libro: “The Data Warehouse Lifecycle Toolkit”.

La siguiente figura muestra de forma esquemática las fases que componen la metodología propuesta por Kimball y los siguientes apartados resumen el contenido de cada una de las fases.

Ciclo de Vida del Negocio Dimensional

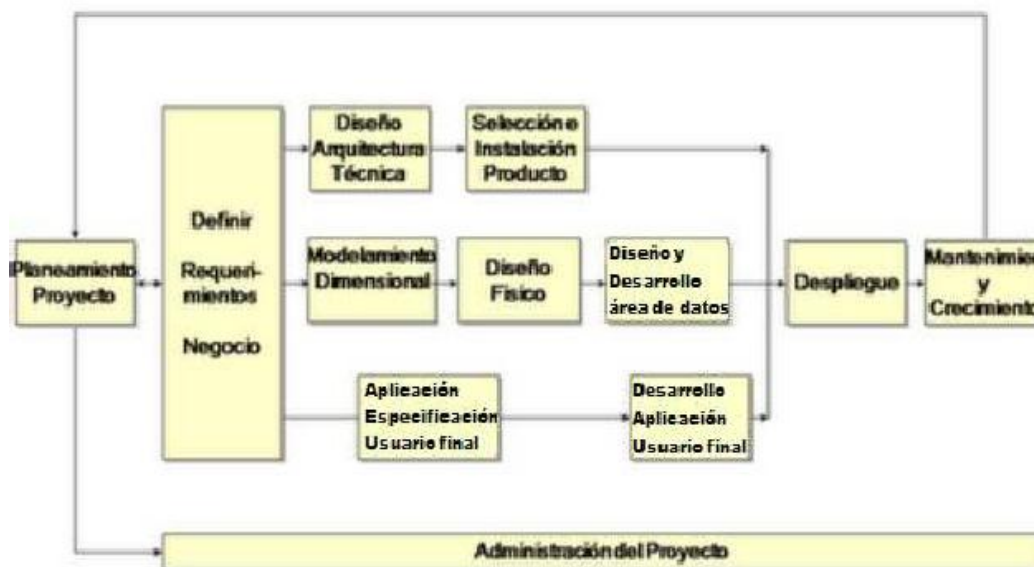


Fig. 6.4: Metodología Kimball
Fuente: (Rodríguez Sanz, 2010)

Planificación del Proyecto

La planificación busca identificar la definición y el alcance del proyecto de DWH, incluyendo las justificaciones del negocio y las evaluaciones de factibilidad.

Definición de los Requerimientos del Negocio

Un factor determinante en el éxito de un proceso de DWH es la interpretación correcta de los diferentes niveles de requerimientos expresados por los distintos grupos de usuarios. La técnica utilizada para revelar los requerimientos de los analistas del negocio difiere de los enfoques tradicionales guiados por los datos. Los diseñadores de los DWH deben entender los factores claves que guían el negocio para determinar efectivamente

Modelado Dimensional

Diseñar los modelos de datos para soportar el análisis de requerimientos demanda un enfoque diferente al usado en los sistemas operacionales. Básicamente, se comienza

con una matriz donde se determina la dimensionalidad de cada indicador y luego se especifica a detalle dentro de cada concepto del negocio, así como la granularidad de cada indicador y las diferentes jerarquías que dan forma al modelo dimensional del negocio o mapa dimensional.

Diseño Físico

El diseño físico de la base de datos se focaliza sobre la selección de las estructuras necesarias para soportar el diseño lógico. Un elemento principal de este proceso es la definición de estándares del entorno de la base de datos. La indexación y las estrategias de particionamiento se determinan también en esta etapa.

Diseño y Desarrollo de la Presentación de Datos

Esta etapa es típicamente la más subestimada de las tareas en un proyecto de DWH. Las principales actividades de esta fase del ciclo de vida son: la extracción, la transformación y la carga (ETL).

Diseño de la Arquitectura Técnica

Los entornos de DWH requieren la integración de numerosas tecnologías. Se deben tener en cuenta tres factores: los requerimientos del negocio, los actuales entornos técnicos y las directrices técnicas y estratégicas futuras planificadas por la compañía para poder establecer el diseño de la arquitectura técnica del entorno de DWH.

Selección de Productos e Instalación

Utilizando el diseño de arquitectura técnica como marco es necesario evaluar y seleccionar los componentes específicos de la arquitectura, como la plataforma de hardware, el motor de base de datos, la herramienta de ETL, las herramientas de acceso, etc. Una vez evaluados y seleccionados los componentes determinados se procede con la instalación y prueba de los mismos en un ambiente integrado de DWH.

Especificación de Aplicaciones para Usuarios Finales

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

Desarrollo de Aplicaciones para Usuarios Finales

A continuación de la especificación de las aplicaciones para usuarios finales, el desarrollo de las aplicaciones de los usuarios finales involucra configuraciones de los metadatos y construcción de reportes específicos.

Implementación

La implementación representa la convergencia de la tecnología, los datos y las aplicaciones de usuarios finales accesibles para el usuario del negocio. Hay varios factores extras que aseguran el correcto funcionamiento de todos estos elementos, entre ellos se encuentran la capacitación, el soporte técnico, la comunicación y las estrategias de realimentación. Todas estas tareas deben tenerse en cuenta antes de que cualquier usuario pueda tener acceso al DWH.

Mantenimiento y crecimiento

Como se remarca siempre, la creación de un DWH es un proceso que acompaña a la evolución de la organización durante toda su historia. Se necesita continuar con las actualizaciones de forma constante para poder seguir la evolución de las metas por conseguir.

Gestión del Proyecto

La gestión del proyecto asegura que las actividades del ciclo de vida se lleven a cabo de manera sincronizada. Como se indica en la figura 6.4, la gestión del proyecto acompaña todo el ciclo de vida.

Metodología propuesta por el Instituto SAS (EEUU)

“Rapid Warehousing Methodology (RWM)” (SAS Institute, 2001)

Como su nombre lo indica es una metodología que compacta todas las áreas tomadas en cuenta por (Kimball,1998), y las plantea en 5 fases:

- ✚ Definicion de objetivos
- ✚ Definicion de requerimientos
- ✚ Diseño y modelizacion
- ✚ Implementacion
- ✚ Implantacion o Gestion del proyecto.

La propuesta se base en esta metodología con una observación que plantea una sola etapa para la definición de objetivos y requerimientos y que en la propuesta está definida como **análisis**.

Metodología

El proceso completo de construcción abarca cuatro grandes etapas identificadas como: **Análisis, Diseño e Implementación, Implantación** cada una de las cuales a su vez incluye una serie de sub-etapas que se propone deberán ejecutarse conforme se señala:

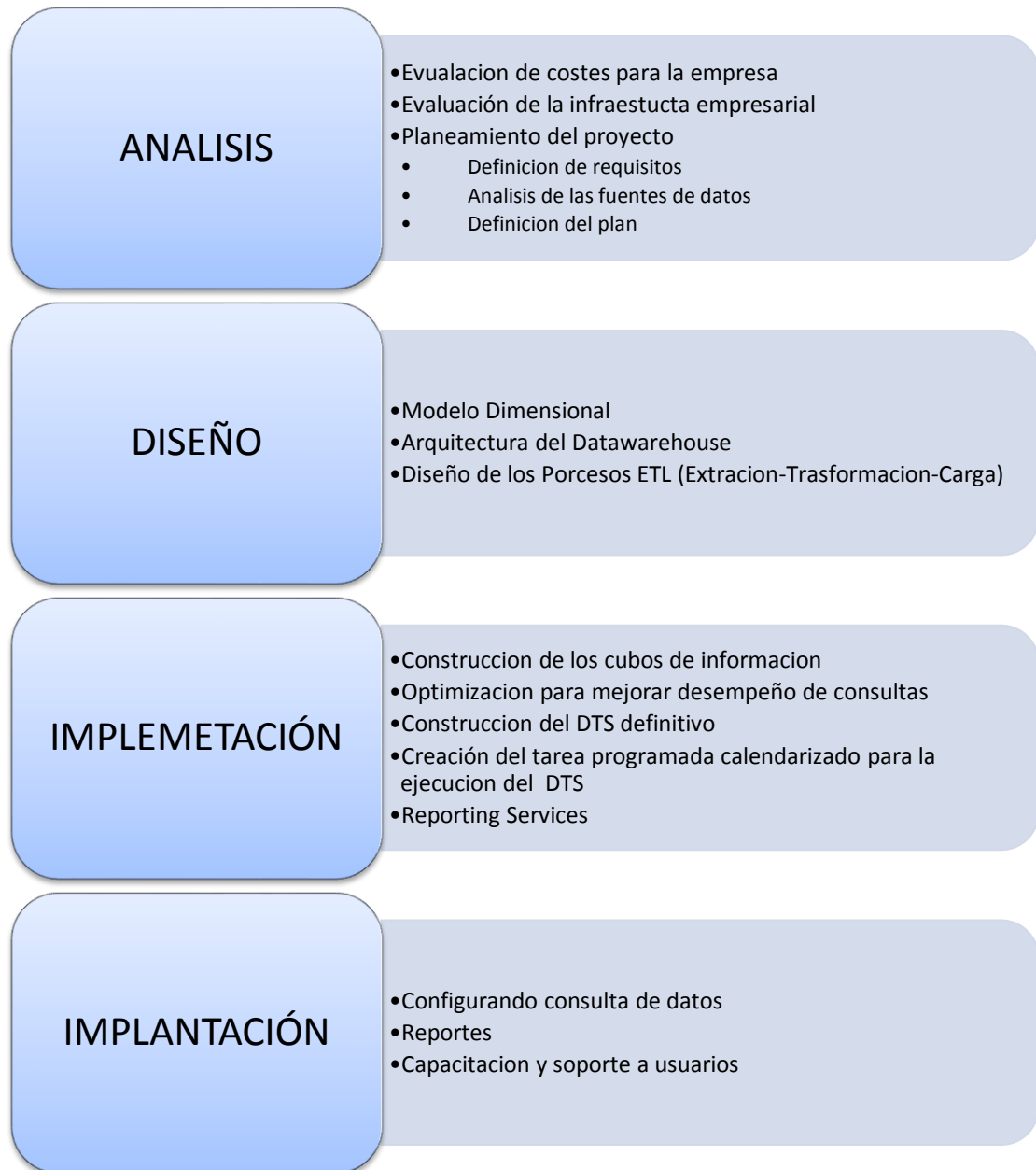


Fig. 6.5: Metodología del datawarehouse
Elaborado por: Ing. Elena Sánchez

Análisis

En esta etapa de la metodología se evalúa a la organización, su infraestructura además de definir la planeación del proyecto, es decir sus requisitos así como las herramientas a utilizar, análisis de datos.

Evaluación de costes en la empresa

El problema del giro del negocio de la empresa está definido, y la solución de un sistema para el apoyo en la toma de decisiones es propuesta.

Para cualquier proyecto es importante realizar un análisis que involucre tantos costos como beneficios. Dado que lograr una cuantificación económica de los factores resulta ser subjetivo. Una alternativa es hacer una valoración desde la perspectiva de los “costos de no disponer en la empresa la información adecuada” sumado a esto los costos de construcción, y de operación según lo clasifica Wol (2002):

De construcción:

Rrhh: Personal de la empresa involucrado.

Tecnología: Nuevas tecnologías cuyo costo es determinado.

Tiempo: El costo de oportunidad asociado a no poseer el sistema.

De operación:

Ya construido e implantado el datawarehouse debe ser mantenido para que sostenga su valor empresarial, estas actividades de soporte son justamente el costo que se relaciona con la evolución, crecimiento y cambio del datawarehouse.

En contraste con el costo tenemos los beneficios del datawarehouse, definidos según Wol (2002):

Mejorar la entrega de información: completa, correcta, consistente, oportuna, accesible.

Mejora el proceso de toma de decisiones: con un mayor soporte de información, se toman decisiones más rápidas, confiables.

Impacto positivo sobre los procesos empresariales: puede llegar a eliminar retardos en los procesos, elimina procesamiento de datos que no son usados, ya que el personal tiene acceso a una mejor calidad de información.

Evaluación de la infraestructura de la empresa

Debido a que le empresa posee ya algunos componentes tecnológicos que soportan la aplicación datawarehouse, se desarrolló el proyecto dentro de este jardín tecnológico para evitar costos adicionales.

Datos técnicos:

Hardware:

- + Servidor Hp proliant 6G,
- + Procesador 1.6 Xeon,
- + Memoria Ram 8Gb,
- + Disco duro 80 GB.

Software:

- + Sistema Operativo - Windows Server 2008
- + Base de datos – Sql server 2008
- + Herramientas (ETL) - DTS Sql Server 2000
- + Repositorios de Metadata – Sql Analysis Services
- + Plataforma Web – Visual Studio 2008- Reporting services.
- + Utilitarios para reportes – Microsoft Excel 2010.

Citamos también el personal definido estructuralmente en la organización:



ORGANIGRAMA ESTRUCTURAL /
ADMINISTRACIÓN

JULIO 2012

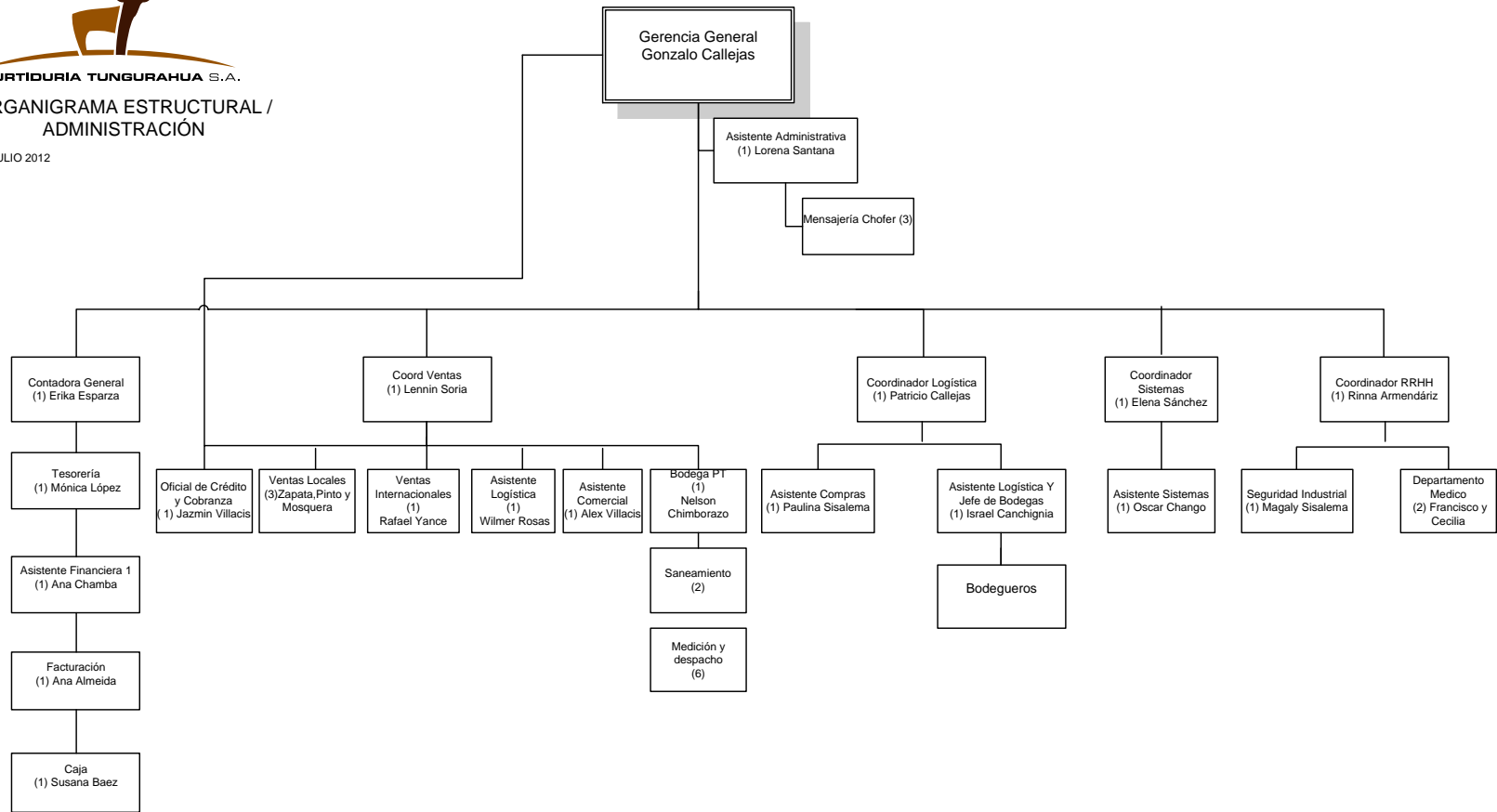


Fig. 6.6: Organigrama Estructural Administración
Fuente: Curtiduría Tungurahua



**ORGANIGRAMA ESTRUCTURAL /
PRODUCCIÓN**

JULIO 2012

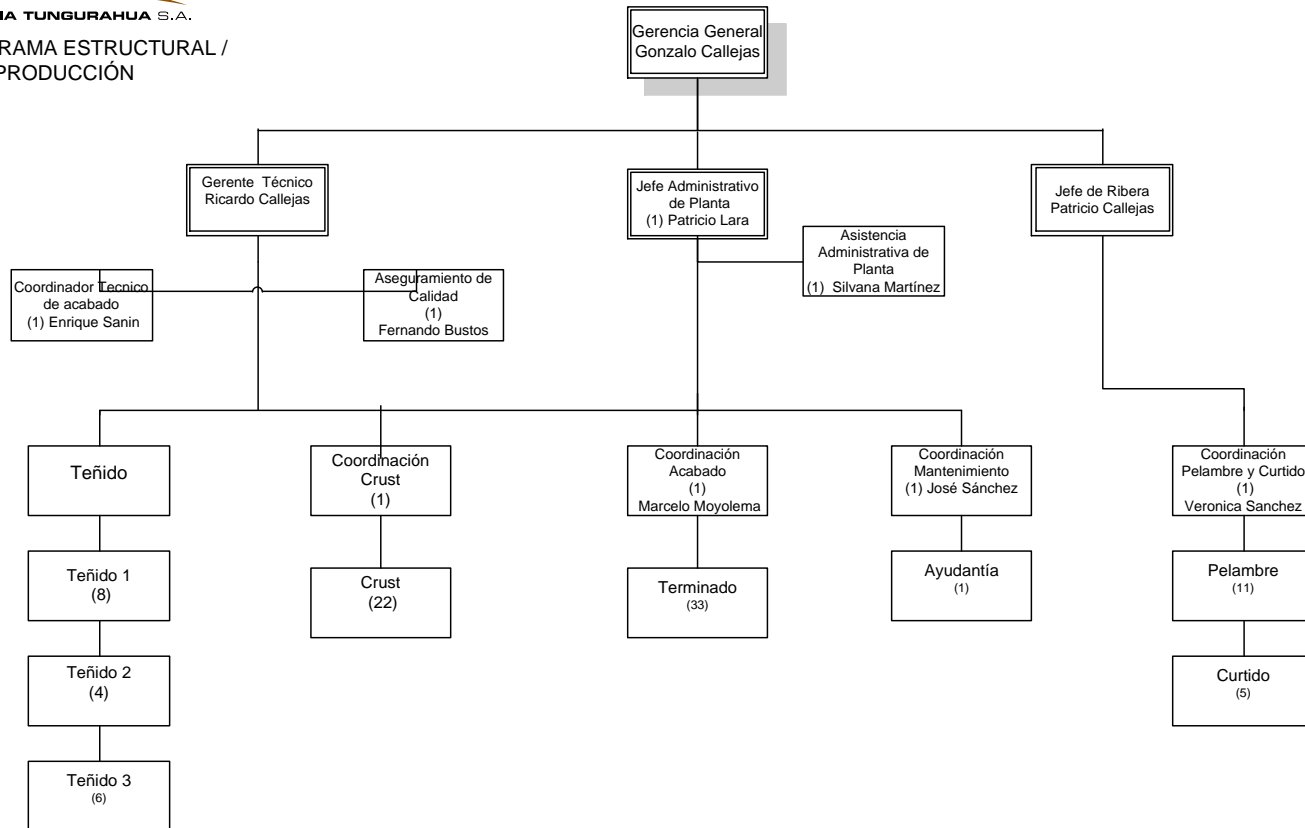


Fig. 6.7: Organigrama Estructural Producción
Fuente: Curtiduría Tungurahua

Planeamiento del proyecto

La clave principal para el éxito del proyecto de la correcta definición de los requerimientos de los usuarios del negocio, puesto que esto es lo que lo afecta en gran medida, describiremos los mismos en el siguiente análisis:

Definición de Requisitos:

Decidir el alcance del proyecto es una de las tareas más difíciles de un sistema de soporte de decisiones, conforme se desarrolló las encuestas y entrevistas con los usuarios de procesos y jefaturas, se definió los siguientes requerimientos, de acuerdo a las posibilidades del cumplimiento:

Del análisis de esta investigación determinamos principalmente que el problema radica en no tener información en línea que apoye en el proceso de toma de decisiones, por lo tanto se requiere la siguiente información:

- ✚ Un histórico de pedidos de materia prima, definidos por centros de costos.
- ✚ Un histórico de productos producidos, vendidos, e incluso reprocesados desde la bodega de producto terminado.
- ✚ Dada la tardanza del actual sistema de gestión, se requiere un inventario a la fecha especialmente para el área de ventas, es decir un stock actual de producto terminado.
- ✚ Un histórico de ventas por cliente, por producto donde se pueda apreciar la cantidad vendida y su área.
- ✚ Un histórico de consumos de materia prima por centro de costo
- ✚ Un histórico de ingresos de materiales a los centros de costos correspondientes.
- ✚ Históricos de los movimientos de préstamos, devoluciones y ventas por centros de costo, sean estos en materia prima como en producto terminado.

Histórico de costes de fórmulas

Análisis de las fuentes de datos:

Validados los requerimientos de los usuarios, verificamos que los datos que describieron en las encuestas y entrevistas existan en la fuente de datos, para convertirla en información.

Auditoria de datos

La fuente es un sistema de gestión desarrollado a medida cuya base de datos está alojado en Oracle 9i.

Detalle de tablas:

<i>Nombre</i>	<i>Descripción</i>	<i>Área</i>
CTNG_T_ALMACENES	Guarda información sobre los Almacenes y las bodegas creadas en la empresa	General
CTNG_T_CENTROS	Aquí se encuentran definidos los centros de costos, especialmente para el proceso de producción	General
CTNG_T_CLIENES	Hace referencia a la ficha principal de los clientes con datos generales.	General
CTNG_T_EMPRESAS	Información principal de	General

	la empresa	
PRO_T_BOD_PEDIDOS	Almacena información de cabecera sobre los pedidos entre bodegas	Pedido materiales
PRO_T_BOD_PEDIDOS_DETALLE	Almacena información detallada sobre los pedidos entre bodegas	Pedido materiales
PRO_T_CAL_MATERIAL	Clasificación de materiales.	Materiales
PRO_T_CALIBRES	Tabla Maestra Calibres, ficha principal con información general	Materiales
PRO_T_COLORES	Tabla Maestra colores, ficha principal con información general	Materiales
PRO_T_MATERIALES	Tabla Maestra de materiales, ficha principal con información general	Materiales
PRO_T_ORD_PRO	Información sobre las ordenes de producción cabecera	Producción
PRO_T_ORD_PRO_DETALLE	Información detallada sobre las ordenes de producción	Producción
PRO_T_ORD_PRO_FORMULACION_APL	Formulación aplicada a una orden de producción por detalle	Producción
PRO_T_PAQU_PIEL	Armada de paquete	PT – Paquetes

	inventario pt	
PRO_T_PAQ_PIEL_DET	Detalle de paquetes en unidades con area.	PT – Paquetes
PRO_T_TRANS_ALM	Cabecera de transferencia por almacén	Inventario
PRO_T_TRANS_ALM_DET ALLES	Detalle de transferencia por almacén	Inventario
PRO_T_VDP_EMP	Cabecera para venta, devolución, préstamo de PT	PT –Inventario
PRO_T_VDP_EMP_DET	Detalle para venta, devolución, préstamo de PT	PT – Inventario
PRO_T_VDP_PAQ_PIEL	Detalle de la venta paquete – piel	PT – Inventario
PRO_T_VDP_TIPO	Tipo de acción en tabla (Venta – Devolución – Préstamo)	PT – Inventario

Tabla 6.1: Detalle de tablas sistema de gestión
Elaborado por: Ing. Elena Sánchez

Después del análisis de los datos hemos determinado que es posible obtener la información solicitada, por lo que los archivos fuentes, serán archivos planos, obtenidos a través de un Tarea programada de Oracle, programado para que diariamente genere las tablas requeridas en archivos csv.

Definición del plan:

1. Extraer de forma adecuada los datos de las fuentes, los que serán archivos planos.
2. Crear un repositorio llamado STAGE para la carga de estos datos a Sql Server.
3. Crear el Datawarehouse DWH (tabla de hechos, procedimientos almacenados para carga y validación de datos) es decir la carga de información relevante para los propósitos requeridos desde el repositorio STAGE
4. Crear un repositorio en base de datos (Sql server 2008) para almacenar los metadatos de los procesos ETL llamado DW1.
5. Crear Datamarts a partir del DWH según los requerimientos
6. Crear procesos ETL para la carga de datos absoluta
7. Automatizar el proceso ETL para cada Datamart.
8. Elaborar los reportes que presenten la información requerida para el apoyo de toma de decisiones, tanto en Microsoft Excel como en Internet Explorer.

Diseño

Esta etapa describe características, arquitectura y modelamiento utilizado.

Cada vista contienen información de una área específica y equivale a un Datamart, se ha optado por utilizar el enfoque de modelo estrella dado que no son muchos datos y posee una alta granularidad en las dimensiones, una de las dimensiones Materiales, deberá ser del tipo Copo de nieve ya que necesitamos la característica de normalización ya que se posee más de un tabla como indicaremos en el modelo dimensional.

En cuanto a la arquitectura el proyecto utiliza la opción MOLAP por las siguientes razones:

- Ofrece el mayor rendimiento de recuperación de datos (cache con el cubo previamente cargado)
- No depende de la cache, si necesita un dato y no lo encuentra, busca en la base de datos relacional (DWH).

Los estándares principalmente utilizados para el diseño de las vistas son:

- dm_nombre_vista: para las dimensiones.
- fc_nombre_vista: para tablas de hechos.
- DM_nombre_datamart: Para los cubos

Modelo dimensional

Esta sección presenta las diversas vistas del modelo.

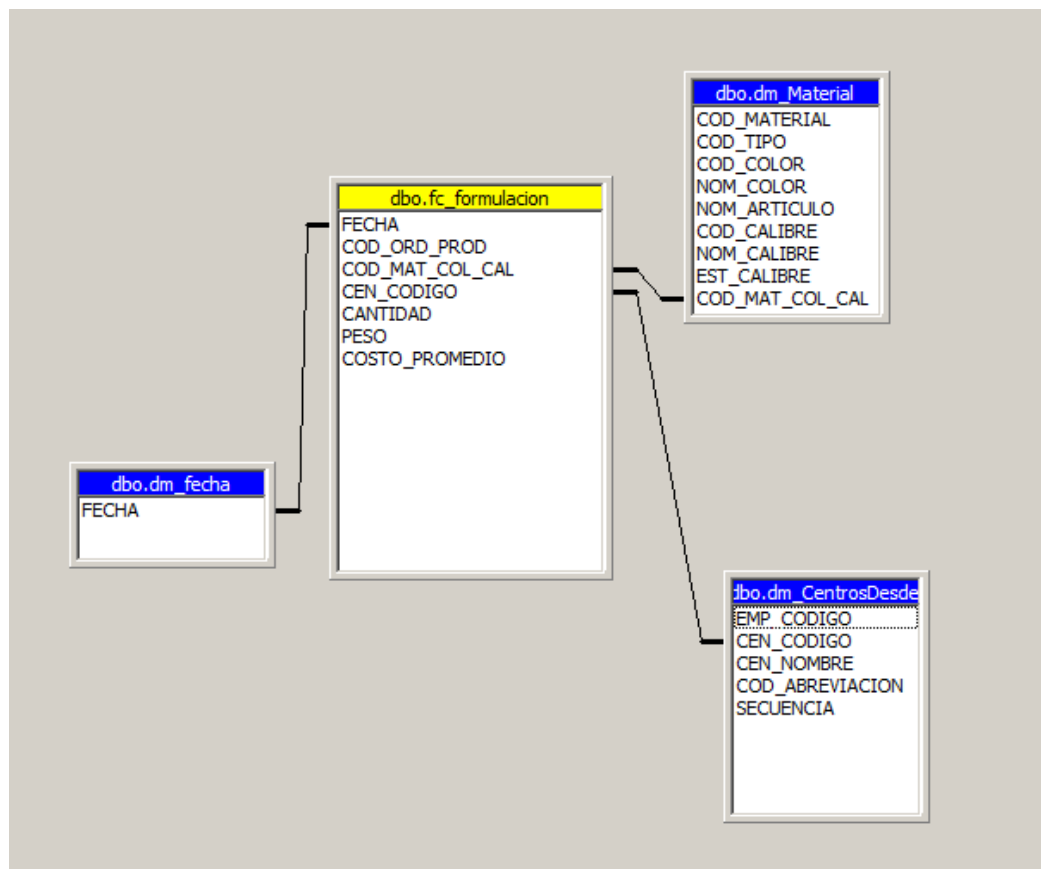


Fig. 6.8: DataMart Formulación

Capturado por: Ing. Elena Sánchez

En el datamart Formulación (DM_Formulación), se dispone de información histórica sobre los costos de haber realizado un producto por órdenes de producción.

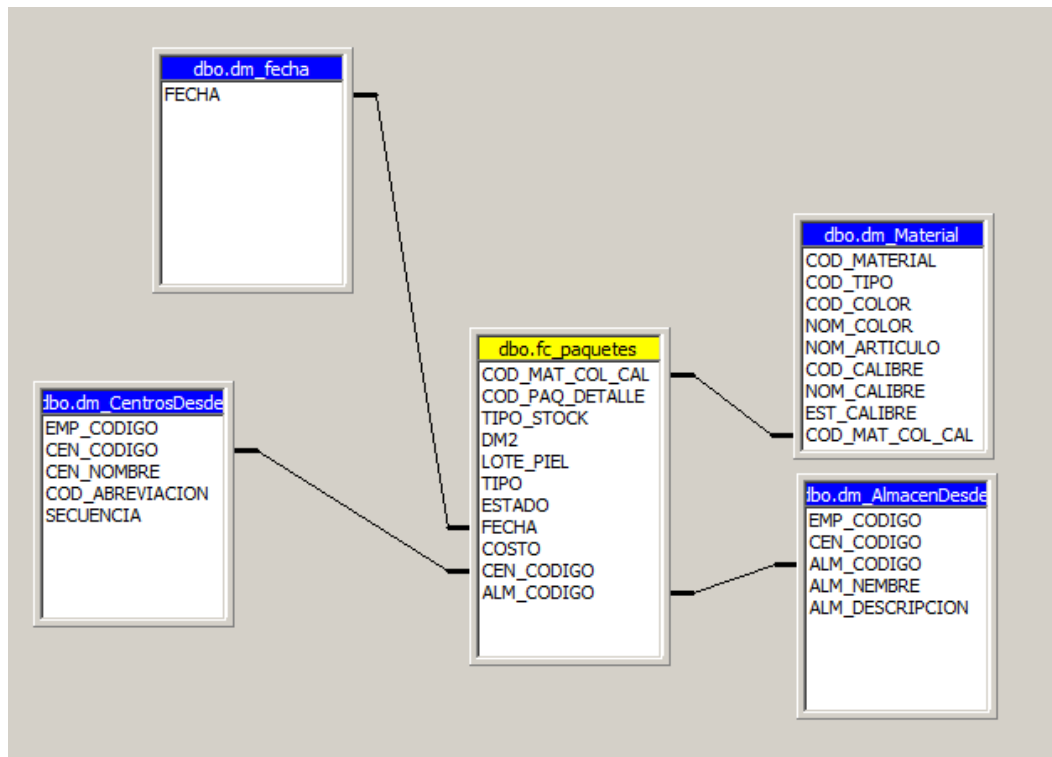


Fig. 6.9: DataMart Paquetes

Capturado por: Ing. Elena Sánchez

El datamart Paquetes (DM_Paquetes), pone a disposición información sobre productos producidos, vendidos, e incluso reprocesados desde la bodega de producto terminado.

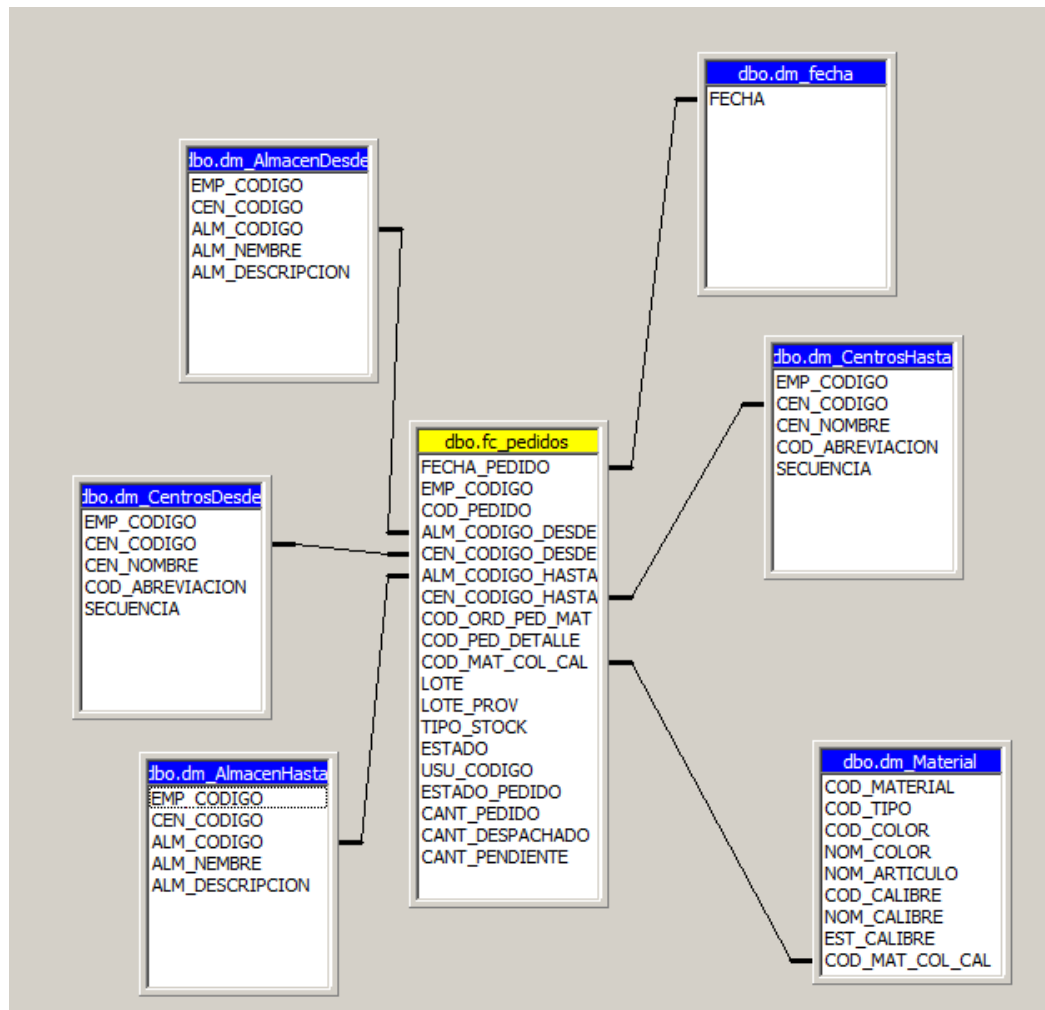


Fig 6.10: DataMart Pedidos

Capturado por: Ing. Elena Sánchez

En el datamart pedidos (DM_pedidos), la informacion que se visualiza es un descriptivo historico de las pedidos de Materia Prima por centros de costos.

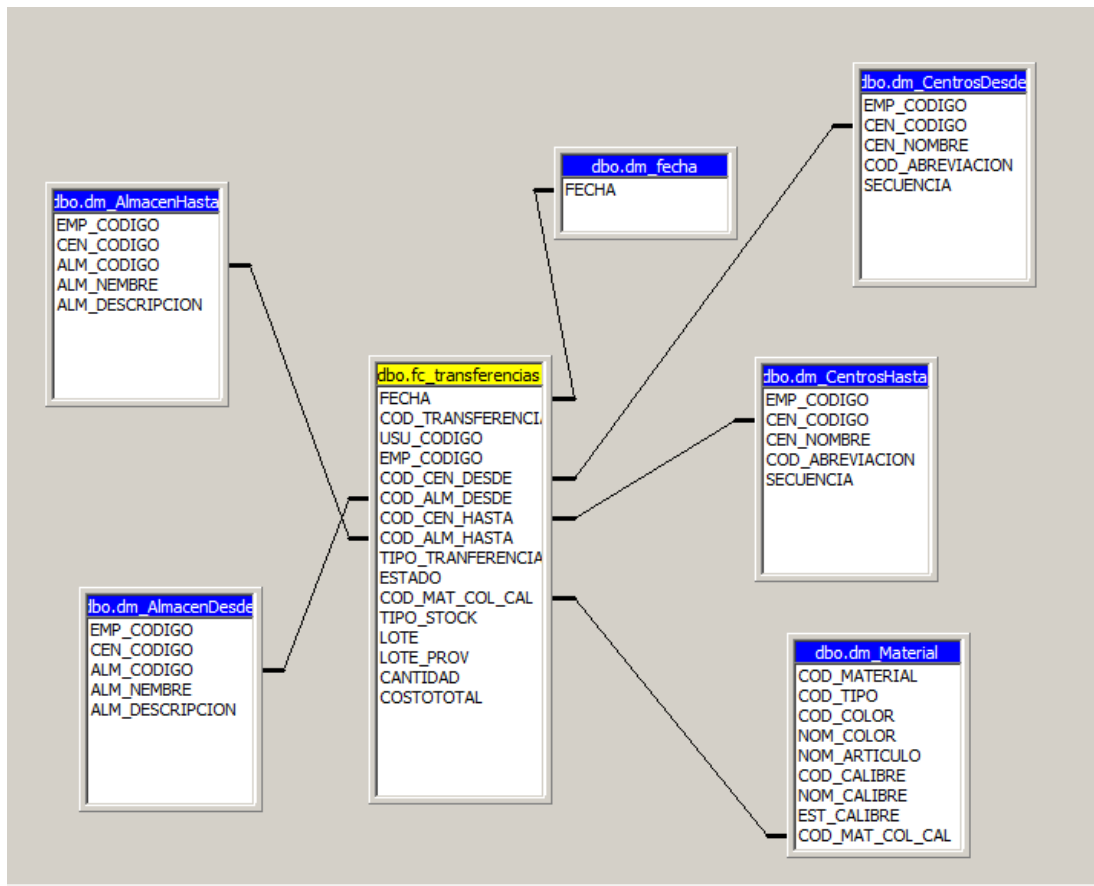


Fig 6.11: DataMart Traslferencias
 Capturado por: Ing. Elena Sánchez

El datamart traslferencias (DM_Traslferencias) contiene los históricos de los movimientos de préstamos, devoluciones y ventas por centros de costo, consumos e ingresos, por centros de costo.

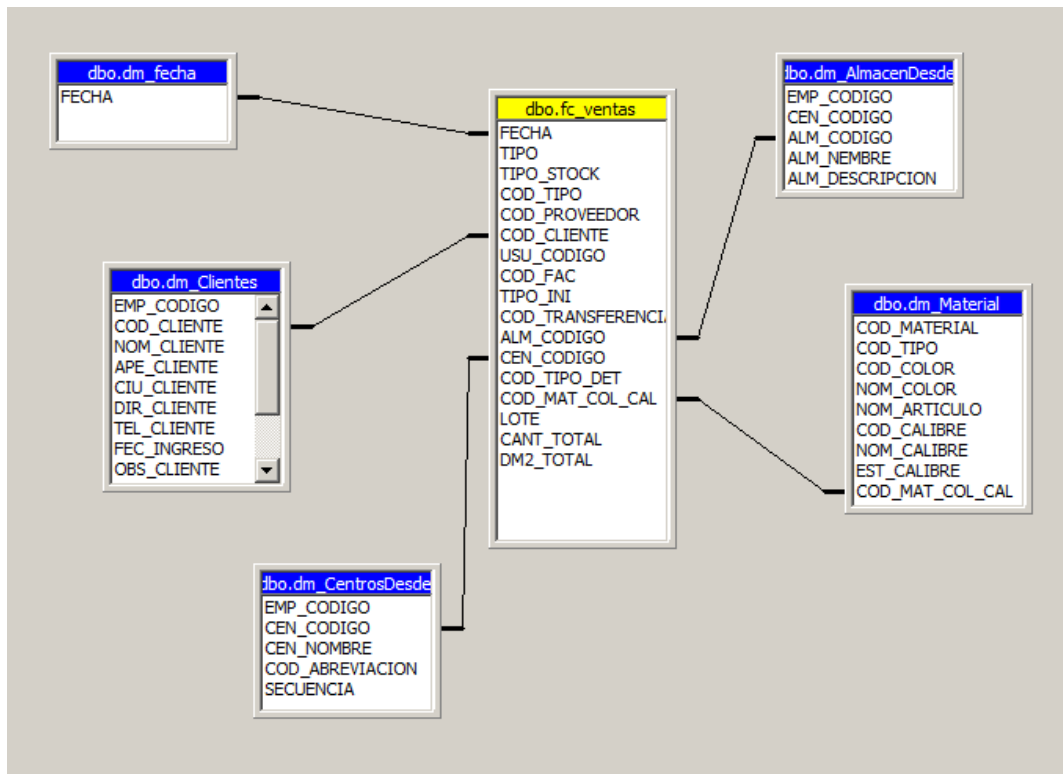


Fig 6.12: DataMart Ventas

Capturado por: Ing. Elena Sánchez

Datamart de Ventas (DM_Ventas) contiene un histórico de ventas por cliente, por producto donde se pueda apreciar la cantidad vendida y su área.

Arquitectura del Datawarehouse

El proyecto está construido en base al siguiente diseño:

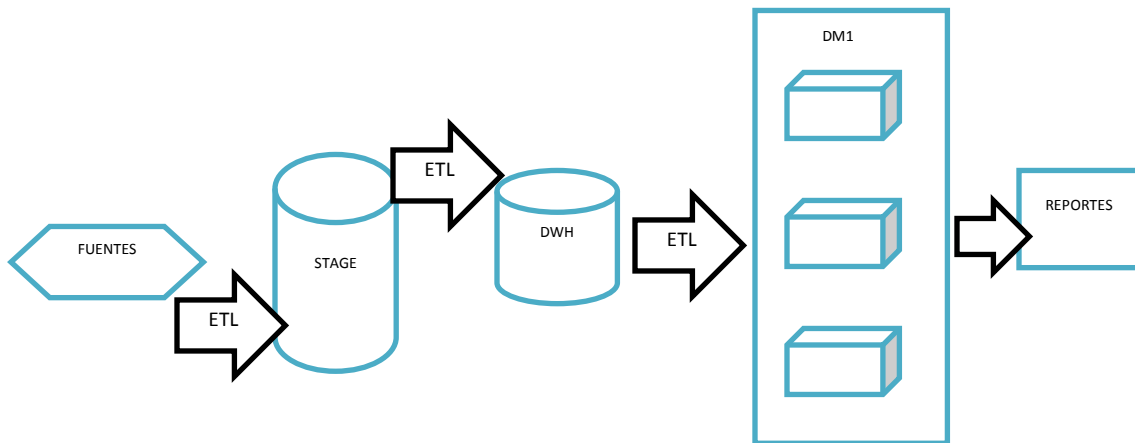


Fig.6.13: Arquitectura del datawarehouse

Elaborado por: Ing. Elena Sánchez

Donde, los Fuentes son los datos de un sistema operacional que captura las transacciones del negocio, desde el stage donde se cargaron los datos a través de etls hacia el dwh, es el área de organización de datos, que extrae y transforma los mismos para pasarlos al servidor de presentación DM1.

El proyecto entonces, define una arquitectura de tres pasos.

- ✚ El sistema fuente
- ✚ Área de organización
- ✚ Servidor de Presentación

Arquitectura que se ajusta a la documentada por (Kimball & Ross, The Data Warehouse Toolkit: The Complete Guide to Dimensional Modeling (2da. Edición), 2002).

Diseño Procesos ETL (Extracción, Transformación y Carga de datos)

Una vez definidos, las tablas de hechos, y las dimensiones que forman el datawarehouse, procedemos a estructurar los procesos que permitan poblar el datawarehouse desde la fuente de datos.

Procedemos con el siguiente diseño:

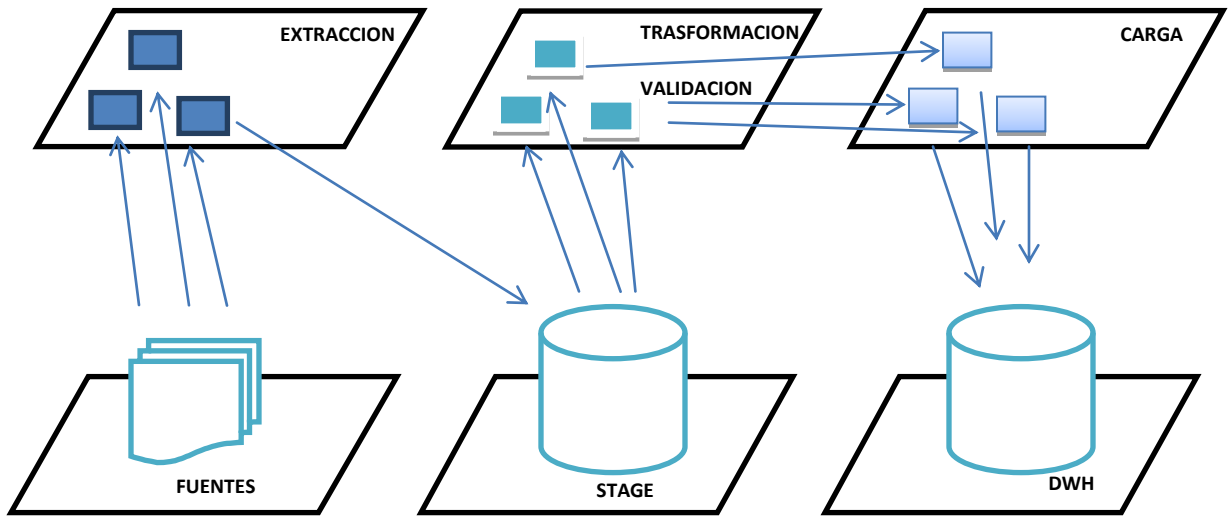


Fig. 6.14: Diagrama de procesos ETL

Elaborado por: Ing. Elena Sánchez

Extracción

Se procede a obtener los datos a través del proceso de extracción para lo cual los archivos fuentes son archivos planos del sistema de Curtiduría Tungurahua S.A que está alojado en una base de datos ORACLE 9i, mismos que son obtenidos a través de un tarea programada en Oracle para que los genere con diariamente a las 12 pm.

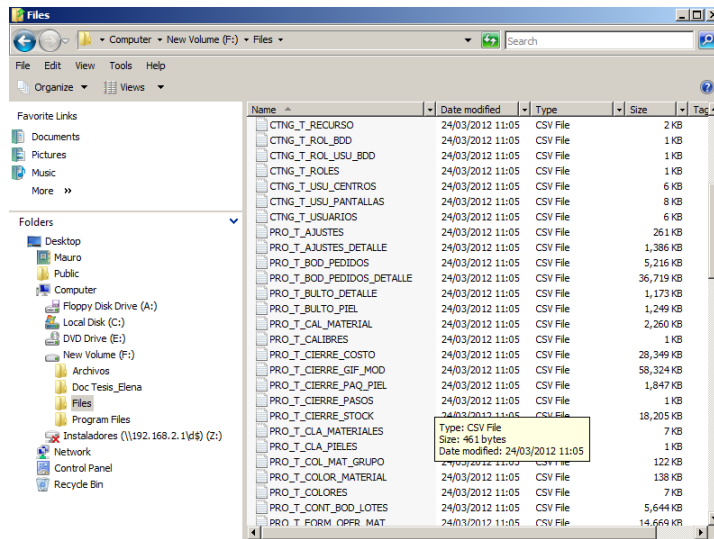


Fig. 6.15: Archivos planos de Oracle
Capturado por: Ing. Elena Sánchez

Extracción de los datos del sistema transaccional desde el SQL Server Enterprise manager.

1. Para lo cual ingresamos SQL Server Enterprise manager y procedemos a Crear un nuevo paquete.

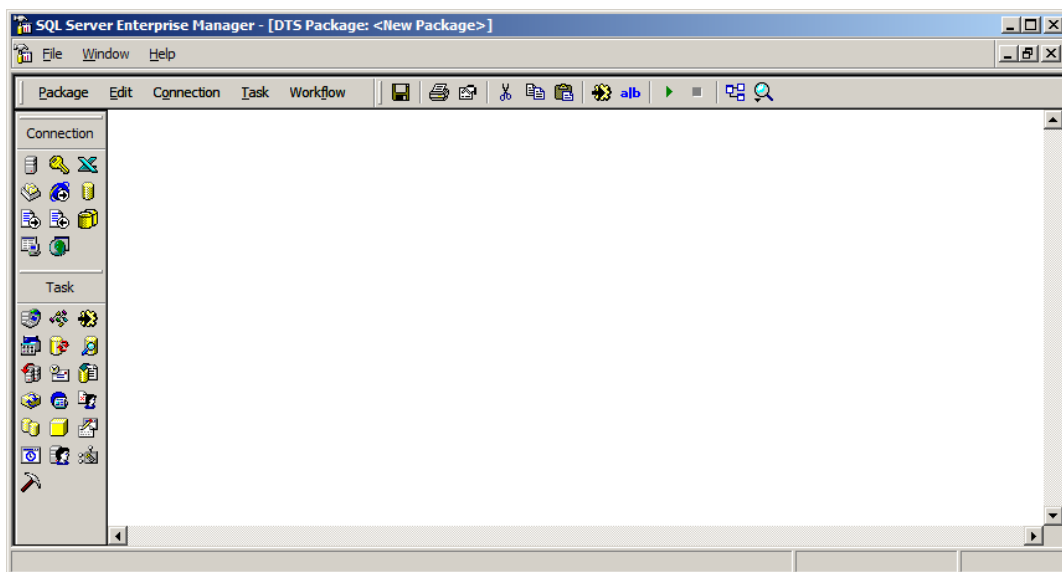


Fig. 6.16: Creación de un paquete
Capturado por: Ing. Elena Sánchez

2. Creamos los enlaces a los archivos planos, a través de una herramienta denominada Text File (Source).

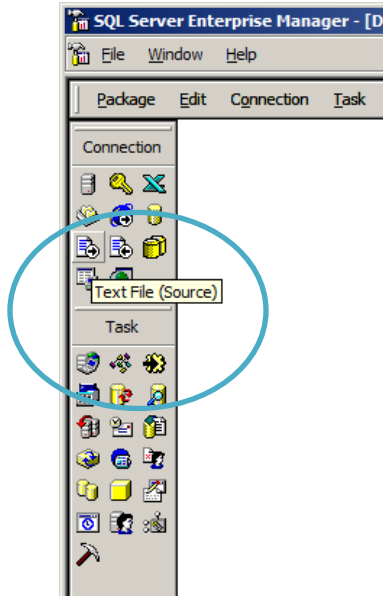


Fig. 6.17: Creación de enlaces a archivos planos
Capturado por: Ing. Elena Sánchez

Una vez creado el enlace nos pedirá realizar la conexión para lo que genera la siguiente imagen, donde en file name: se debe apuntar al directorio que contiene el archivo plano.

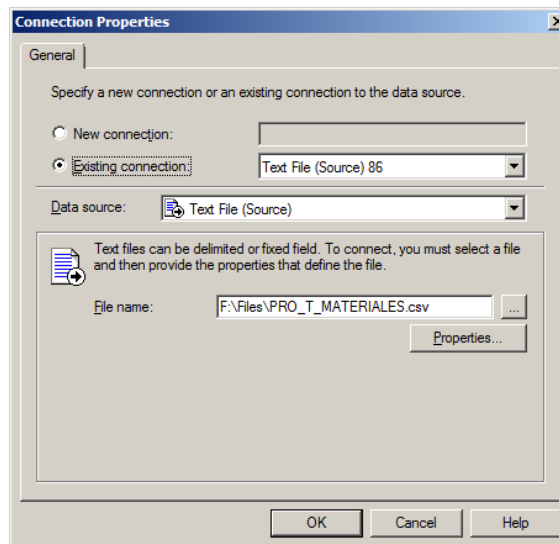


Fig. 6.18: Conexión archivos planos
Capturado por: Ing. Elena Sánchez

- Configuración de separadores de columnas, si el archivo plano fue configurado como separador de columnas, con punto y como o espacios, la misma configuración debe ser configurada en la siguiente ventana:

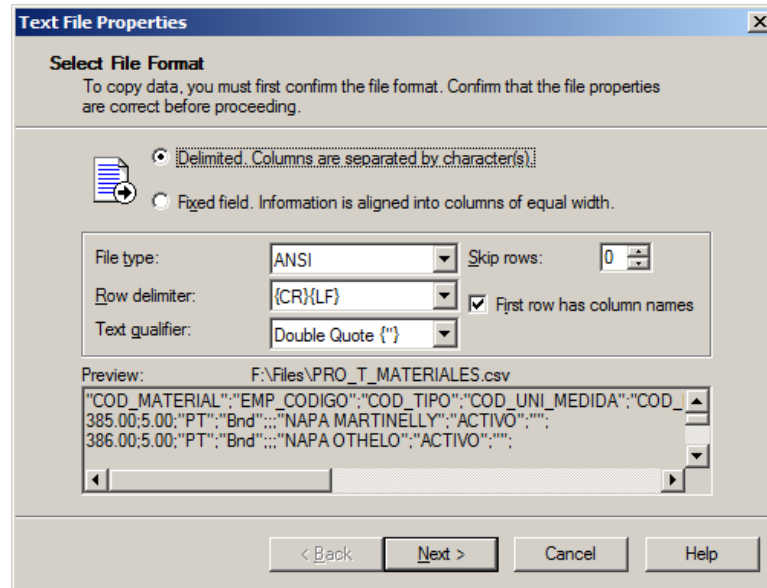


Fig. 6.19: Configuración de separadores de columna

Capturado por: Ing. Elena Sánchez

- Revisión de texto en columnas, aquí se procede a delimitar las columnas con sus respectivos campos

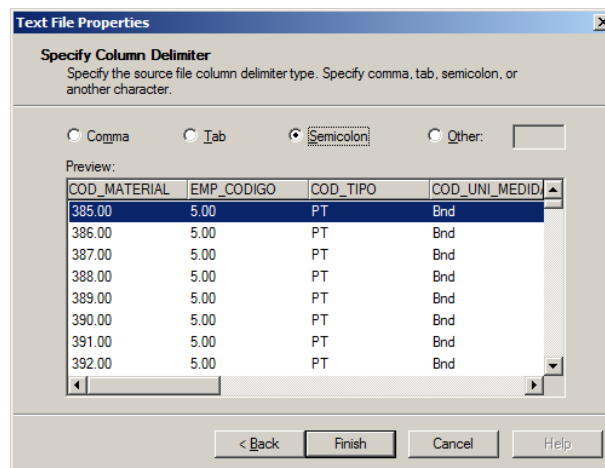


Fig. 6.20: Delimitación de columnas

Capturado por: Ing. Elena Sánchez

5. Configuración del servidor de destino, se escoge la herramienta Microsoft OLE DB Provider for Sql Server para definir las propiedades del destino donde se subirán las fuentes, se define la base de datos, ubicación y tipo de autenticación:

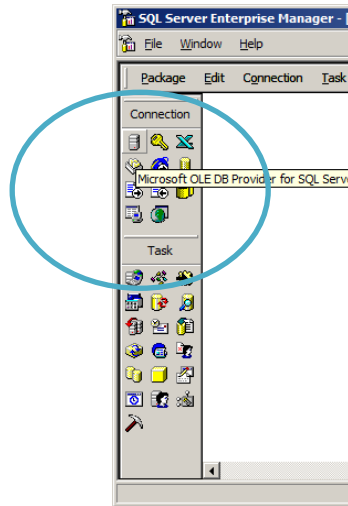


Fig. 6.21: Herramienta para configurar destino

Capturado por: Ing. Elena Sánchez

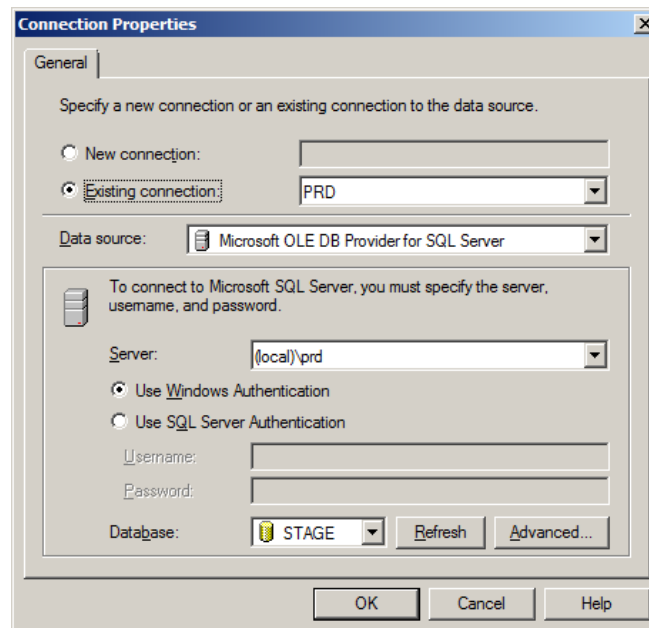


Fig. 6.22: Configuración del servidor destino

Capturado por: Ing. Elena Sánchez

6. Diagrama final de la extracción

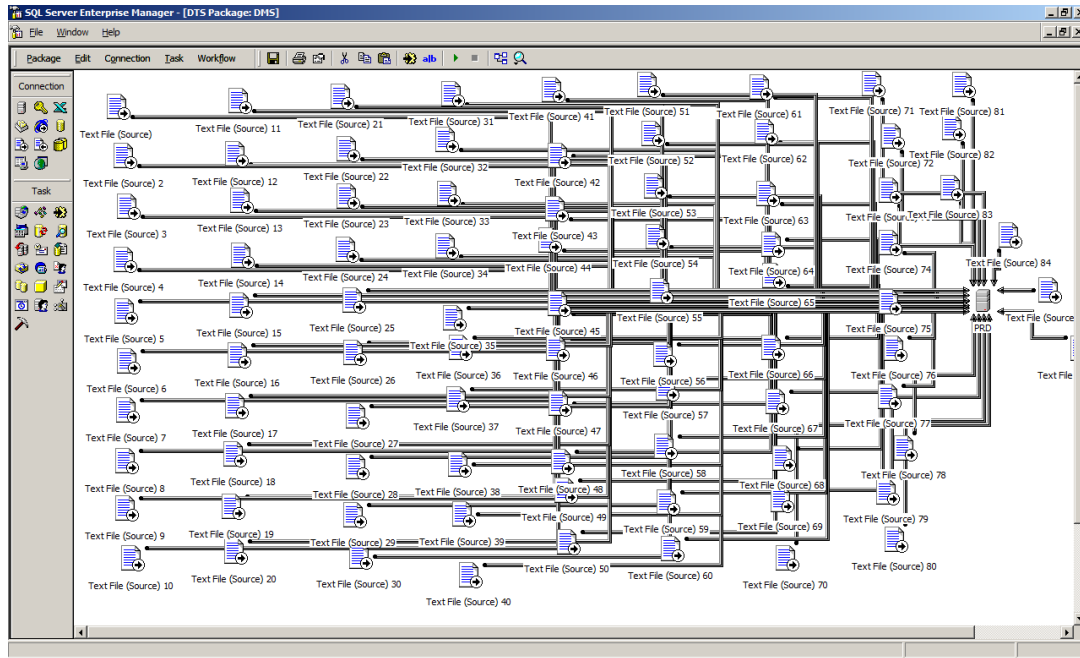
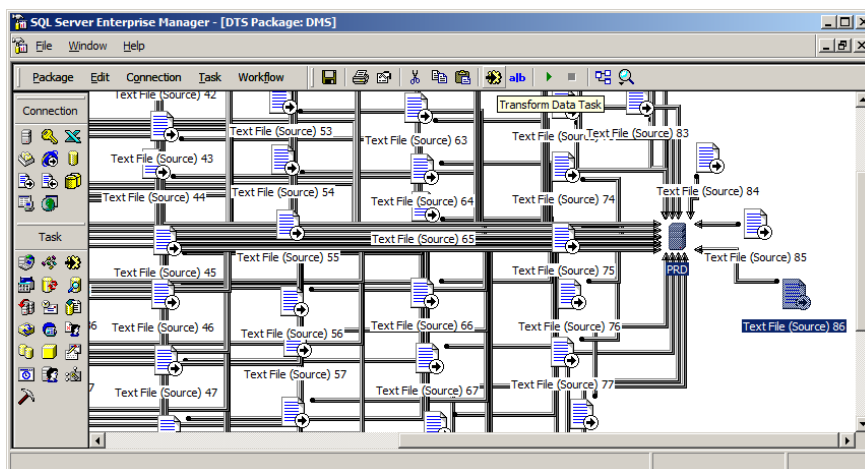


Fig. 6.23: Diagrama Final de extracción

Capturado por: Ing. Elena Sánchez

7. Se utiliza la herramienta text file (source) para definir el proceso de migración, desde las fuentes hasta la base de datos sql server en el server llamado PRD¹⁶



¹⁶ PRD, Servidor que almacena las bases de datos en sql server para almacenar los fuentes

Donde se define las características de la migración, el origen: si es una tabla o vista, la ubicación de la misma.

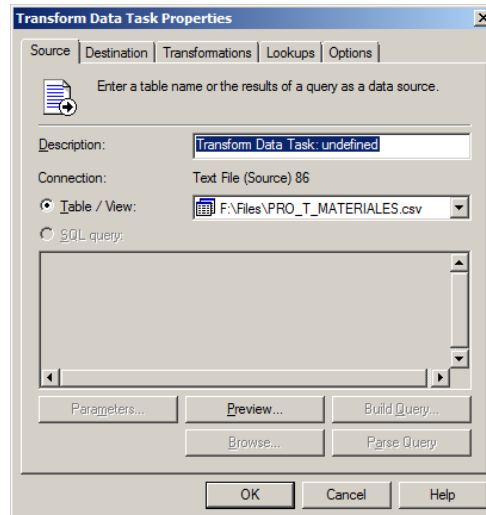


Fig. 6.24: Propiedades de transformación origen.

Capturado por: Ing. Elena Sánchez

Se define también el servidor PRD destino en la base de datos STAGE la tabla que le corresponde:

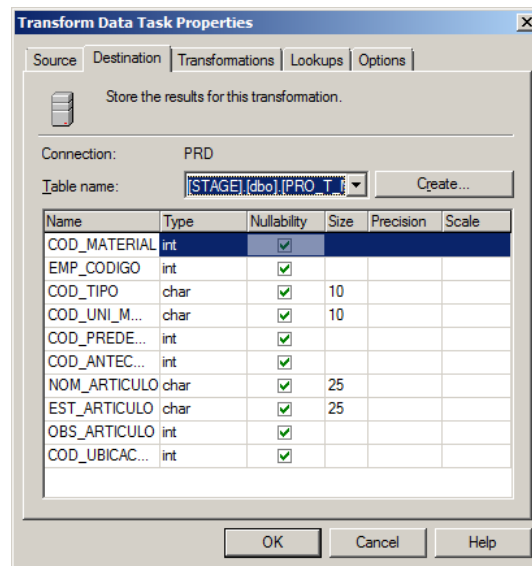


Fig. 6.25: Propiedades de transformación destino.

Capturado por: Ing. Elena Sánchez

Se define la correspondencia de campos entre el origen y el destino, igualando cada columna de la tabla de origen, con la columna correspondiente en la tabla de destino en la base STAGE, en este caso los campos poseen los mismos nombres, pero dependiendo del diseño el nombre de las columnas pueden variar.

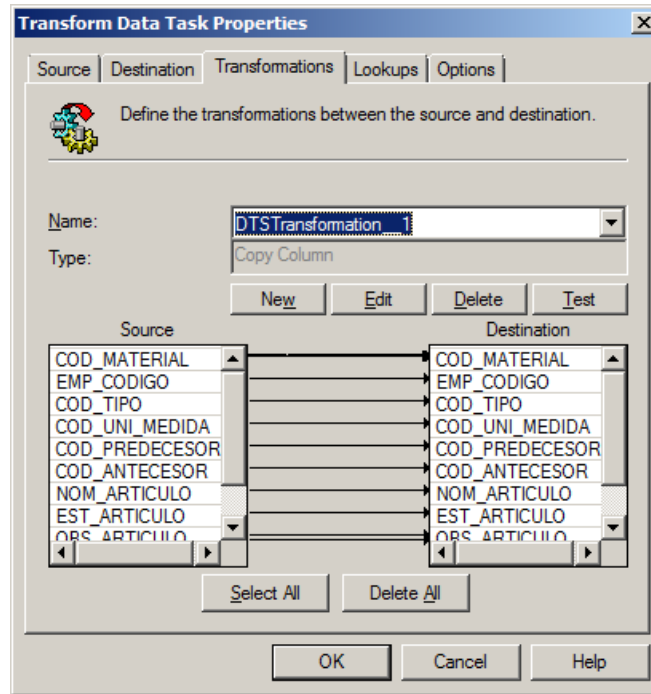


Fig. 6.26: Propiedades de transformación.

Capturado por: Ing. Elena Sánchez

Transformación y Carga de Datos

Se procede a realizar el proceso de transformar oportunamente los datos que son requeridos, a través de procedimientos almacenados en la una nueva base de datos denominada DWH¹⁷, además se procede con el siguiente proceso que la carga de datos ya procesados una vez terminada la operación de transformación en una base de datos creada especialmente para esto denominada STAGE¹⁸ para los cual se procede a:

1. Creación de la base de datos repositorio de paso

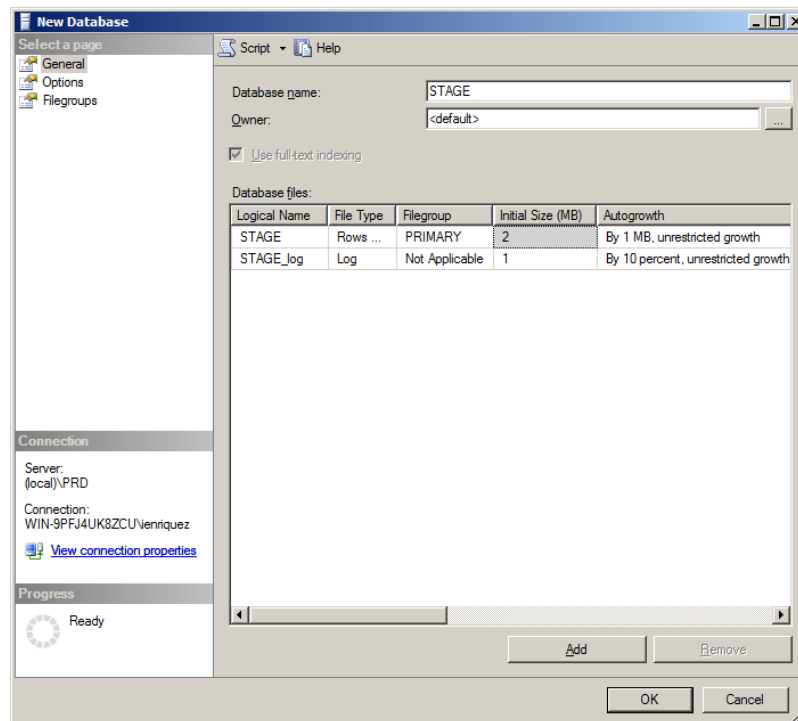


Fig. 6.27: Creación de BDD Stage.

Capturado por: Ing. Elena Sánchez

¹⁷ DWH, base de datos donde se cargarán los datos procesados después de las operaciones de transformación

¹⁸ STAGE, Base de datos para la transformación y carga de datos

2. Creación de la base de datos para el Datawarehouse

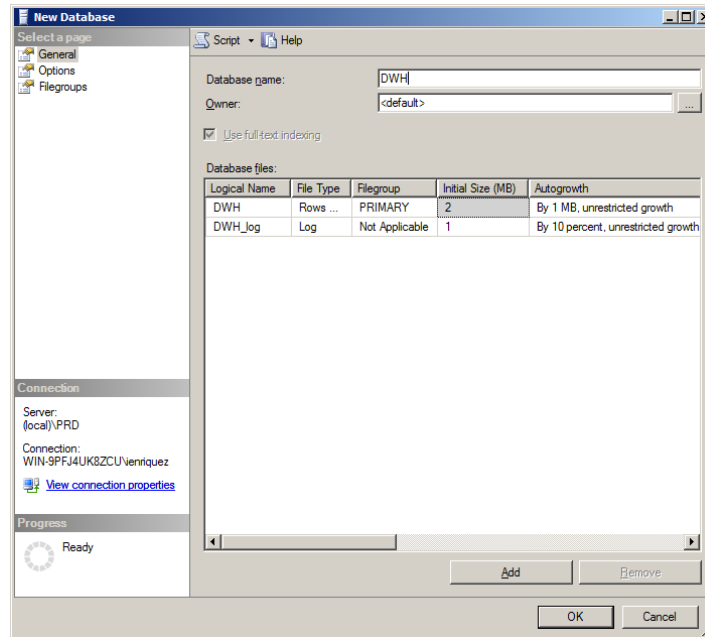


Fig. 6.28: Creación de BDD Dwh.

Capturado por: Ing. Elena Sánchez

Además se procede a crear la base de datos que almacenara la información ya dimensionada, para que las consultas sql se puedan realizar incluso directamente sobre esta base a la que se denominará DW1¹⁹

3. Creación de la base de datos resumen donde se alojara los cubos de información.

¹⁹ DW1, Base de datos

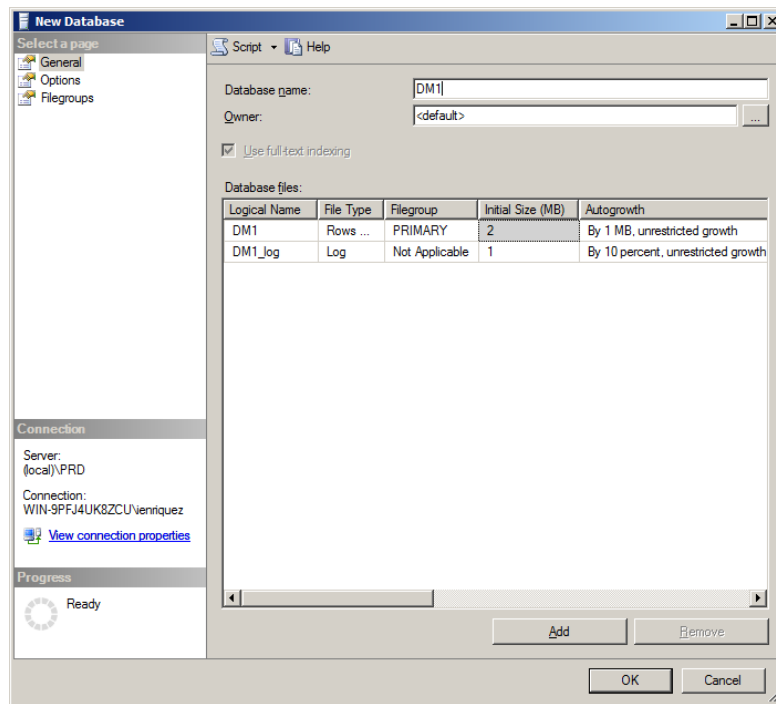


Fig. 6.29: Creación de BDD DW1.

Capturado por: Ing. Elena Sánchez

- Creación de procedimientos de transformación y carga de datos, como se definió anteriormente están creados sobre la base DWH

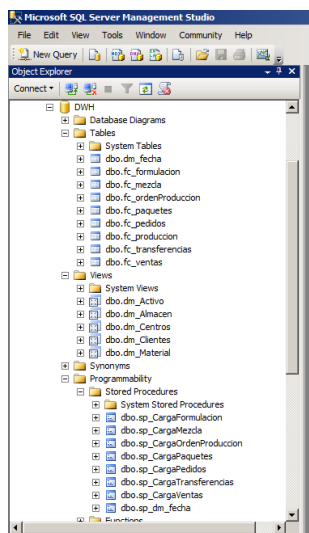


Fig. 6.30: Listado de procedimientos ETL

Capturado por: Ing. Elena Sánchez

```

Microsoft SQL Server Management Studio
File Edit View Query Project Debug Tools Window Community Help
New Query DWH Execute
SQLQuery6.sql - (local)\... (56) | SQLQuery4.sql - (local)\... (56)
USE [DWH]
GO
/***** Object: StoredProcedure [dbo].[sp_CargaTransferencias] Script Date: 04/05/2012 01:16:30 *****/
SET ANSI_NULLS ON
GO
SET QUOTED_IDENTIFIER ON
GO
ALTER PROCEDURE [dbo].[sp_CargaTransferencias] as
truncate table fc_transferencias
insert into fc_transferencias
select convert(char,FECHA,112) FECHA,
t1.COD_TRANSFERENCIA,
USU_CODIGO,
t1.EMP_CODIGO,
COD_CEN_DESDE,
COD_ALM_DESDE,
COD_CEN_HASTA,
COD_ALM_HASTA,
TIPO_TRANSFERENCIA,
ESTADO,
COD_PEDIDO,
COD_ORD_PED_MAT,
COD_MAT_COL_CAL,
TIPO_STOCK,
LOTE,
LOTE_PROV,
CANTIDAD,
COSTO
from STAGE..PRO_T_TRANS_ALM t1 join STAGE..PRO_T_TRANS_ALM_DETALLES t2 on t1.COD_TRANSFERENCIA = t2.COD_TRANSFERENCIA

```

Connected: (1/1) (local)PRD (10.0 SP2) WIN-9FF34JK3ZCU\jennq... DWH 00:00:00 0 rows
Ready Ln 39 Col 1 Ch 1 IN6

Fig. 6.31: Procedimiento almacenado para cargar Transferencias.

Capturado por: Ing. Elena Sánchez

```

Microsoft SQL Server Management Studio
File Edit View Query Project Debug Tools Window Community Help
New Query DWH Execute
SQLQuery7.sql - (local)\... (59) | SQLQuery6.sql - (local)\... (58) | SQLQuery4.sql - (local)\... (56)
USE [DWH]
GO
/***** Object: StoredProcedure [dbo].[sp_CargaVentas] Script Date: 04/05/2012 01:17:21 *****/
SET ANSI_NULLS ON
GO
SET QUOTED_IDENTIFIER ON
GO
ALTER procedure [dbo].[sp_CargaVentas] as
truncate table fc_ventas
insert INTO fc_ventas
select convert(char,FECHA,112),
t1.TIPO,
TIPO_STOCK,
t1.COD_TIPO,
t1.EMP_CODIGO,
COD PROVEEDOR,
COD_CLIENTE,
USU_CODIGO,
COD_FAC,
TIPO_INI,
COD_TIPO_INI,
COD_TRANSFERENCIA,
ALM_CODIGO,
CEN_CODIGO,
COD_ALTERNO,
COD_TIPO_DET,
COD_MAT_COL_CAL,
LOTE,
CANT_TOTAL,
DM2_TOTAL

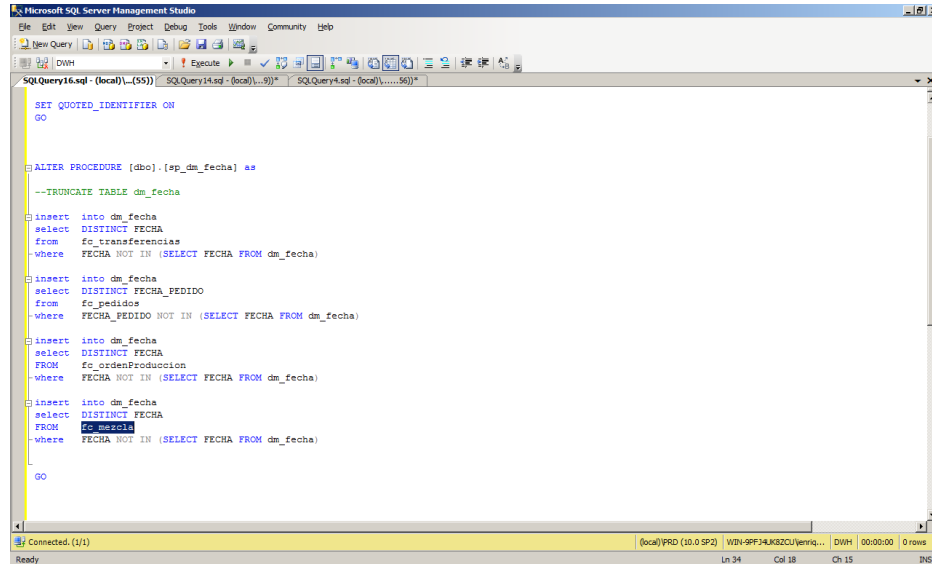
```

Connected: (1/1) (local)PRD (10.0 SP2) WIN-9FF34JK3ZCU\jennq... DWH 00:00:00 0 rows
Ready Ln 42 Col 1 Ch 1 IN6

Fig. 6.32: Procedimiento almacenado para cargar Ventas

Capturado por: Ing. Elena Sánchez

5. Procedimiento para carga de dimensiones es decir las entidades por la que la empresa desea organizar su información



```
SET QUOTED_IDENTIFIER ON
GO

ALTER PROCEDURE [dbo].[sp_dm_fecha] as
--TRUNCATE TABLE dm_fecha

insert into dm_fecha
select DISTINCT FECHA
from fc_transferencias
where FECHA NOT IN (SELECT FECHA FROM dm_fecha)

insert into dm_fecha
select DISTINCT FECHA_PEDIDO
from fc_pedidos
where FECHA_PEDIDO NOT IN (SELECT FECHA FROM dm_fecha)

insert into dm_fecha
select DISTINCT FECHA
FROM fc_ordenProduccion
where FECHA NOT IN (SELECT FECHA FROM dm_fecha)

insert into dm_fecha
select DISTINCT FECHA
FROM fc_ordenCaja
where FECHA NOT IN (SELECT FECHA FROM dm_fecha)

GO
```

Fig. 6.33: Procedimiento almacenado para cargar dimensión fecha

Capturado por: Ing. Elena Sánchez

La dimensión de material, es una dimensión que puede ser creada como un modelo copo de nieve, se ha realizado otro tipo de esquema que mejora el manteniendo del DWH, des normalizando el modelo entidad relación con el fin de tener una sola tabla con la jerarquía requerida.

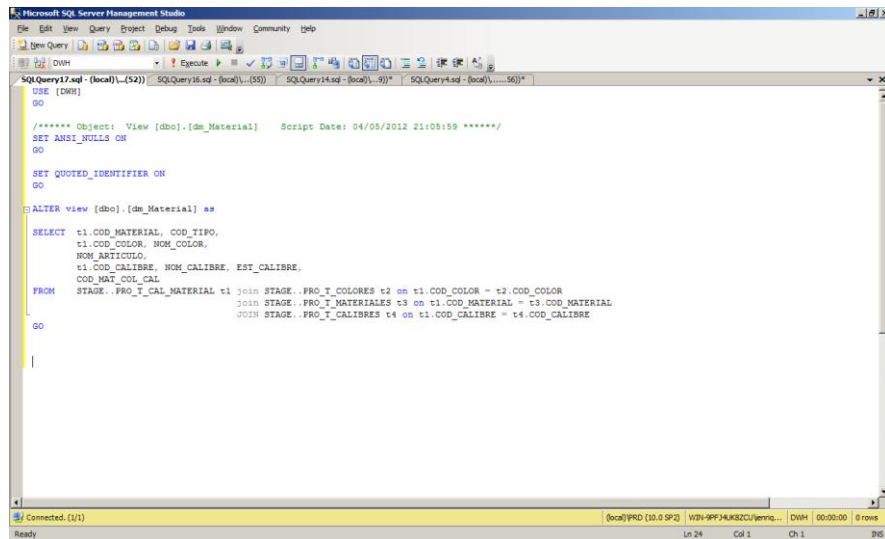


Fig. 6.34: Vista para cargar dimensión material (modelo Copo de nieve)

Capturado por: Ing. Elena Sánchez

- Finalmente se crea el paquete de transformación y carga a través de la herramienta sql execute que se encarga de borrar los datos a través de sentencias truncate de la tablas para evitar duplicados, un execute package que ejecuta la tarea DMS²⁰, una vez recargado la base de datos con los datos se inserta un sql execute para ejecutar los procedimientos almacenados de transformación y carga para terminar con el proceso ETL completo.

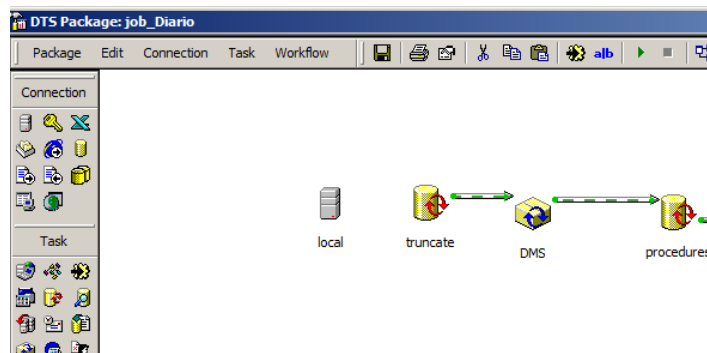


Fig. 6.35: Diagrama final procesos ETL

Capturado por: Ing. Elena Sánchez

²⁰ DMS, execute package que ejecuta el proceso de carga de las fuentes

Implementación

Con los requerimientos capturados, los modelos de datos generados y los procesos de carga definidos implementamos el sistema final que estará disponible para los usuarios

Construcción de los cubos de Información

Procedemos a la creación de la base de datos multidimensional desde Analysis manager.

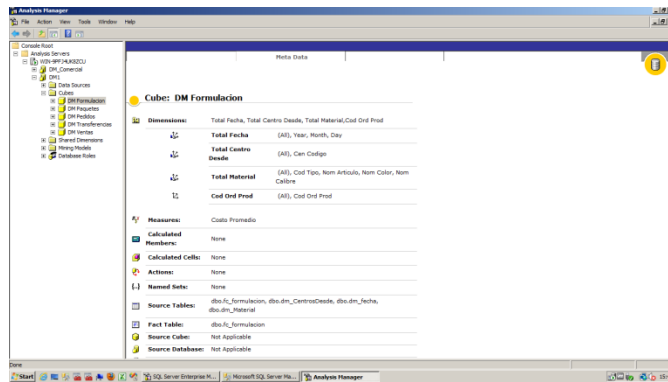
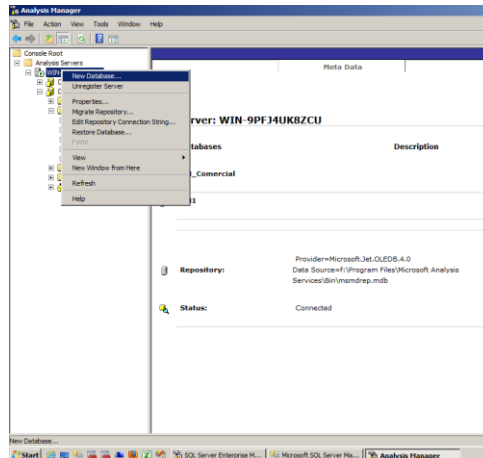


Fig. 6.36: Cubos de información

Capturado por: Ing. Elena Sánchez

Creamos la base de datos DW1 al Analysis Manager.



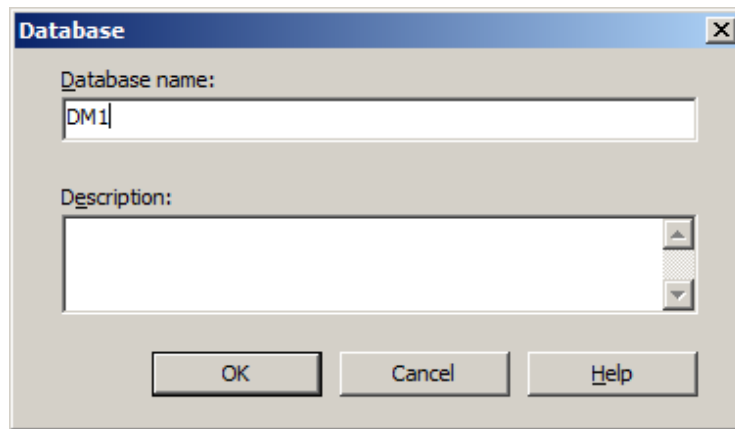


Fig. 6.37: Base DW1 en Analysis Services.

Capturado por: Ing. Elena Sánchez

Creación de la conexión al DWH desde el Analysis Manager

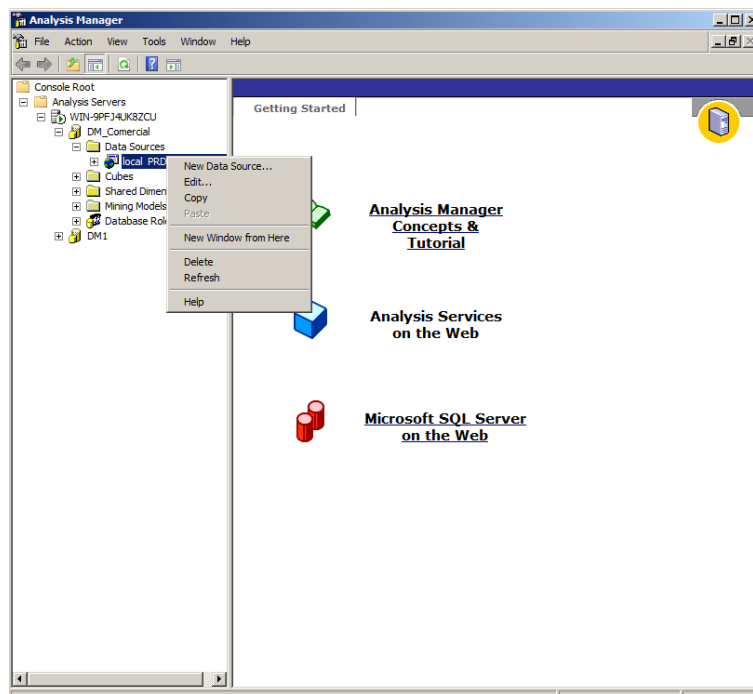


Fig. 6.38: Conexión al DWH desde Analysis Manager

Capturado por: Ing. Elena Sánchez

En las propiedades de conexión escogemos el proveedor en este caso Microsoft OLE Provider for ODBC Driver, en conexión escogemos el servidor PRD y escogemos la base de datos donde se almacena los datos dimensionales es decir la base DM1.

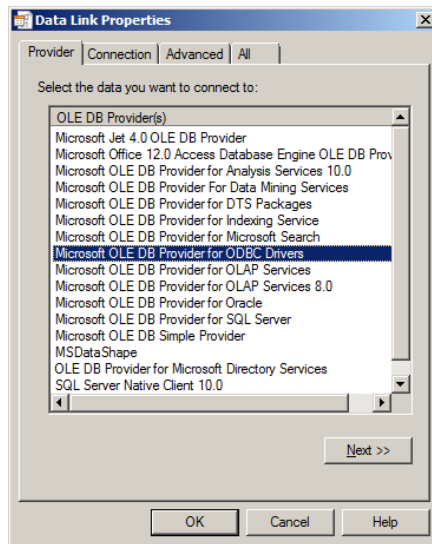


Fig. 6.39: Proveedor de Conexión al DWH desde Analysis Services

Capturado por: Ing. Elena Sánchez

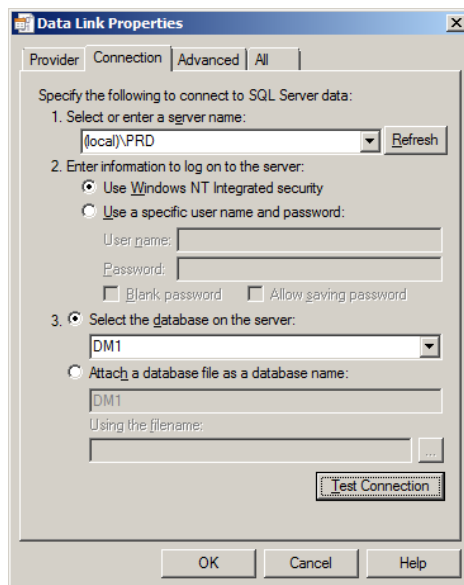
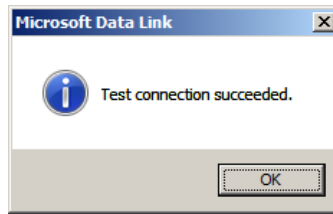


Fig. 6.40: Cadena de Conexión al DWH desde Analysis Services

Capturado por: Ing. Elena Sánchez



Se requiere crear la dimensión de tiempo, una muy importante ya que esto permitirá generar los reportes en la línea de tiempo como históricos, la granularidad más usada es el día por tanto se procede de la siguiente manera:

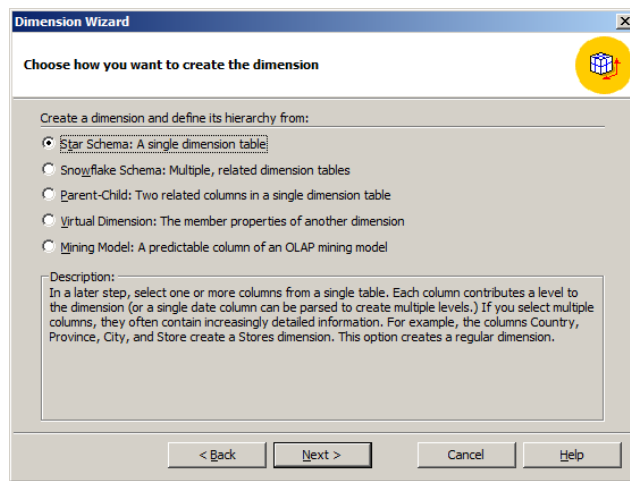


Fig. 6.41: Opciones de creación de dimensiones

Capturado por: Ing. Elena Sánchez

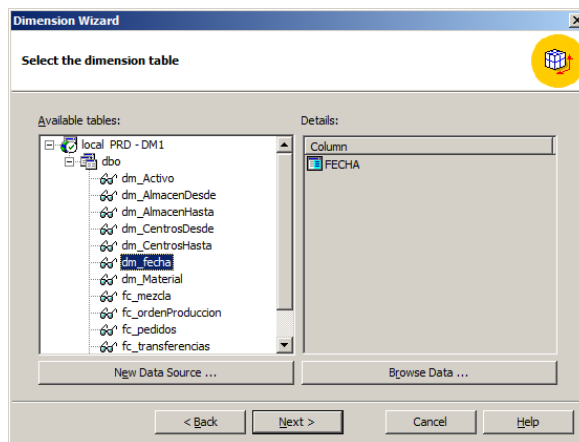


Fig. 6.42: Selección de la tabla o vista que conformara la dimensión

Capturado por: Ing. Elena Sánchez

Una vez relacionando con la tabla dimensión que le permitirá obtener los datos de fecha, definimos que la dimensión es tipo Time dimensión, cuyo dato de columna es tipo Fecha.

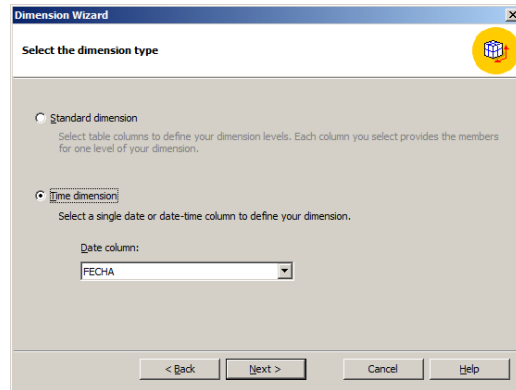


Fig. 6.43: Selección tipo de dimensión
Capturado por: Ing. Elena Sánchez

Definimos los niveles en la estructura de la dimensión tipo, Año-Mes-Día.

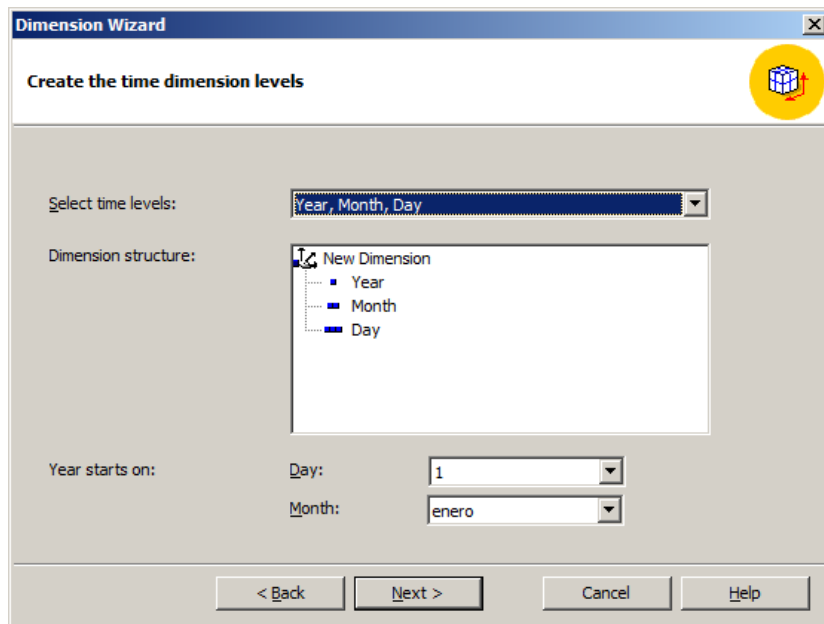


Fig.6.44: Niveles de una dimensión
Capturado por: Ing. Elena Sánchez

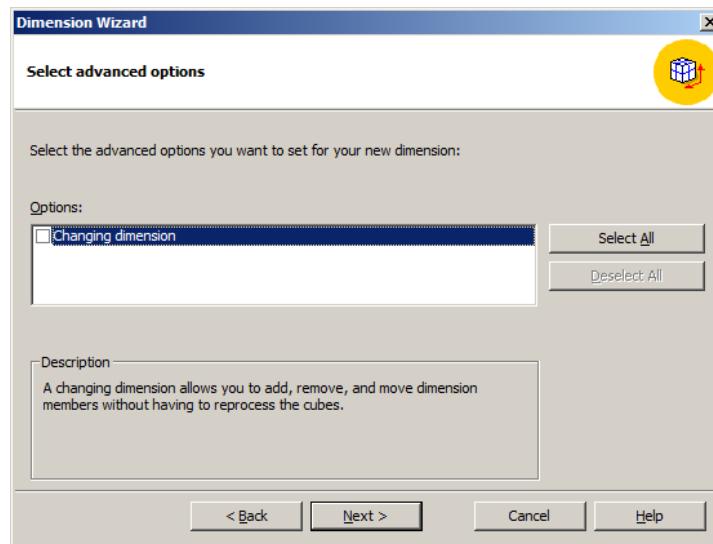


Fig. 6.45: Opciones avanzadas de una dimensión

Capturado por: Ing. Elena Sánchez

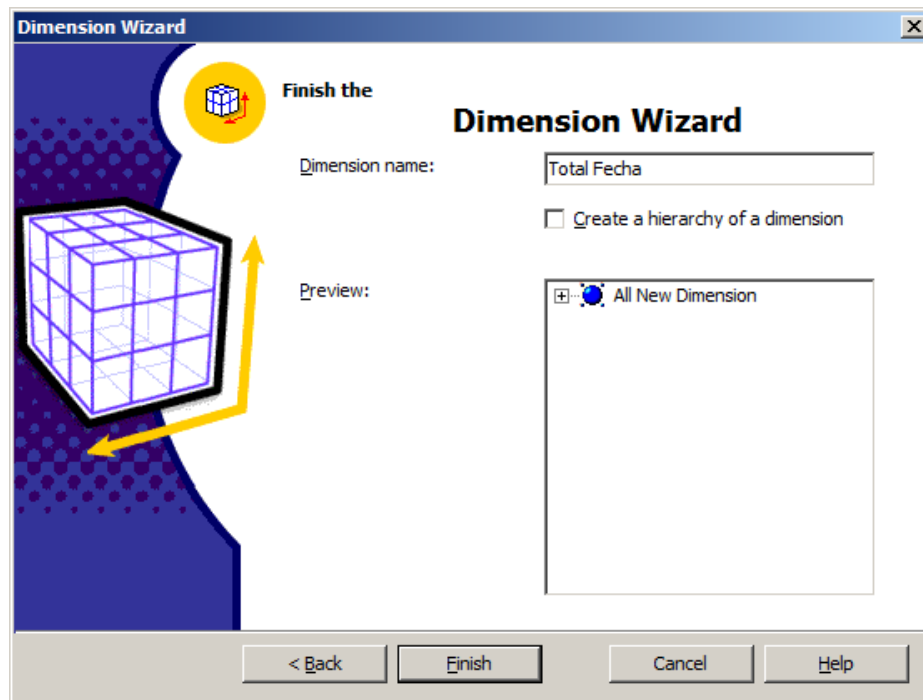


Fig. 6.46: Dimensión creada

Capturado por: Ing. Elena Sánchez

Creación de dimensiones públicas es decir una dimensión que será compartido por todo los cubos, este es el caso de la dimensión material.

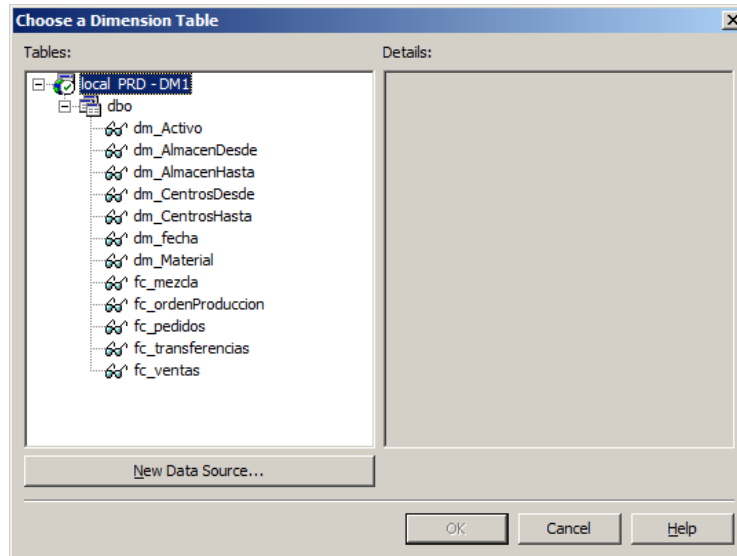


Fig. 6.47: Dimensiones publicas

Capturado por: Ing. Elena Sánchez

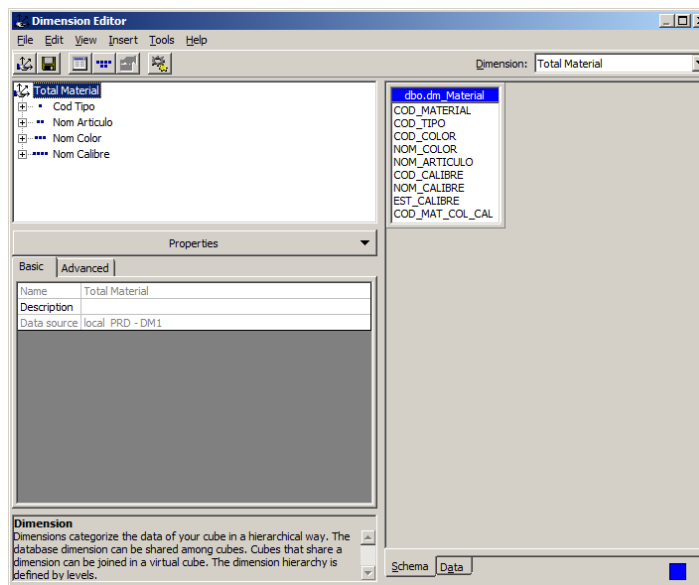


Fig. 6.48: Edición de dimensiones

Capturado por: Ing. Elena Sánchez

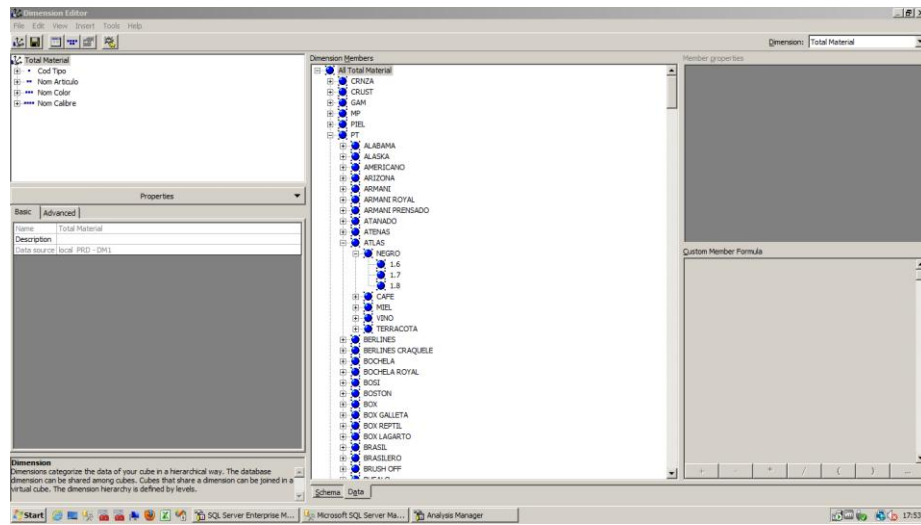


Fig. 6.49: Pre visualización de datos

Capturado por: Ing. Elena Sánchez

Se procede a crear el cubo de información, en el mismo Analysis Manager el árbol organizativo nos ubicamos en cube, con el menú emergente seleccionamos New Cube.

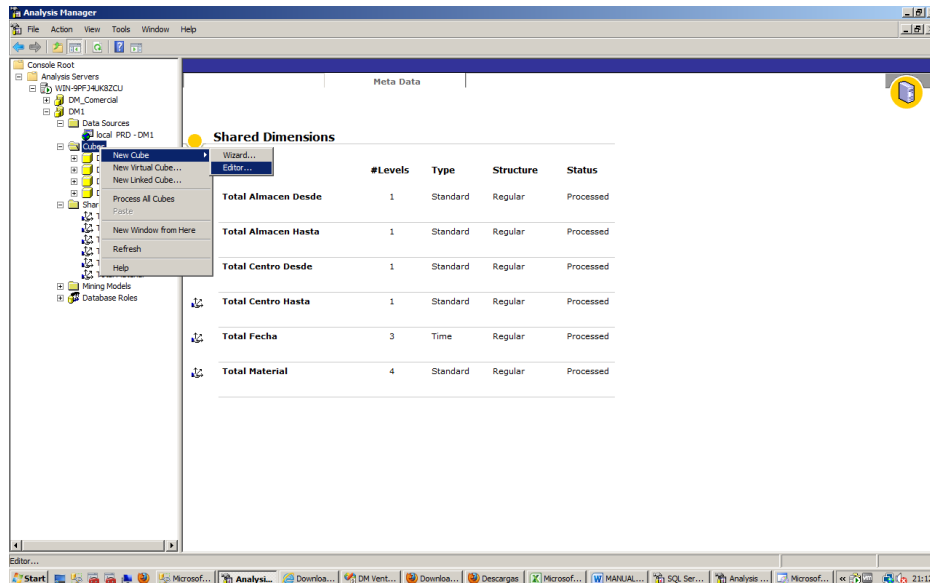


Fig. 6.50: Nuevo cubo

Capturado por: Ing. Elena Sánchez

Seleccionamos la tabla de hechos que conformara el cubo en construcción en este caso es la tabla ventas.

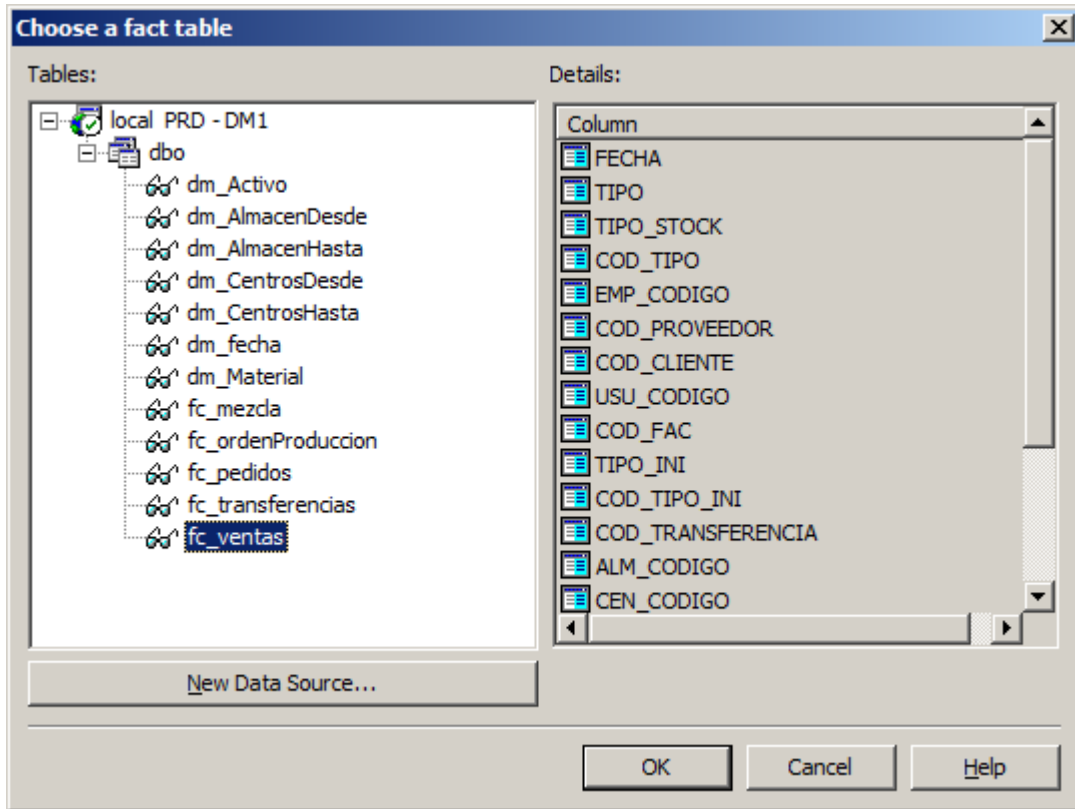
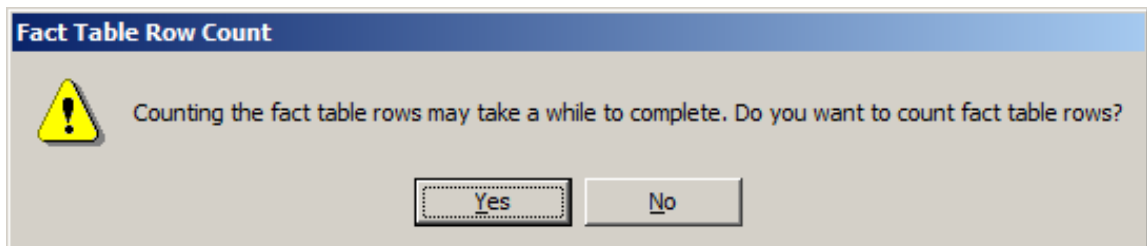


Fig. 6.51: Selección de la tabla de hechos

Capturado por: Ing. Elena Sánchez



Se agrega la tabla al editor del cubo y se procede a agregar las dimensiones y medidas.

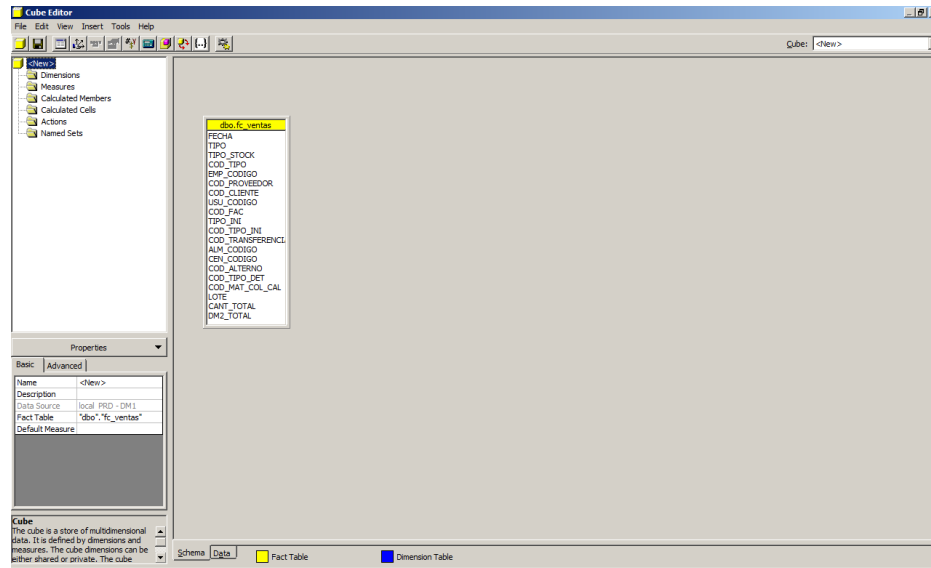


Fig.6.52: Editor de cubos
Capturado por: Ing. Elena Sánchez

Para agregar las dimensiones se ubica en el árbol organizativo de la izquierda en dimension

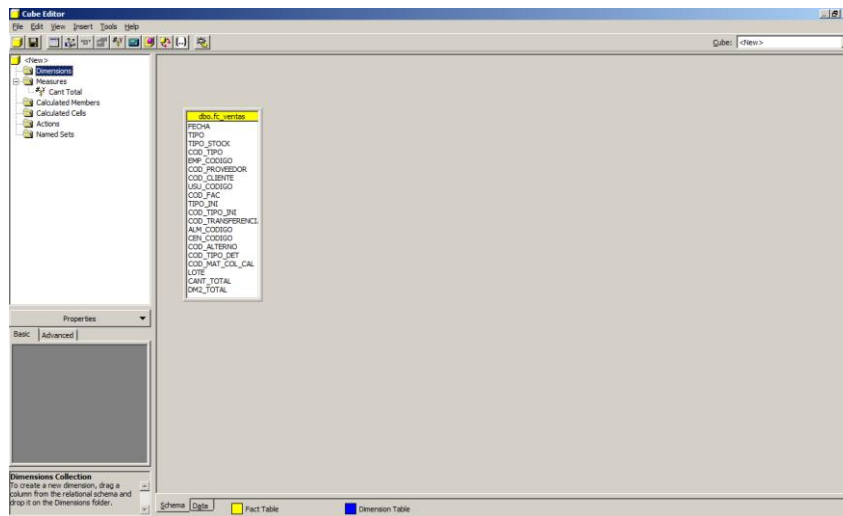


Fig. 6.53: Editor cubos
Capturado por: Ing. Elena Sánchez

Agregar dimensiones públicas, desde el menú emergente en existing dimensión y escogemos las dimensiones que se requieren para formar el diagrama tipo estrella.

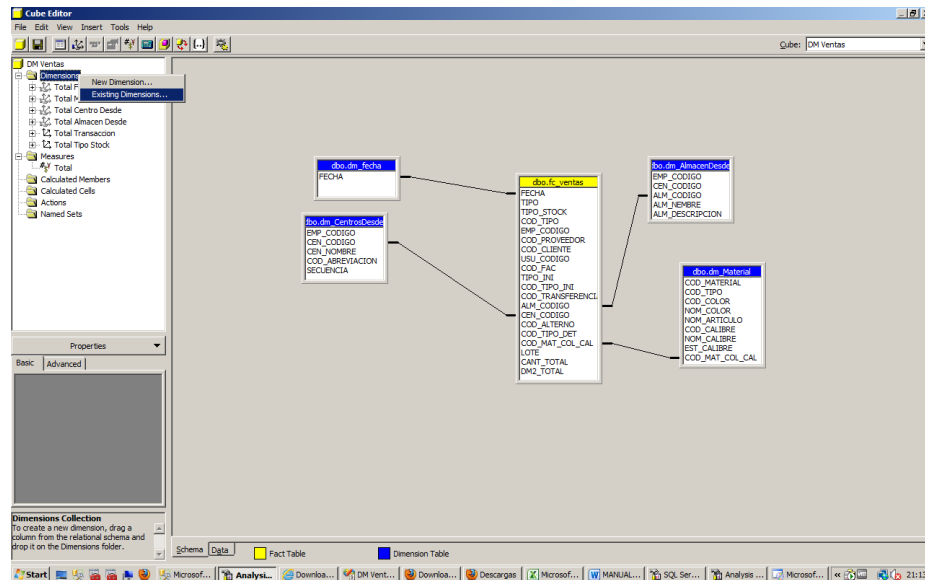


Fig. 6.54: Editor cubos- Agregar dimensiones

Capturado por: Ing. Elena Sánchez

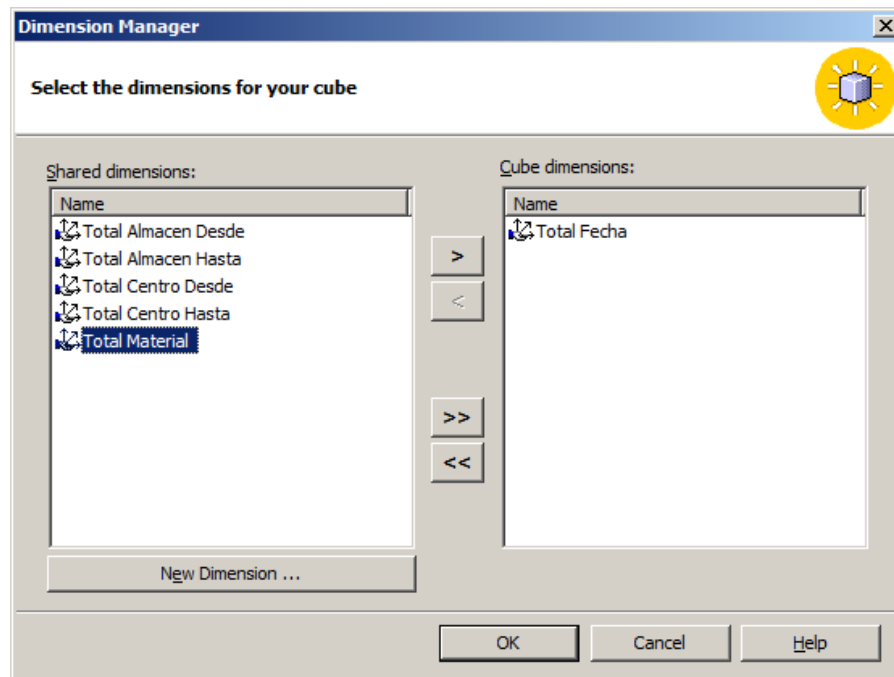


Fig. 6.55: Selección de la dimensión para el cubo

Capturado por: Ing. Elena Sánchez

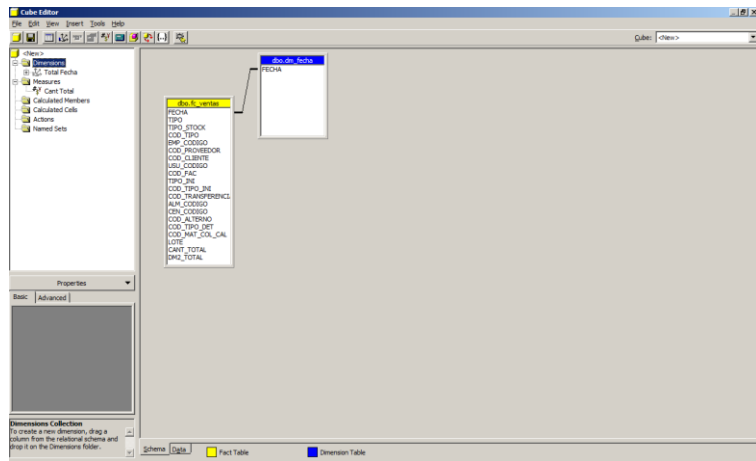


Fig. 6.56: Dimensión Agregada
 Capturado por: Ing. Elena Sánchez

Para agregar dimensiones privadas, se lo realiza arrastrando el campo de la tabla de hechos hasta el área de dimensiones en el árbol organizativo de la izquierda.

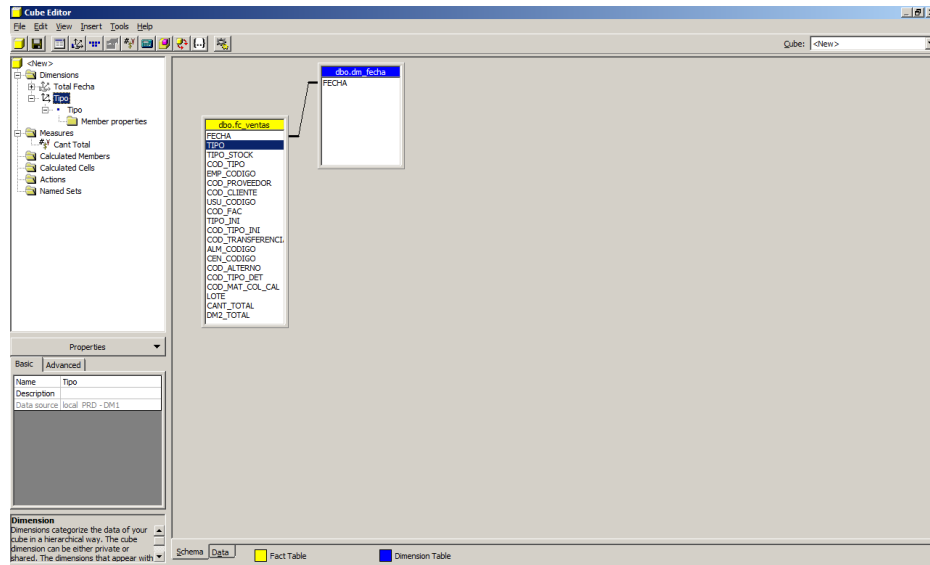


Fig. 6.57: Agregar dimensiones privadas
 Capturado por: Ing. Elena Sánchez

El cubo final con esquema estrella.

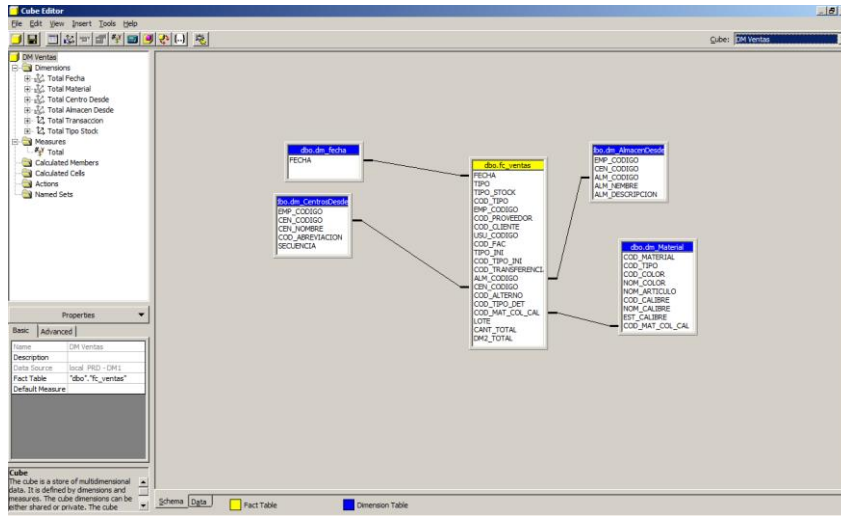


Fig. 6.58: Cubo ventas final tipo estrella

Capturado por: Ing. Elena Sánchez

Una vez creado el cubo se realiza el procesamiento del mismo de la siguiente manera.

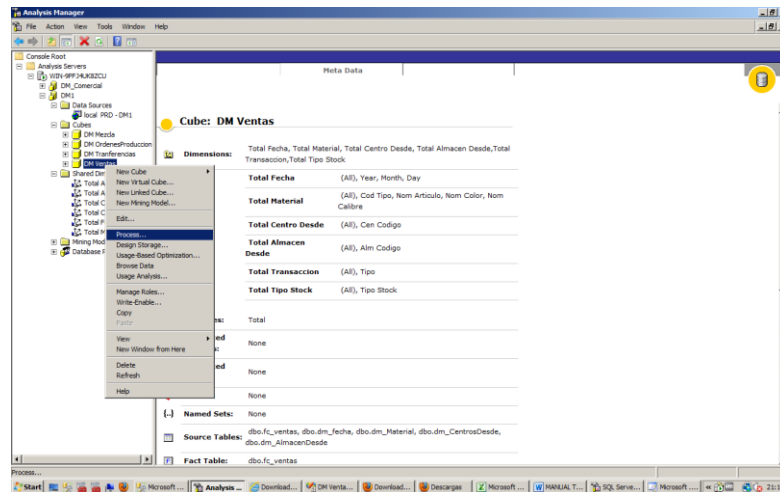


Fig. 6.59: Procesamiento del cubo

Capturado por: Ing. Elena Sánchez

El procesamiento full, es el recomendado ya que ofrece la ventaja de que todos los datos de la tabla de hechos van a ser actualizados, la actualización incremental de las

dimensiones ayuda a que sus datos no sean reprocesados sino solo agregados los últimos que se fueron creando.

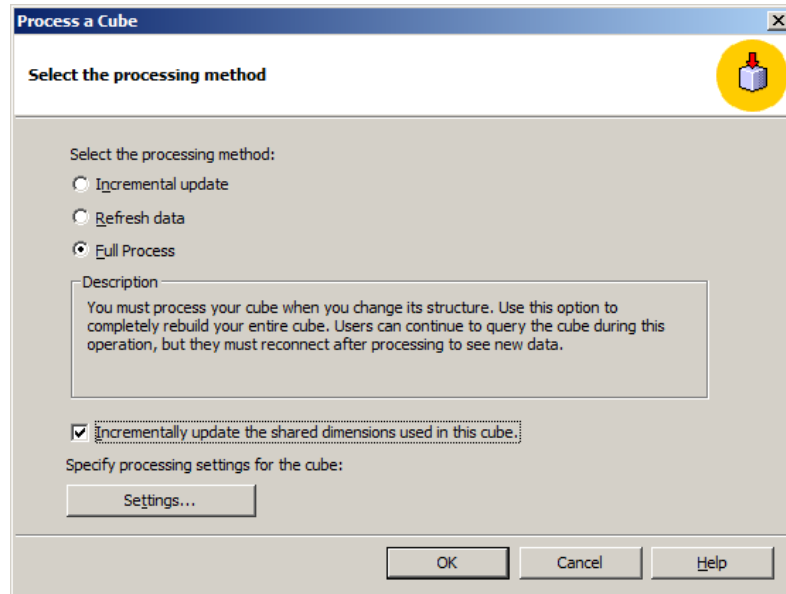


Fig. 6.60: Método de procesamiento del cubo

Capturado por: Ing. Elena Sánchez

Optimización para mejorar desempeño de consultas.

Existen distintas formas de acceder a la información a la hora de darle tratamiento a los cubos.

MOLAP

En este esquema, los datos son pre calculados y almacenados multidimensionalmente en cubos de datos. Es decir, los datos se toman del datawarehouse, se pre calculan y se almacenan en cubos. El resultado es un tiempo de respuesta más rápido ya que los datos se encuentran pre calculados y no necesitan ser recalculados ante cada requerimiento. Lo negativo de este esquema es que almacenar datos multidimensionalmente requiere mucho más espacio.

ROLAP

Las implementaciones ROLAP son similares en funcionalidad a las MOLAP. Sin embargo, estos sistemas utilizan un motor de base de datos relacional en lugar de un motor de base de datos multidimensional especializado. Esto da mayor escalabilidad dado que permiten manejar más grandes volúmenes de datos que un esquema MOLAP. Además, las implementaciones ROLAP generalmente tienen un mejor esquema de drilling (análisis de datos a diferentes niveles), Las estructuras de datos utilizadas requieren, al igual que en el esquema MOLAP, de un tiempo inicial de diseño e implementación bastante grande. Adicionalmente, dado que las estructuras de datos son relacionales, para acceder a los registros detallados se debe utilizar SQL. De esta manera, el motor ROLAP debe realizar trabajos adicionales para realizar comparaciones, tales como comparar el importe de ventas del mes actual con el del mismo mes del año pasado.

HOLAP

Algunas empresas de tecnología venden soluciones capaces de acceder bases de datos relacionales directamente desde un motor de base de datos multidimensional, apareciendo así el concepto de HOLAP (Hybrid OLAP). Estas implementan el concepto de drilling, el cual automáticamente genera SQL para recuperar registros detallados para mayor análisis. Esto le da al usuario final la percepción de que está pasando de la base de datos multidimensional a la relacional. El esquema HOLAP combina la performance y funcionalidad de una base de datos multidimensional con la capacidad de visualizar datos más detallados (drilling).

DOLAP (Desktop OLAP)

El Mercado OLAP desktop resultó de la necesidad de los usuarios de correr consultas de negocios usando sets de datos relativamente pequeños extraídos de los sistemas productivos. Muchos sistemas OLAP desktop fueron desarrollados como extensiones de sistemas de reportes de producción, mientras otros fueron desarrollados en la época temprana de los sistemas client/server para utilizar las ventajas de la naciente PC. Los sistemas OLAP de desktop son populares y típicamente requieren poca inversión en IT para ser implementados. También proveen operaciones OLAP

altamente móviles para aquellos usuarios que trabajan remotamente o viajan en forma permanente.

En este proyecto el más recomendable es el MOLAP.

Diseño de agregaciones

Uno de los requerimientos básicos de una datawarehouse tiene que ver con su rendimiento, es decir la capacidad de responder las consultas en tiempos razonables y aceptables.

La solución ideal es implementar un conjunto de resúmenes por cada tabla de hechos; esto es, precalcular algunos valores agregados y almacenarlos en la base de datos para incrementar el rendimiento, de manera que cada vez que se ejecute una consulta sobre el datawarehouse éste seleccionará adecuadamente de dónde extraer sus resultados, eligiendo los resúmenes correctos cada vez que sea posible.

Existen principalmente dos técnicas, para la implementación de resúmenes, la primera consiste en almacenar estos registros precalculados en nuestra tabla de hechos y dimensiones original, esto implica que debemos de alguna manera introducir un nuevo campo “NIVEL” encada una de las tablas afectadas con el objetivo de distinguir los registros resumidos de los originales, y no cometer errores en los cálculos.

La segunda alternativa, más recomendada, consiste en seleccionar cada resumen y ubicarlo en nuevas tablas de hechos, que ahora se rodearán de nuevas tablas dimensionales más pequeñas que contendrán sólo aquellos atributos que tengan sentido para cada nivel de resumen. Ambas técnicas producen el mismo número extra de registros, la diferencia es sólo dónde los almacenamos.

Para diseñar una agregación se realiza las siguientes tareas:

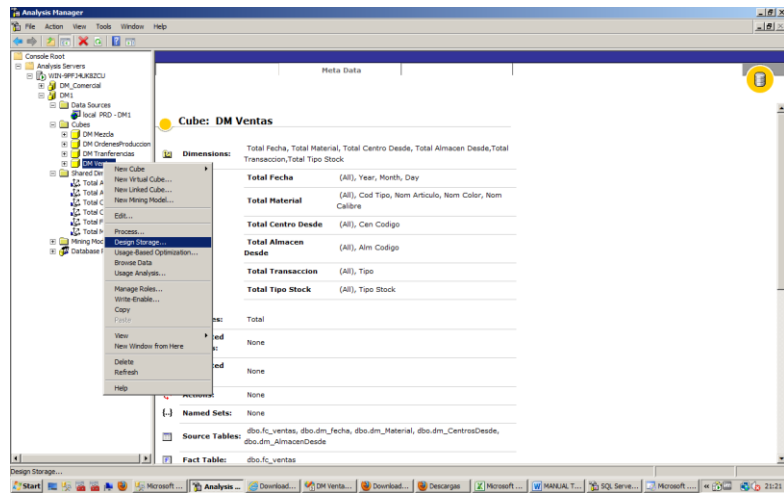
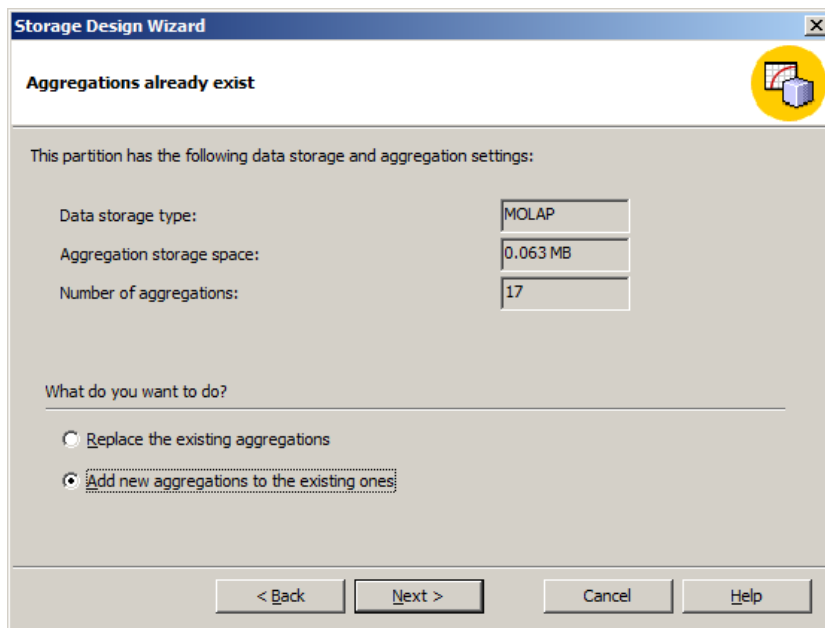


Fig. 6.61: Diseño de agregaciones

Capturado por: Ing. Elena Sánchez



Lo recomendado es diseñar las agregaciones en base a un porcentaje de crecimiento o un tamaño de almacenamiento.

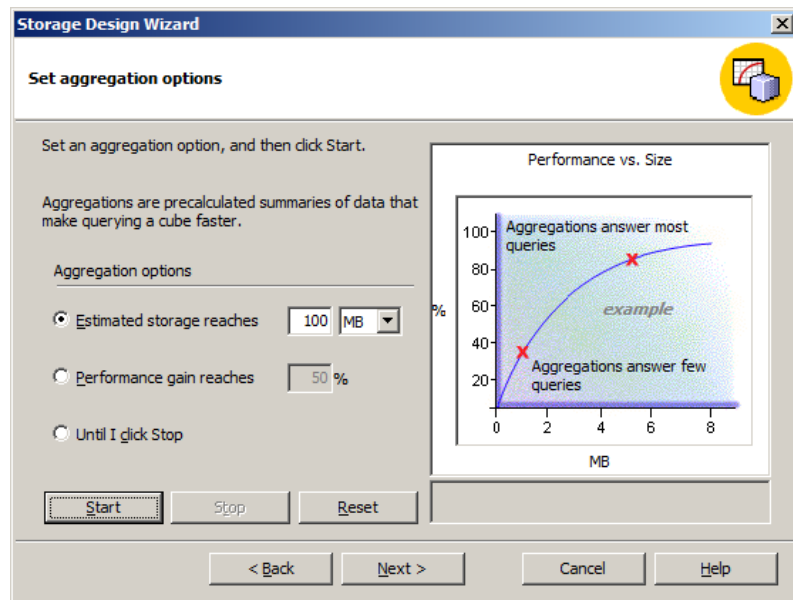


Fig. 6.62: Rendimiento cubo de información
 Capturado por: Ing. Elena Sánchez



Fig. 6.63: Procesamiento de las agregaciones en el cubo
 Capturado por: Ing. Elena Sánchez

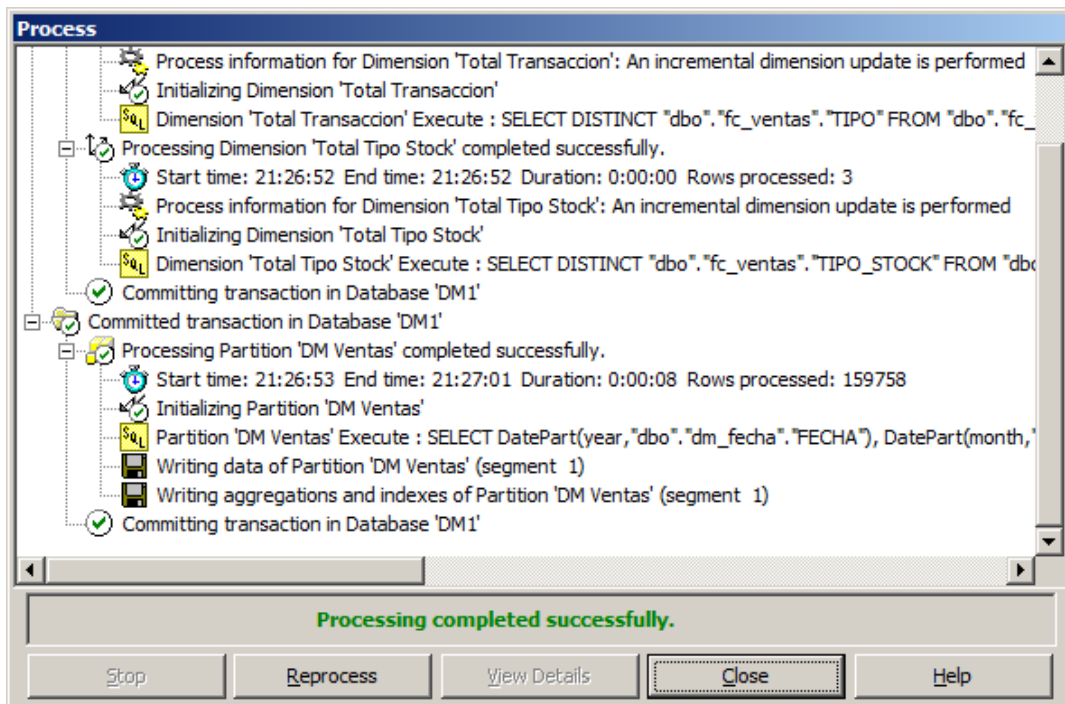
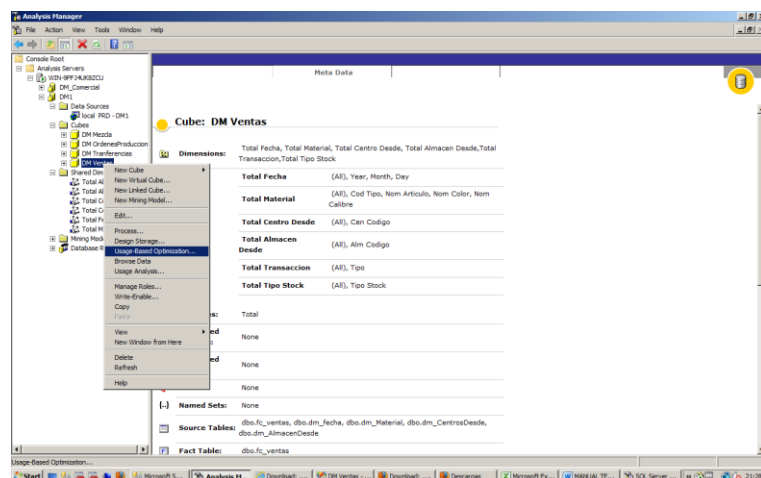


Fig. 6.64: Procesamiento del cubo

Capturado por: Ing. Elena Sánchez

Optimización basada en el uso.

Se puede procesar el cubo optimizándolo según el uso que se le dé:



Capturado por: Ing. Elena Sánchez

Se puede optar por varias opciones entre ellas, consultas entre fechas, consultas por usuarios, consultas con procesamiento Molap, Rolap, o cuyas cache este en el servidor mismo, la que más utiliza es la de consultas más frecuentes, e incluso se puede definir el número de veces que corrieron dichas consultas.

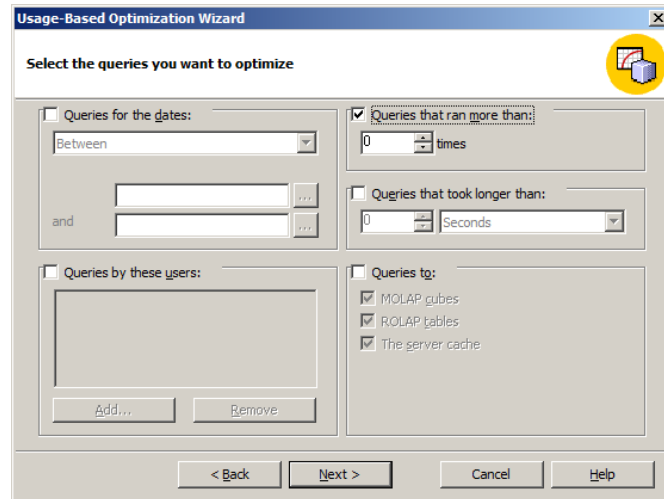


Fig. 6.65: Selección de los consultas sql a optimizar
Capturado por: Ing. Elena Sánchez

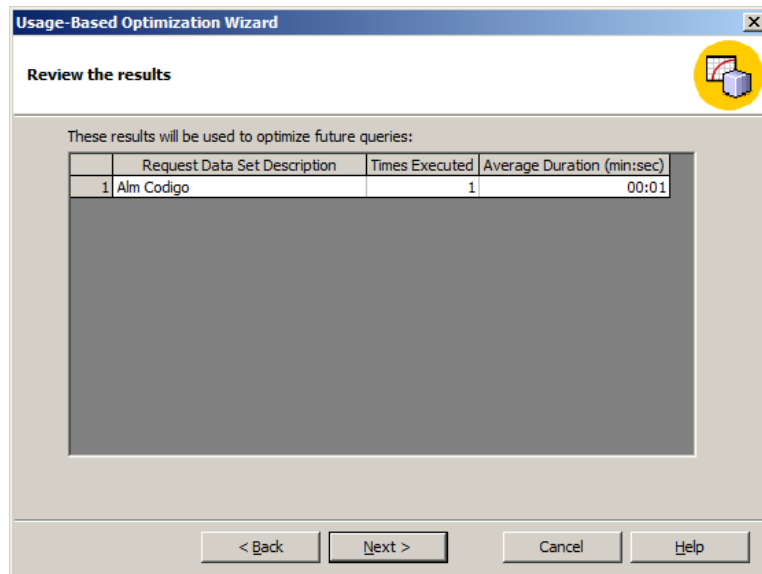


Fig. 6.66: Revisión consulta optimizado
Capturado por: Ing. Elena Sánchez

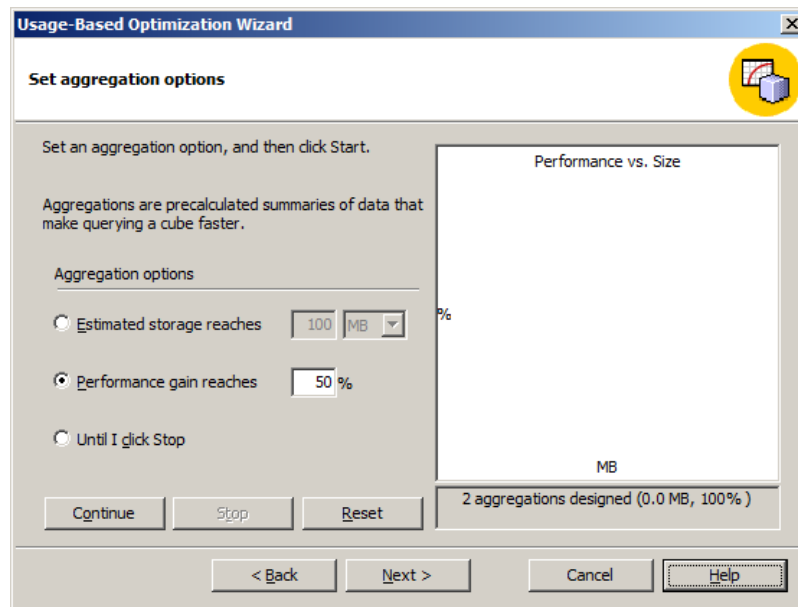


Fig. 6.67: Opciones para procesar las agregaciones

Capturado por: Ing. Elena Sánchez

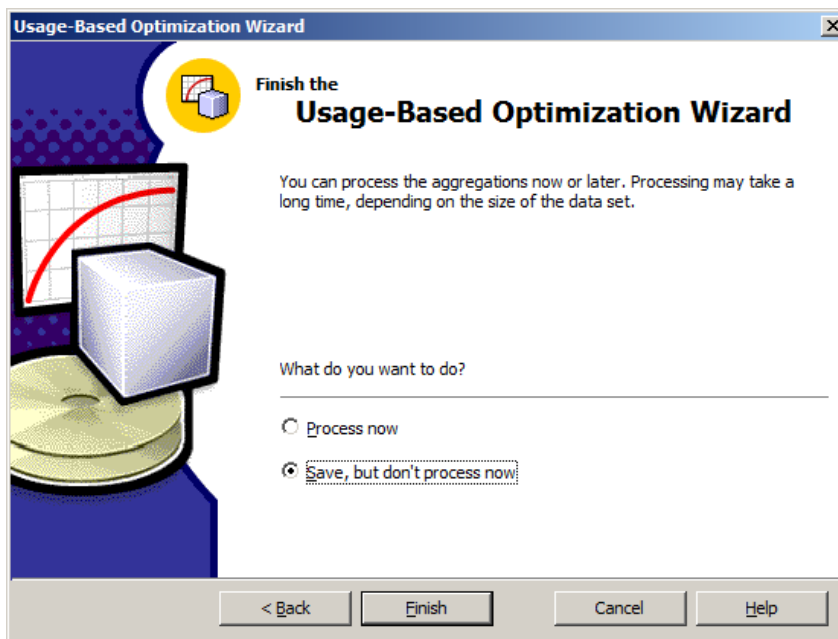


Fig.6.68: Procesamiento de la optimización basada en el uso

Capturado por: Ing. Elena Sánchez

Construcción del DTS²¹ definitivo

Para evitar que los de datos se dupliquen, creamos un Tarea ejecutable SQL que realice un truncate a las tablas.

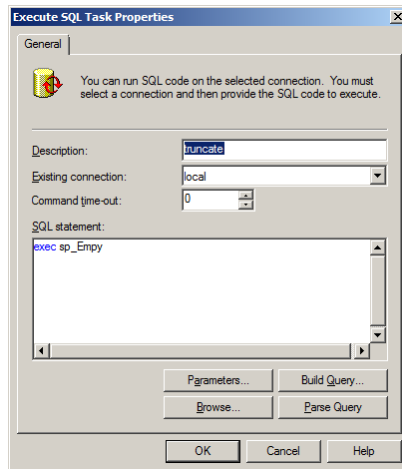


Fig. 6.69: Tarea de ejecución de limpieza de tablas

Capturado por: Ing. Elena Sánchez

Ejecutar carga

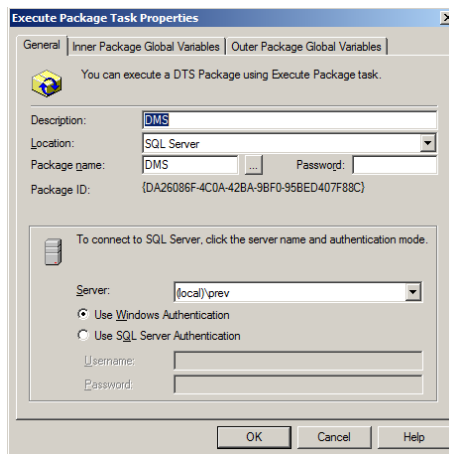


Fig. 6.70: Tarea de ejecución de paquetes

Capturado por: Ing. Elena Sánchez

²¹ DTS, Servicio de transformación de datos

Ejecuta los procedimientos de carga, se enumera cada uno de los procedimientos almacenados para cargar las tablas de hechos.

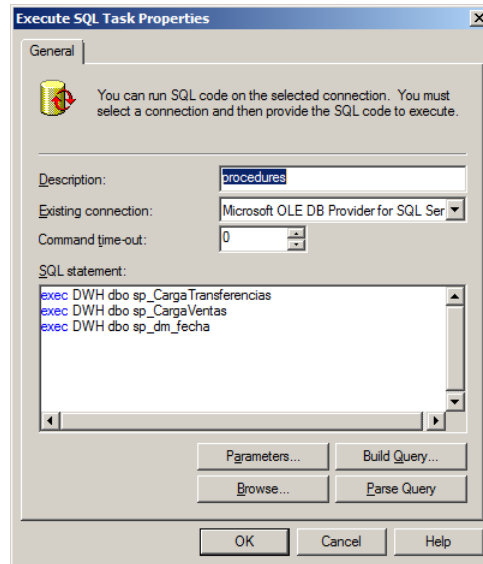


Fig. 6.71: Tarea de ejecución de procedimientos de carga
Capturado por: Ing. Elena Sánchez

Ejecuta el cubo de información.

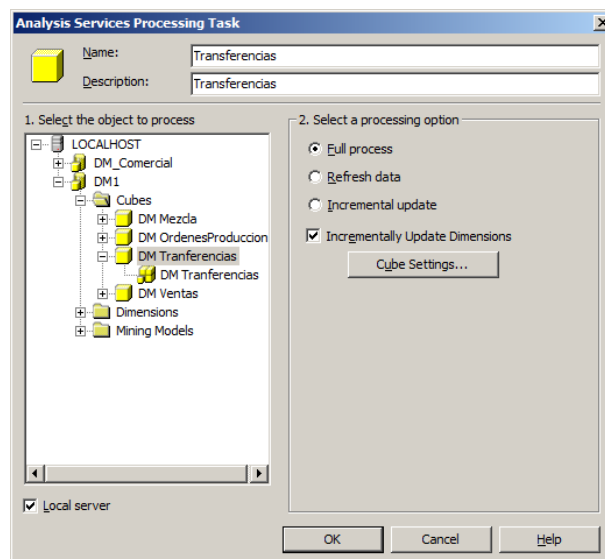


Fig. 6.72: Tarea de ejecución del cubo

Capturado por: Ing. Elena Sánchez

Finalmente se tiene el diagrama DTS completo:

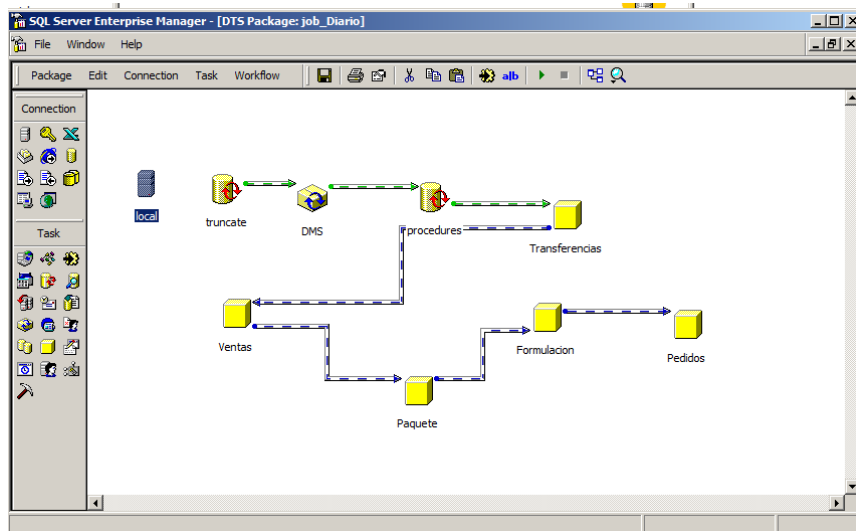


Fig. 6.73: Diagrama final de ejecución

Capturado por: Ing. Elena Sánchez

Si deseamos procesarlo manualmente se lo ejecuta con **F5**

Calendarizando la tarea programada diariamente.

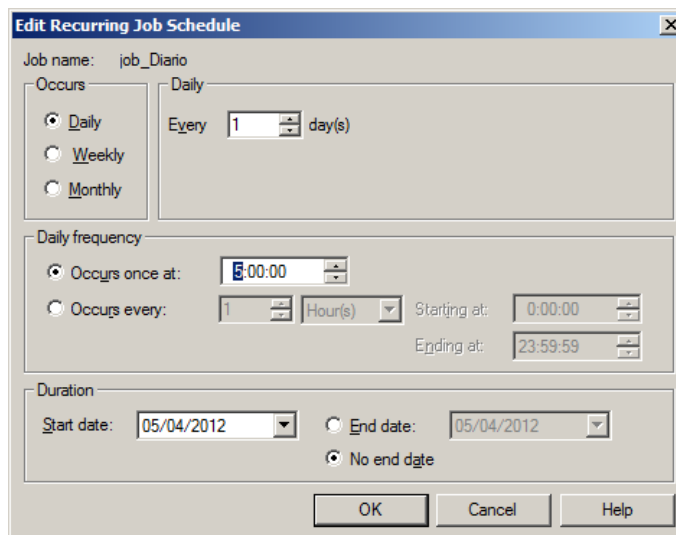


Fig. 6.74: Calendarización tarea programada de procesamiento

Capturado por: Ing. Elena Sánchez

Panel de monitoreo.

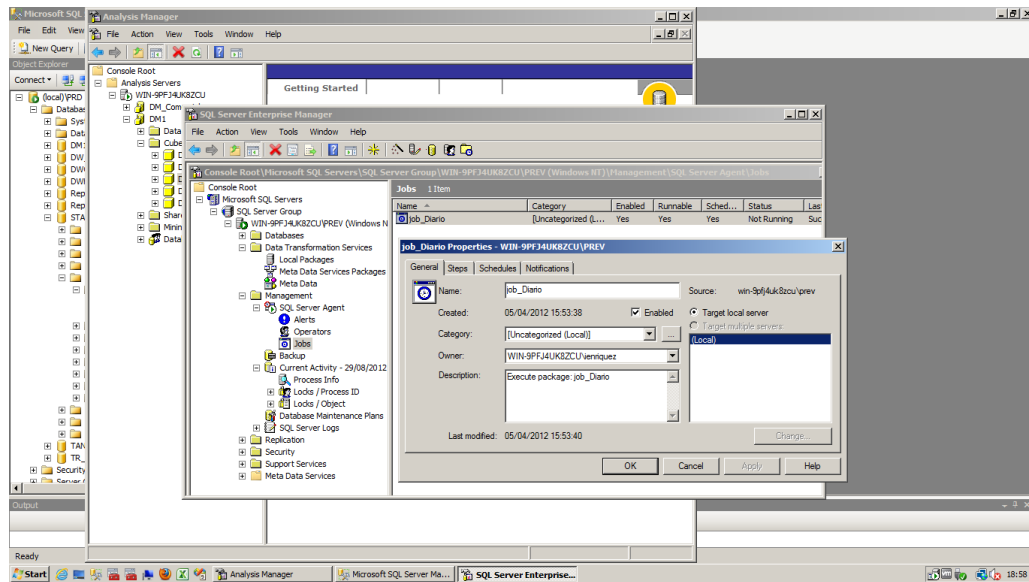


Fig. 6.75: Panel de monitoreo del tarea programada de procesamiento

Capturado por: Ing. Elena Sánchez

REPORTING SERVICES

Reporting Services es una plataforma que permite definir, administrar y distribuir distintos formatos de reportes dentro de una organización o a través de múltiples organizaciones.

Este servicio es una extensión a las capacidades Inteligencia de Negocios que nos provee de herramientas para almacenar información (Report Server), herramientas para crear reportes (Report Designer) y herramientas para administrar reportes (Report Manager).

Ventajas de Reporting Services:

- Una de las principales, es que se cuenta con una interface Web para lo que es la administración de los reportes, desde esta interface se puede determinar en qué formato debe llegar el reporte, es decir, podemos decidir que el reporte

llegue a una de las gerencias en formato PDF y para el departamento de consolidación y validación de datos podría enviarse la información en formato XML.

- ✚ Cuenta con un lenguaje de especificación estándar denominado Report Definition Language o simplemente RDL, el cual es un lenguaje de formato XML, que se encarga de definir el reporte.
- ✚ Con SQL Server Reporting Services, podemos conectarnos a cualquier repositorio de datos, a través de un .NET Data Provider, un proveedor OLE DB provider o uno de tipo ODBC.
- ✚ Para la distribución, los usuarios pueden acceder a los reportes en base a la infraestructura existente. Es decir, que los usuarios pueden acceder a los reportes a través de una barra de herramientas en el browser. Los reportes son accesados desde un repositorio centralizado, presentado como un folder en orden jerárquico.
- ✚ Otra de las grandes características de Reporting Services, es que puede distribuir el reporte en distintos formatos, como hojas de excel, documentos pdf, texto, XML, etc.
- ✚ La arquitectura de Reporting Services, permite a los desarrolladores preparar aplicaciones personalizadas que accedan a los reportes a través de una API que está expuesta como un Web service.

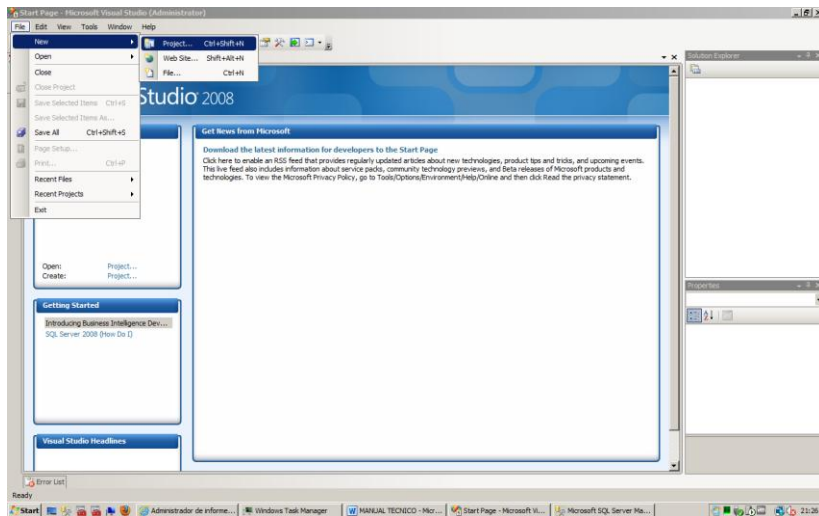


Fig. 6.76: Creación de un proyecto de report services

Capturado por: Ing. Elena Sánchez

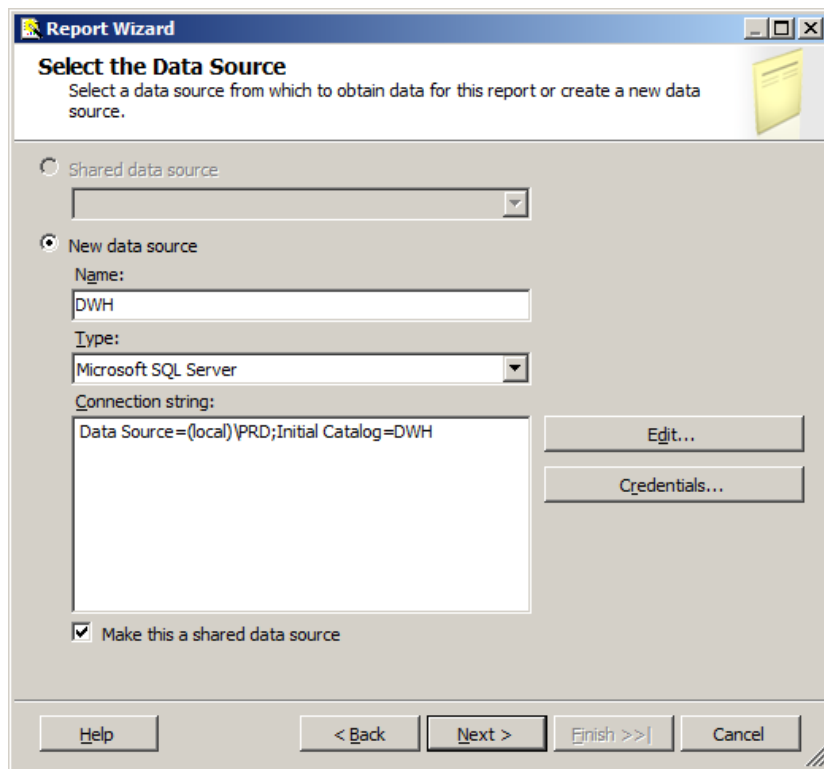


Fig. 6.77: Definición del origen para el report services

Capturado por: Ing. Elena Sánchez

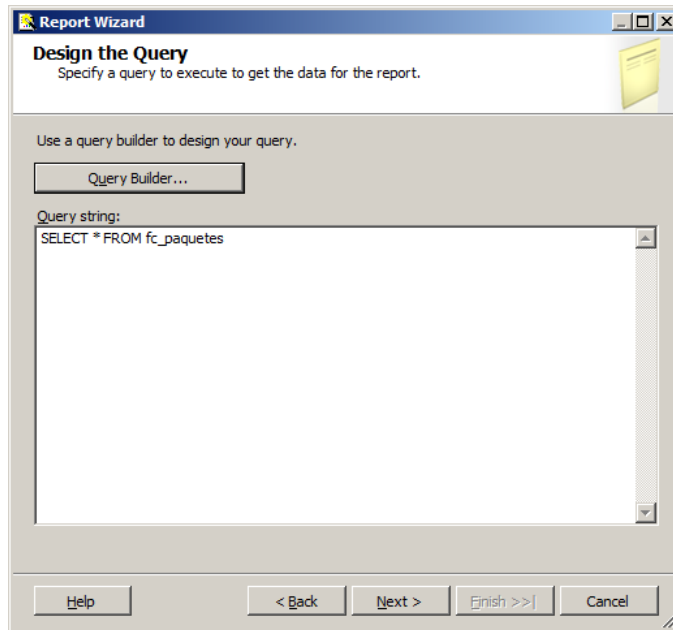


Fig. 6.78: Definición de la consulta a reportearse

Capturado por: Ing. Elena Sánchez

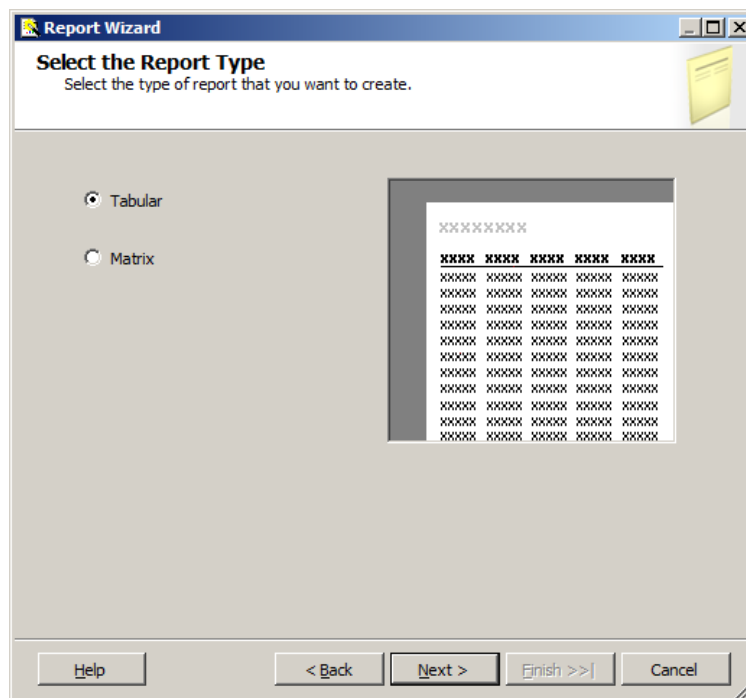


Fig. 6.79: Tipo de reporte

Capturado por: Ing. Elena Sánchez

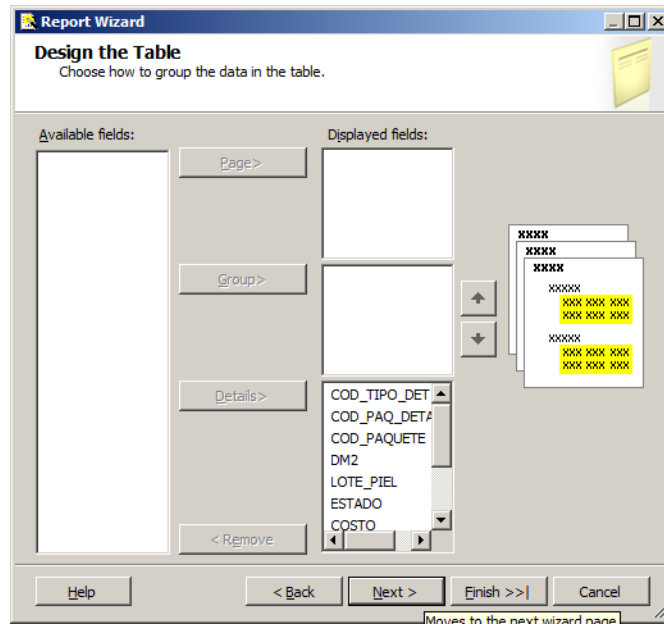


Fig. 6.80: Definición del diseño del reporte

Capturado por: Ing. Elena Sánchez

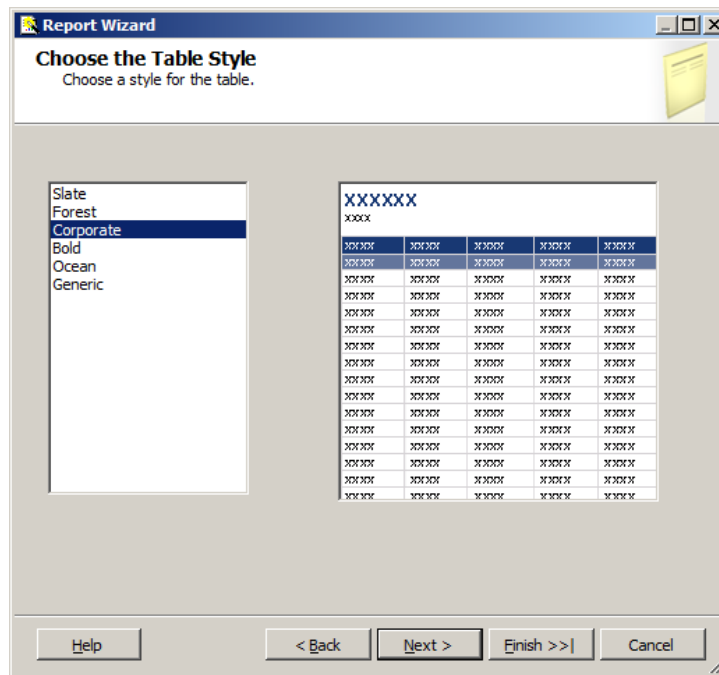


Fig. 6.81: Definición del estilo del reporte

Capturado por: Ing. Elena Sánchez

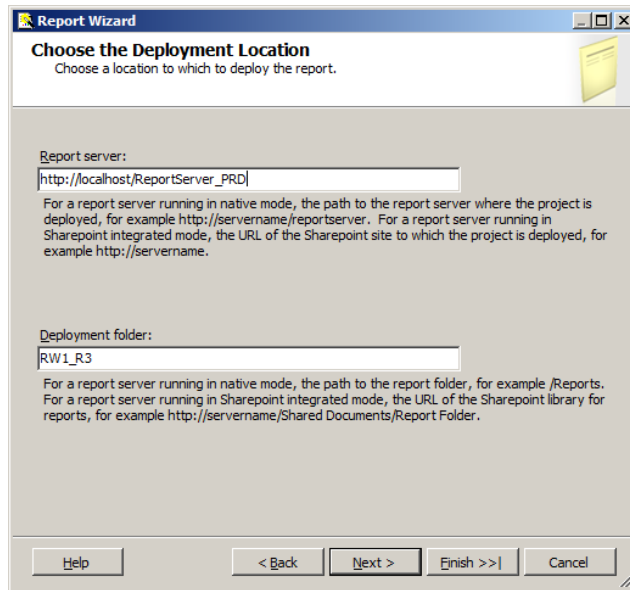
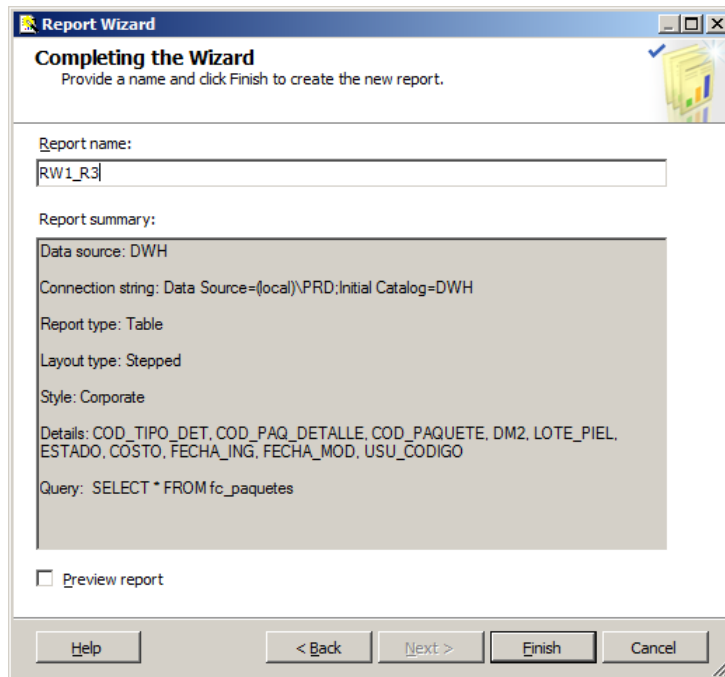


Fig. 6.82: Definición de la ubicación del reporte

Capturado por: Ing. Elena Sánchez



Implantación

Esta es la última fase de desarrollo del DWH, para lo cual empezamos con la configuración desde un cliente en este caso Microsoft Excel 2010, que es la herramienta más explotada por la empresa

Configurando la consulta de los datos.

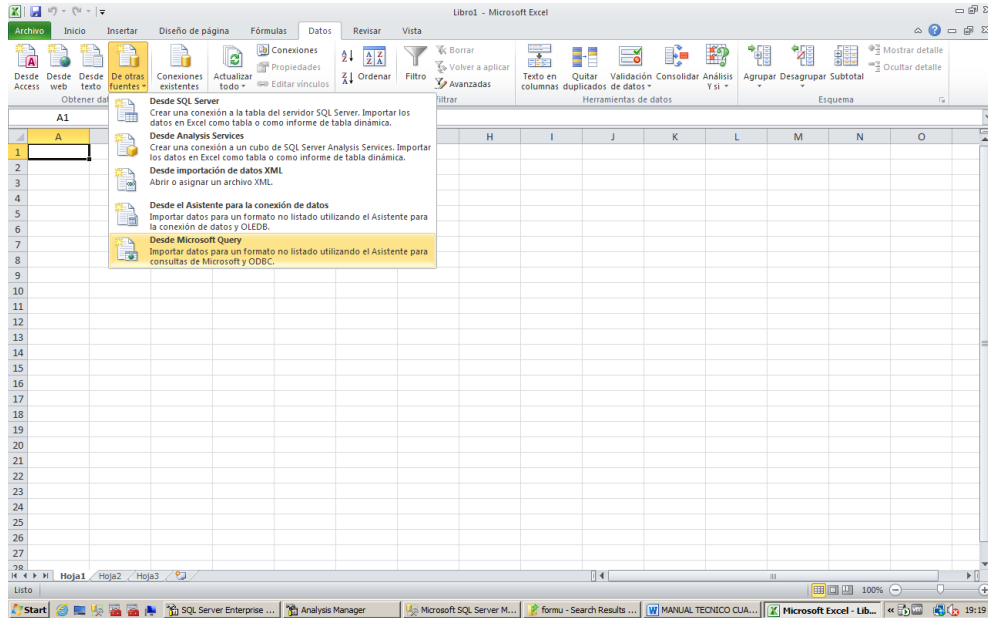


Fig. 6.83: Configuración de Ms Excel para la obtención del reporte

Capturado por: Ing. Elena Sánchez

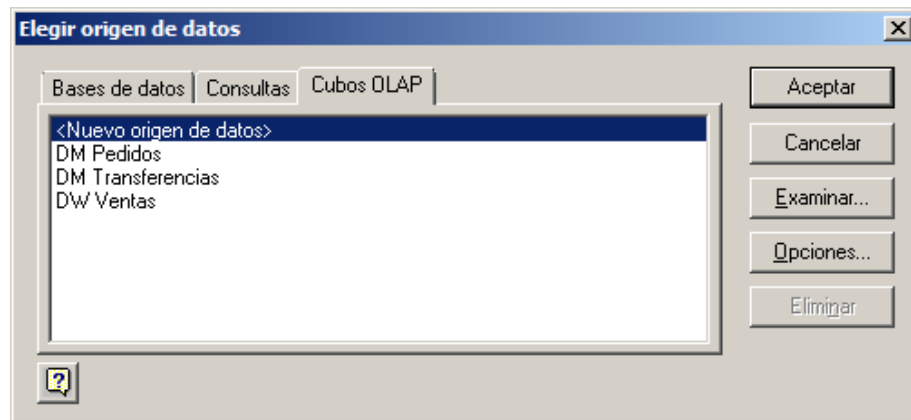


Fig. 6.84: Elección del origen de datos

Capturado por: Ing. Elena Sánchez

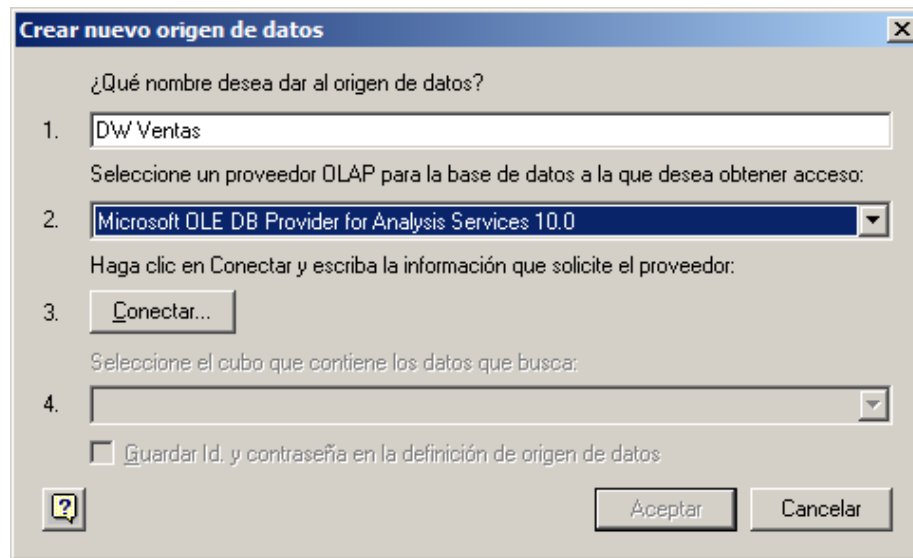


Fig. 6.85: Creación del origen
Capturado por: Ing. Elena Sánchez

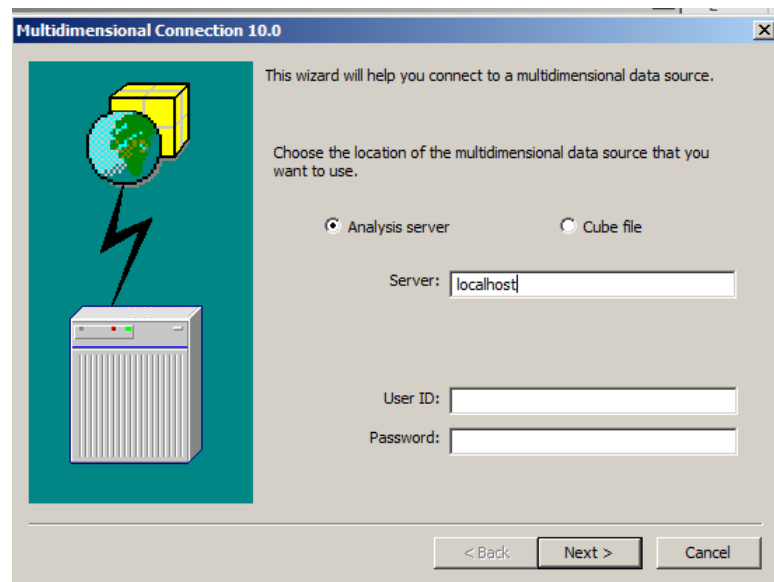


Fig. 6.86: Conexión con analysis server
Capturado por: Ing. Elena Sánchez

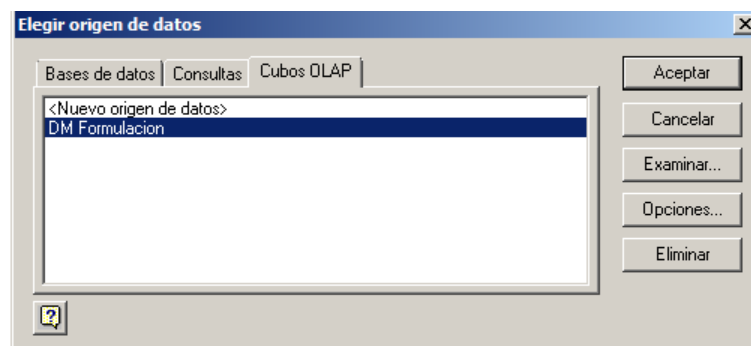
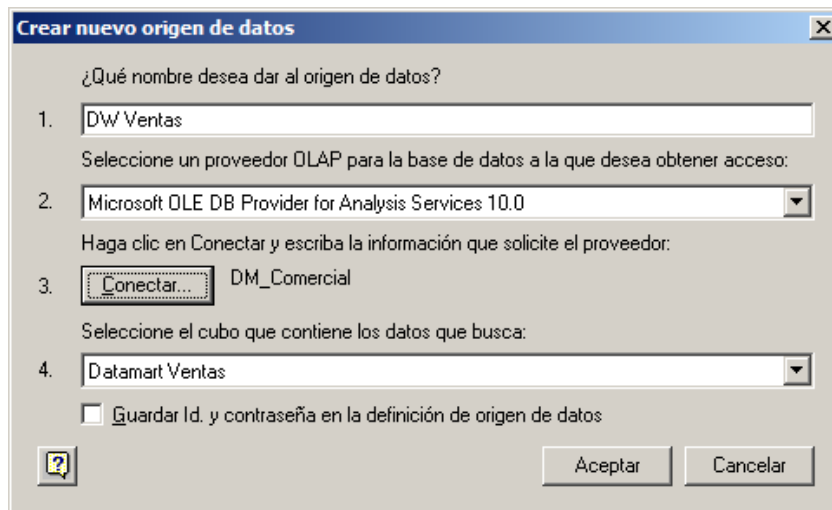


Fig. 6.87: Selección del cubo

Capturado por: Ing. Elena Sánchez

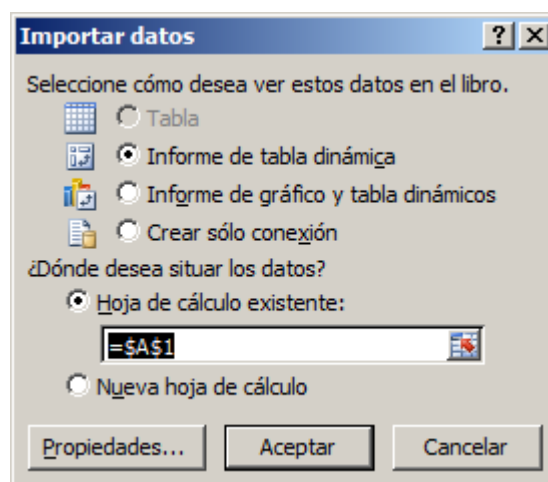


Fig. 6.88: Importar datos

Capturado por: Ing. Elena Sánchez

Reportes

A continuación se detalla los reportes obtenidos e instalados a cada uno de los usuarios según sus necesidades.

Historico de pedidos de materia prima, definidos por centros de costos

Etiquetas de fila	2007	2008	2009	2010	2011	2012	Total general
GAMB	25						25
MP	10555487.41	12802101.18	16694619.35	18299774.11	1980486.23	4820516.7	83977395
ACIDO FORMICO	7508.81	14374.97	19422.39	24561.5	2575.32	6907.32	58352.31
ACIDO OXALICO	3.35	15.3	36.36	28.17	46.34	21.2	152.72
BETINGAN R-7	7702.38	14072.9	5618.18	15485.5	16693.68	4054.25	63721.99
QUEBRACHO ATO	14247.38	17907.3	21069.2	23702.8	17687.55	4460.81	99075.44
TANGAN OS	7080	11240.27	7949.45	19069.94	12975.52	3354.37	61709.75
TANGAN PR	1880.56	4074.7	9736.24	12051.13	9774.44	126.24	37901.31
LEUKOTAN 1084	9069.97	11513.6	18177	22414.33	12813.95	4557.85	78544.21
TRUPON DKA	14178.98	16962.7	23006.05	23118.11	17095.65	4388.39	98949.88
TRUPONDI IMP	952.92	1802.5	1783.15	907.71	224.56	143.45	5914.29
TRUPON DSL	2209.89	3783.04	3077.3	1948.7	670.04	152.31	11841.29
TRUPON PEM	3311.21	5440.56	7150.75	9505.47	11835.22	3203.21	40444.42
TRUPON PA11	1426.14	2096.2	4006.82	5964.89	3763.3	911.81	18163.16
TRUPON DB-80	2296.39	3606.73	1602.56	3937.12	3308.16	469.1	15220.06
FORMIATO DE SODIO	9064.33	14059.18	17338.65	21825.95	20559.08	5269.73	88647.02
ANILINA NEGRO TRUPOCOR T	575.72	547.01	1020.12	1391.02	2911.29	430.2	6875.36
ANILINA NEGRO TRUPOCOR MI	4304.44	6867.8	11928.92	12513.71	8564.89	2285.85	46465.61
ANILINA AMARILLO TRUPOCOR	8.81	16.51	7.2	0.4	11.34	44.26	84.26
ANILINA HABANA TRUPOCOR D	29.78	96.48	33.11	9.91	3.34		172.52

Fig. 6.89: Reporte Histórico de pedidos
Capturado por: Ing. Elena Sánchez

Un historico de productos producidos, vendidos, e incluso reprocesados desde la bodega de producto terminado.

Etiquetas de fila	Estado	Costo
PT	ACTIVO	189,720,925 25,400,206
ALABAMA	INACTIVO	8,892 630
ALASKA	REPROCESO	48,588 1,540
AMERICANO	TRANSFER	1,204 70
ARIZONA		808 70
BERLINES		265,813 700
BERLINES CRAQUELI		5,221 238
BOCHELA		3,505 322
BOSI		13,677 504
BOSTON		5,380 434
BOX		204,784 56
BRASIL		9,344 602
BRASILERO		352 54,508 154
BRUSH OFF		882 85,007 516
BUFALDO		84 15,290 672
BUFALO AMERICANO		9 1,537 126
CAMPERO		71 14,007
COLEGIAL		3,651 492,198 2,021
COLEGIAL ESPECIAL		43,896 8,117,614 3,062
CRAQUELE		96 18,349 630
CRAZY		521 97,660 56
CRISTAL		932 76,299 881

Fig. 6.90: Reporte Histórico de Administración de bodega pt
Capturado por: Ing. Elena Sánchez

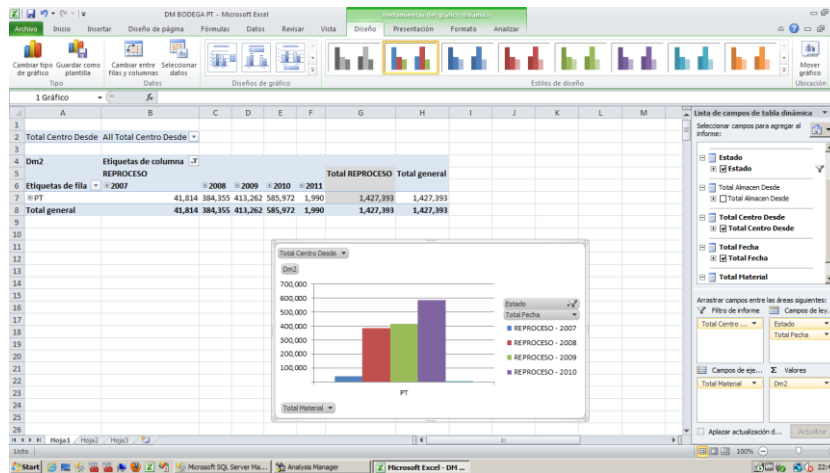


Fig. 6.91: Reporte Histórico de reprocesos

Capturado por: Ing. Elena Sánchez

- ✚ Dada la tardanza del actual sistema de gestión, se requiere un inventario a la fecha especialmente para el área de ventas, es decir un stock actual de producto terminado.

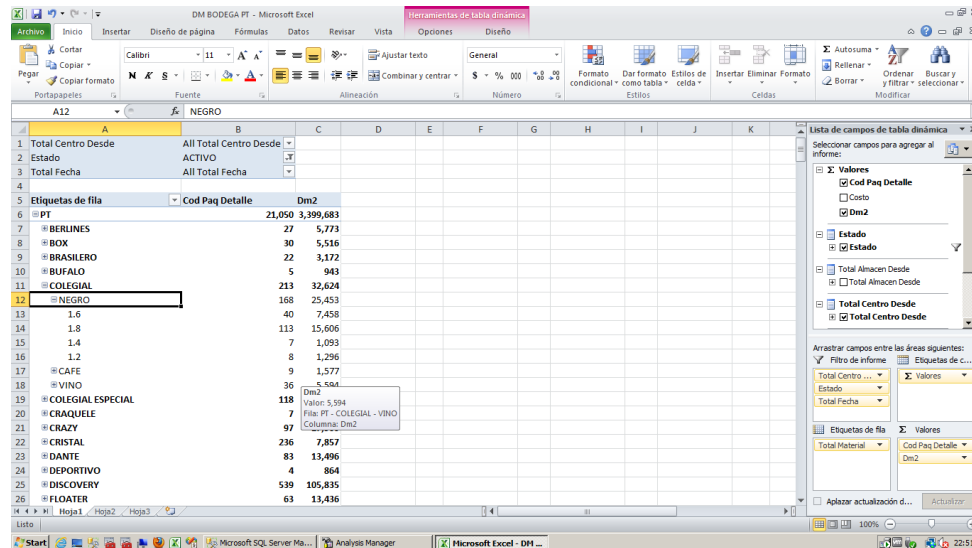


Fig. 6.92: Reporte Stock disponible en bodega

Capturado por: Ing. Elena Sánchez

- ✚ Un histórico de ventas por cliente, por producto donde se pueda apreciar la cantidad vendida y su área.

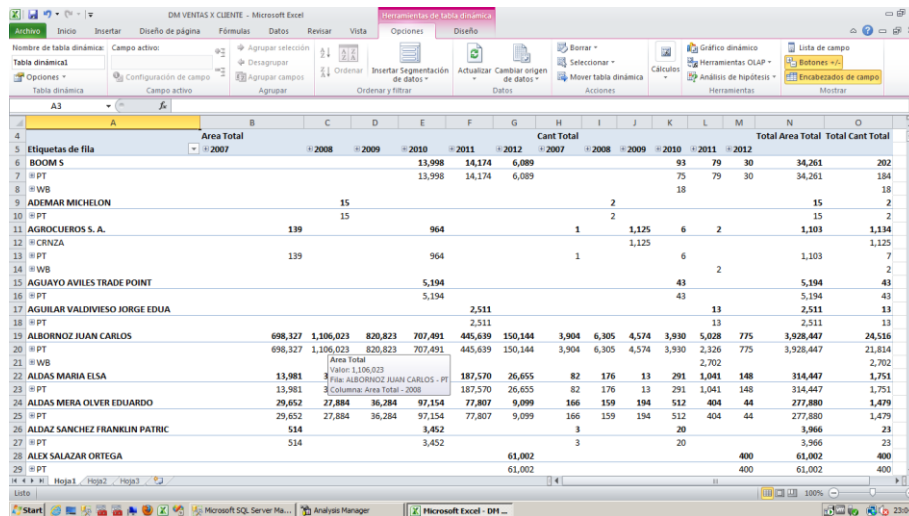


Fig. 6.93: Reporte Histórico de ventas por cliente

Capturado por: Ing. Elena Sánchez

Un histórico de consumos de materia prima por centro de costo

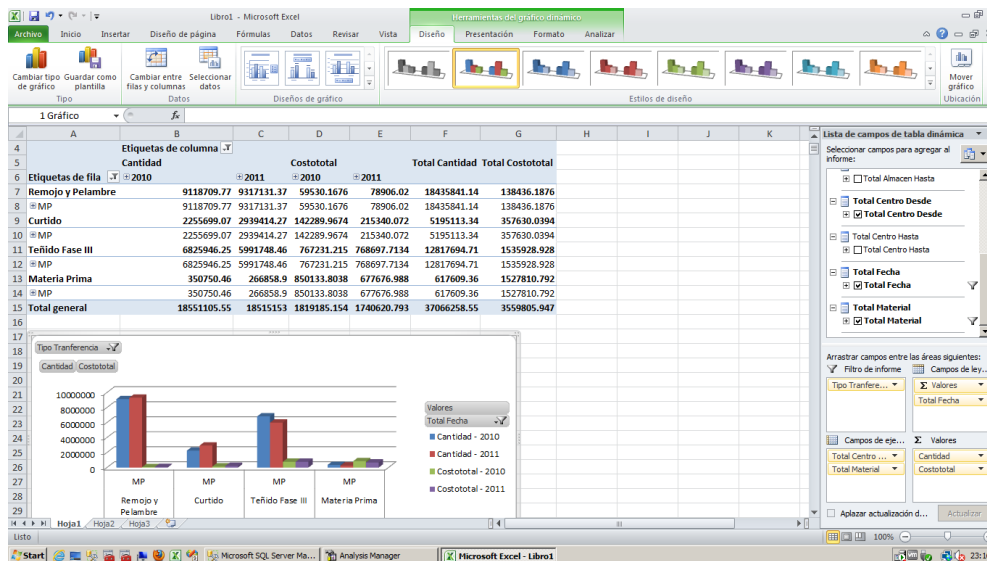


Fig. 6.94: Reporte Histórico de consumos de MP

Capturado por: Ing. Elena Sánchez

Un histórico de ingresos de materiales a los centros de costos correspondientes.

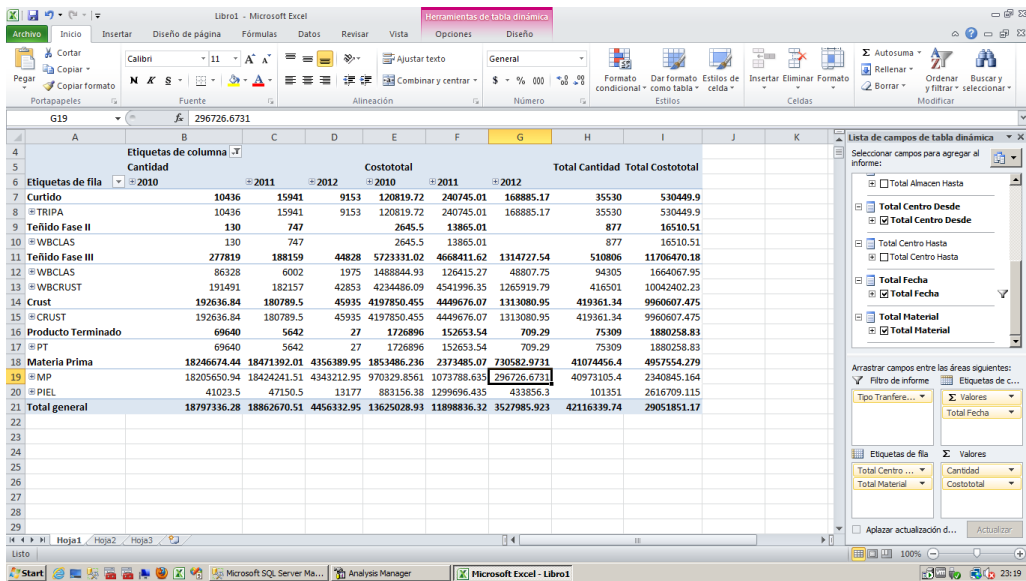


Fig. 6.95: Reporte Histórico de ingresos

Capturado por: Ing. Elena Sánchez

Historicos de los movimientos de préstamos, devoluciones y ventas por centros de costo, sean estos en materia prima como en producto terminado.

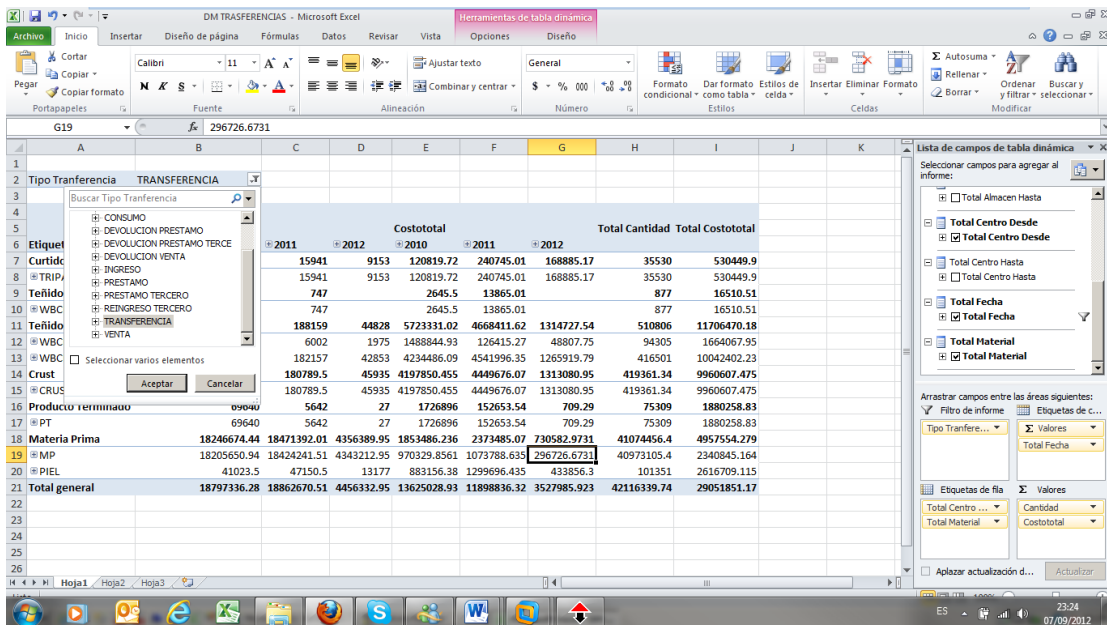


Fig. 6.96: Reporte Histórico de movimientos

Capturado por: Ing. Elena Sánchez

Histórico de costes de fórmulas

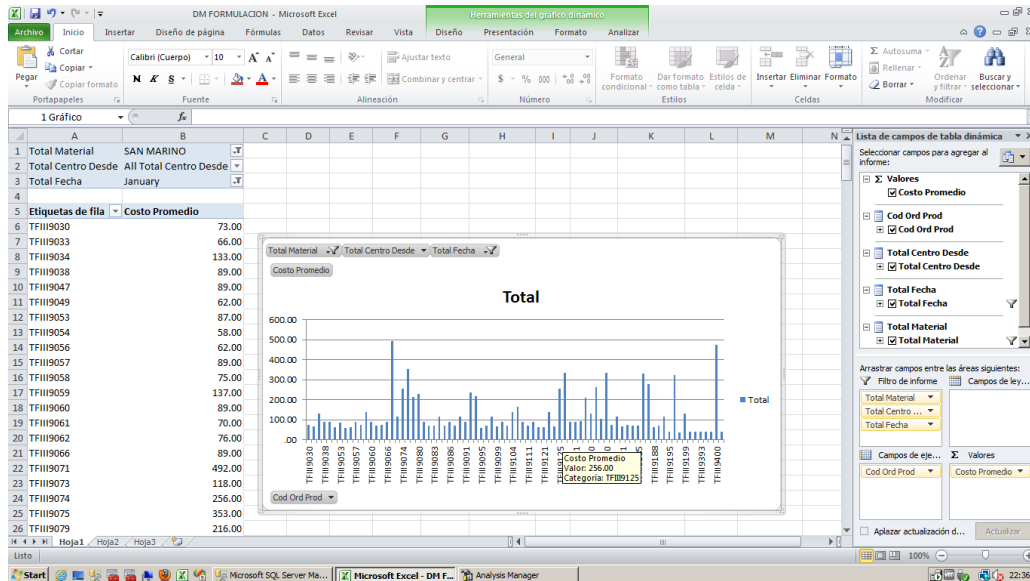


Fig. 6.97: Reporte de costes de formulas

Capturado por: Ing. Elena Sánchez

Panel de Administración de Reporting Services

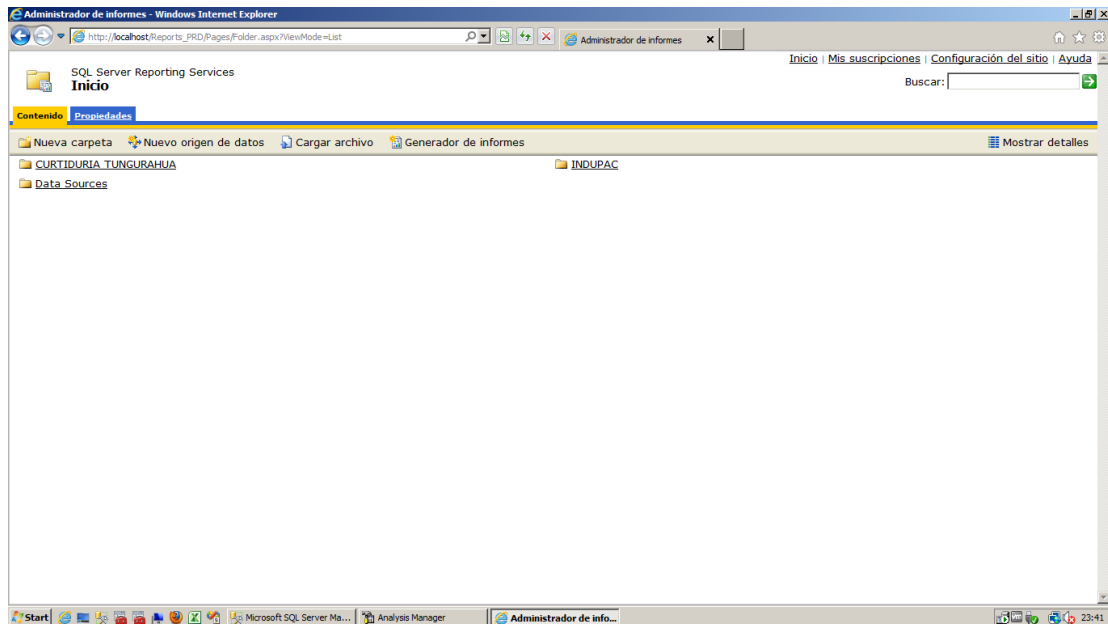


Fig. 6.98: Panel de Administración del Reporting Services

Capturado por: Ing. Elena Sánchez

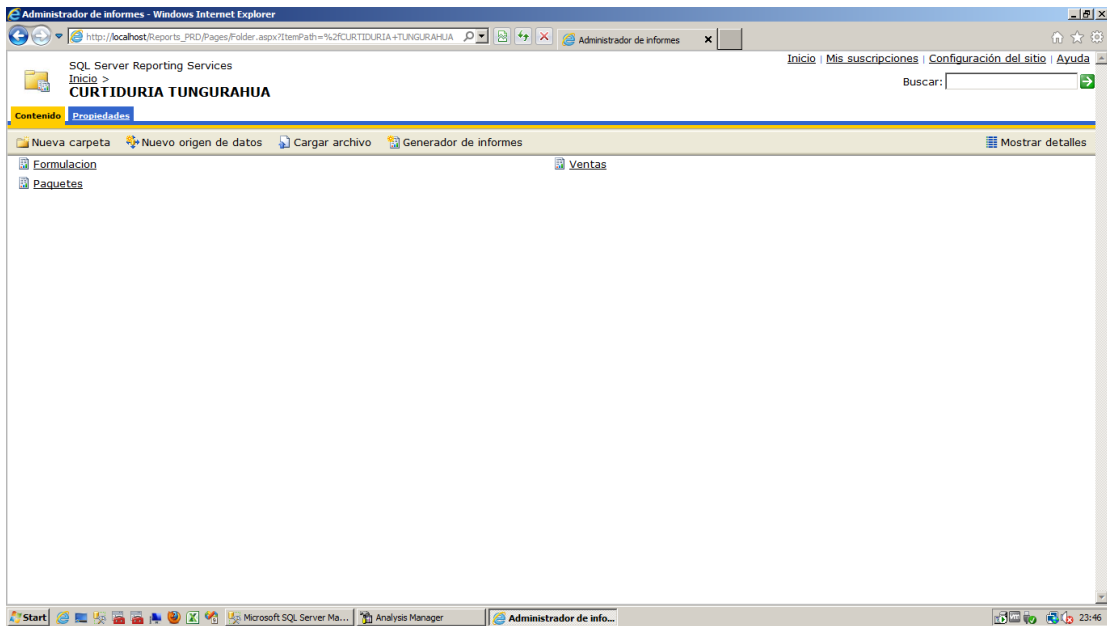


Fig. 6.99: Reporte dentro del reporting services

Capturado por: Ing. Elena Sánchez

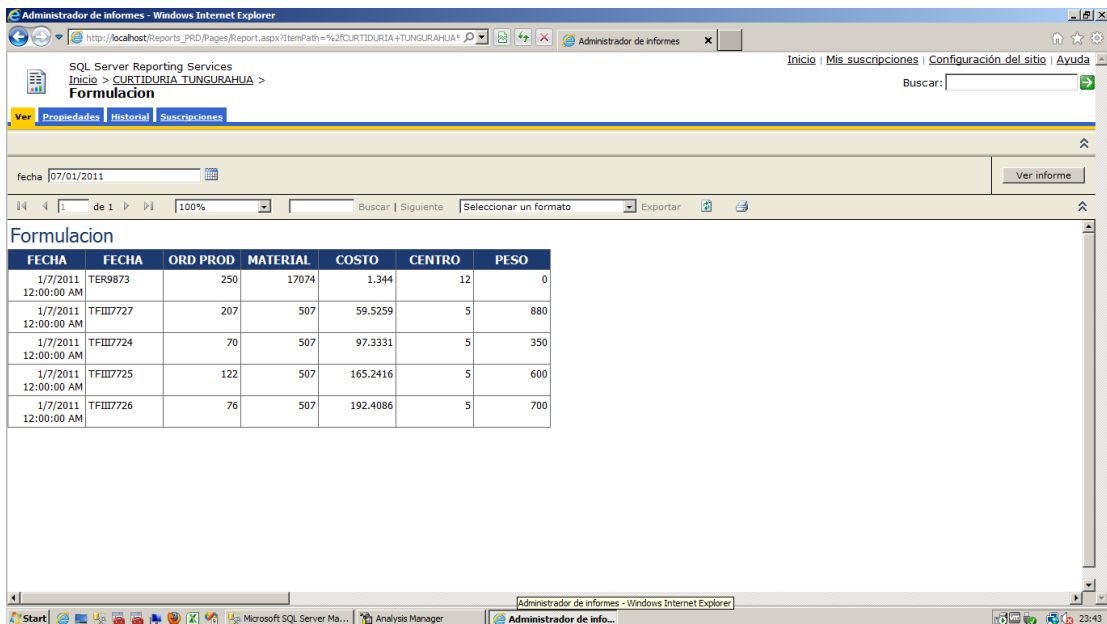


Fig. 6.100: Reporte Histórico de órdenes de producción

Capturado por: Ing. Elena Sánchez

Administrador de informes - Windows Internet Explorer

http://localhost/Reporte_Pages/Report.aspx?ItemPath=%2FCURTIDURIA+TUNGURAHUA

Inicio | Mis suscripciones | Configuración del sitio | Ayuda

SQL Server Reporting Services
Inicio > CURTIDURIA TUNGURAHUA >

Ventas

Ver | Propiedades | Historial | Suscripciones

Cliente: 271

1 de 27 | 100% | Buscar | Siguiente | Seleccionar un formato | Exportar

Ventas

FECHA	TIPO STOCK	COD TIPO	TIPO	COD PROVEEDOR	COD CLIENTE	USU CODIGO	COD FAC	TIPO INI	COD TRANSFERENCIA	ALM CODIGO	CEN CODIGO	COD TIPO DET	COD MAT COL CAL
1/9/2007 12:00:00 AM	PRIMERA	1010	DEVOL		271	DARAUJO		VENTA	6452	16	13	7212	75
3/7/2007 12:00:00 AM	PRIMERA	444	VENTA		271	11			2244	16	13	2557	75
3/7/2007 12:00:00 AM	PRIMERA	444	VENTA		271	11			2244	16	13	2558	75
3/7/2007 12:00:00 AM	PRIMERA	444	VENTA		271	11			2244	16	13	2559	75
3/7/2007 12:00:00 AM	PRIMERA	444	VENTA		271	11			2244	16	13	2561	75
1/9/2007 12:00:00 AM	PRIMERA	1011	DEVOL		271	DARAUJO		VENTA	6453	16	13	7213	75
1/9/2007 12:00:00 AM	PRIMERA	1012	DEVOL		271	DARAUJO		VENTA	6456	16	13	7214	75
3/7/2007 12:00:00 AM	PRIMERA	444	VENTA		271	11			2244	16	13	2560	75
3/7/2007 12:00:00 AM	PRIMERA	445	VENTA		271	11			2246	16	13	2562	75

Start | Microsoft SQL Server Ma... | Analysis Manager | Administrador de info... | 23:46

Fig. 6.101: Reporte Histórico de clientes

Capturado por: Ing. Elena Sánchez

Plan de Capacitación y Soporte a Usuarios

Para la correcta utilización de la información en los cubos es necesario educar a los usuarios con las nuevas características del sistema.

Uno de los recursos importantes en una empresa lo conforma el personal implicado en las actividades laborales por lo tanto un equipo con el conocimiento apropiado ejecutara sus procesos de manera eficiente para lograr esto se define el siguiente plan de capacitación.

Actividad

Obtención de reportes del sistema denominado Datawarehouse

Alcance

El presente plan de capacitación es aplicado a los siguientes grupos conformados por las siguientes personas y/o función que desempeña:

Grupo # 1

ÁREA: PRODUCCION	
INTEGRANTE	FUNCION
Ricardo Callejas	Gerente de Producción
Patricio Lara	Jefe de Planta
Verónica Sánchez	Coordinador de Ribera
Fernando Valenzuela	Coordinador de Teñido-Acabado

Tabla 6.2: Grupo de capacitación Producción
Elaborado por: Ing. Elena Sánchez

Grupo # 2

ÁREA: VENTAS	
INTEGRANTE	FUNCION
Lenin Soria	Coordinador Ventas
Jazmín Villacís	Cobranzas
Alex Villacís	Asistente Comercial
Gilberto Zapata	Vendedor Senior
Julio Pinto	Vendedor Senior

Tabla 6.3: Grupo de capacitación Ventas
Elaborado por: Ing. Elena Sánchez

Grupo # 3

ÁREA: LOGISTICA DE DESPACHOS	
INTEGRANTE	FUNCION
Nelson Chimborazo	Adm. de bodega de PT
Wilmer Rosas	Asistente de Logística
Alex Villacís	Asistente Comercial

Tabla 6.4: Grupo de capacitación Logística de despachos
Elaborado por: Ing. Elena Sánchez

Grupo # 4

ÁREA: LOGISTICA DE ABASTECIMIENTO	
INTEGRANTE	FUNCION
Patricio Callejas	Coordinador Logística
Paulina Sisalema	Asistente Compras
Israel Canchignia	Jefe de Bodegas

Tabla 6.5: Grupo de capacitación Logística de abastecimiento
Elaborado por: Ing. Elena Sánchez

Objetivo de la planificación

Impulsar la eficacia organizacional, elevando el nivel de rendimiento de los colaboradores.

Estrategias

- ✚ Metodología exposición – dialogo
- ✚ Taller personal

Nivel de capacitación

El personal conoce la herramienta básica que es Microsoft Excel 2010, se orientara al personal al uso específico de herramientas OLAP y tablas dinámicas ampliando sus conocimientos y perfeccionando sus habilidades en el manejo de esta herramienta.

Acciones

Las acciones para el desarrollo del plan de capacitación están respaldadas con los temas a impartirse:

Herramienta a usar Microsoft Excel

Contenido para exposición (Ver Anexo 4)

1. Definiciones generales
 - 1.1 Inteligencia de negocios, Datawarehouse
2. Configurar una consulta de Datos
 - 2.1 crear, editar y administrar conexiones a datos externos
3. Obtención y manipulación de un reporte a través de cubos OLAP

2.1 Datos de origen OLAP

2.2 Recuperar datos OLAP

2.3 Tipos de campos

2.4 Cálculos

2.5 Subtotales

Cronograma

	SEPTIEMBRE																			
	SEMANA 1					SEMANA 2					SEMANA 3					SEMANA 4				
	3	4	5	6	7	10	11	12	13	14	17	18	19	20	21	24	25	26	27	28
Grupo # 1	■																			
Grupo # 2						■														
Grupo # 3											■									
Grupo # 4																■				

Tabla 6.6: Cronograma de Capacitación
Elaborado por: Ing. Elena Sánchez

CONCLUSIONES Y RECOMENDACIONES

CONCLUSIONES

- ✚ El trabajo de tesis presenta una solución que Curtiduría Tungurahua S.A va implementar para satisfacer sus necesidades de análisis de información para el apoyo en la toma de decisiones, presentando información en línea que refleja el estado de la empresa.

- ✚ Se concluye que es de mucha importancia el desarrollo de la fase de análisis para evitar que a lo largo del proyecto se presenten necesidades de reestructuración de los procesos, incidencias que causan retrasos en la presentación del proyecto, algunas inconsistencias saltan a la vista cuando ya se obtiene el reporte es aquí donde se debe regresar a procesos anteriores para resolverlos.

- ✚ Un proceso puede estar bien definido y estructurado, sin embargo existieron otros factores como datos “libres“ por llamarlos así ya que son incoherentes pero que no podemos modificarlos ya que no hay información al respecto más que el sistema transaccional actual de la empresa como fuente, estos datos se identificaron durante la implantación del sistema.

- ✚ Con los reportes obtenidos como fruto de la investigación se satisfacen los objetivos planteados en el inicio de la propuesta, sin embargo si la empresa posee información adicional que desearía obtener para su análisis, es posible agregar nuevos procesos etl y/o tablas, ya que el presente trabajo deja abierta esta posibilidad.

- ✚ El presente trabajo investigativo proporciona que los datos estén almacenados ordenadamente en el datawarehouse posibilitando el acceso directo a través de reportadores incluso sql a las tablas de hechos para obtener información personalizado, que no se hayan definido en los Datamart o incluso facilita el desarrollo de nuevos datamarts.

- ✚ Se concluye que el resultado de la experiencia en la empresa Curtiduría Tungurahua S.A como positiva pues se lograron alcanzar los objetivos propuestos obteniendo un beneficio para la empresa y personal ya que se obtuvo el conocimiento para hacer grandes aportes acercando las tecnologías a la empresa.

RECOMENDACIONES

- ✚ Se recomienda al futuro investigador identificar los datos “libres” es decir incoherentes durante la fase de análisis, lo que nos proporcionara un ahorro de tiempo y recursos en el proyecto.
- ✚ Se recomienda al departamento de RRHH planificar siempre una inducción con quien corresponda en el uso de la herramienta para que el personal nuevo pueda usarlo eficientemente, ya que aunque el sistema esté bien estructurado si no se lo utiliza de manera correcta puede fracasar el producto final es decir los reportes.
- ✚ Algunos procesos, tanto en el ETL como en el reporte especialmente de formulaciones, presentan particularmente datos difíciles de solucionar, ya que no existe información confiable desde la fuente, y al tratarse de un histórico es imposible reestructurar esta información, por lo que se recomienda al investigador en la fase de análisis determinar si la utilización de esta información es o no de importancia para el proceso.

Bibliografía

- Business Intelligence. (21 de Octubre de 2009). *Scribd*. Recuperado el 10 de Abril de 2012, de Scribd: <http://es.scribd.com/doc/21355363/Que-es-Bussines-Inteligence>
- Chaudhuri, S., & Dayal, U. (1997). *Data warehousing and Olap for decision support*. ACM SIGMOD Record.
- Contraloria General del Estado, N. V. (abril de 2010). *Controlaroria General del Estado - Normatividad Vigente*. Recuperado el 27 de noviembre de 2012, de Contraloria General del Estado - Normatividad Vigente: http://www.contraloria.gob.ec/documentos/normatividad/constitucion_de_bolsillo.pdf
- Datawarehouse Extracción y Trasformación. (2008). *Extracción y Transformación de los Datos en Data Warehouse*. Recuperado el Mayo de 2012, de <http://www.programatium.com/manuales/Data-Warehouse/estraccion.htm>
- Dimas, Y., & Taborda, C. (2007). *Construcción de cubos de información multidimensionales con OLAPX*. Obtenido de <http://www.monografias.com/trabajos61/cubos-informacion-olapx/cubos-informacion-olapx2.shtml>
- Hackathorn, R., & White, C. (26 de Junio de 2007). *Data Warehouse Appliances: Evolution or Revolution?* Recuperado el Junio de 2012, de <http://www.beyerresearch.com/study/4639>
- IBM, S. L. (2002). *Fundamentals of Data Warehouse and Business Intelligence for Knowledge Management*. IBM Global Service Group.
- INEI, I. N. (s.f.). *Manual de Construcción de un Data Warehouse*. Recuperado el Junio de 2012, de <http://www.ongei.gob.pe/publica/metodologias/Lib5084/INDEX.HTM>

- Inmon, W. H. (1992). *Building the Data Warehouse*. California: Wiley and Son.
- IT News. (2011). Obtenido de IT News: <http://itnews.ec/marco/000040.aspx>
- Kimball, R. (1996). *The Data Warehouse Toolkit*. California: Wiley.
- Kimball, R., & Ross, M. (2002). *The Data Warehouse Toolkit: The Complete Guide to Dimensional Modeling (2da. Edicion)*. California: Wiley.
- Martinez Orol, A. (2010). *olap online analytic processing*. Obtenido de <http://www.gestiopolis.com/canales8/ger/olap-online-analytic-processing.htm>
- Microsoft Books Online. (Junio de 2000). *SQL Server 2000 Books Online -Download Center*. Obtenido de <http://www.microsoft.com/en-us/download/details.aspx?id=18819>
- Microsoft Office OLAP. (2010). *Información general sobre el procesamiento analítico en línea (OLAP)*. Obtenido de <http://office.microsoft.com/es-es/excel-help/informacion-general-sobre-el-procesamiento-analitico-en-linea-olap-HP010177437.aspx>
- Microsoft SQL Server Data Warehousing Training, K. (2000). *Microsoft SQL Server Data Warehousing Training Kit*.
- Pérez, H., & Jiménez, D. (2011). *DATAWAREHOUSE*. Recuperado el Junio de 2012, de [DATAWAREHOUSE: hpez1295.files.wordpress.com/2011/11/datawarehouse-harol2.docx](http://DATAWAREHOUSE:hpez1295.files.wordpress.com/2011/11/datawarehouse-harol2.docx)
- Rodriguez Sanz, M. (22 de Julio de 2010). *Metodologías de Desarrollo de un datawarehouse*. Recuperado el Julio de 2012, de <http://es.scribd.com/doc/52203545/14/Metodologias-de-desarrollo-de-un-Data-Warehouse>
- SAS Institute, I. (2001). En *Rapid Warehouse Methodology*. EEUU: SAS Institute White Paper.

- Todo BI. (07 de Octubre de 2007). *Todo BI:Business Intelligence*. Recuperado el Junio de 2012, de <http://todobi.blogspot.com/2007/10/tipos-de-data-warehouse-appliances.html>
- Toma de Decisiones. (2010). *www.Itescam.edu.mx*. Obtenido de www.itescam.edu.mx/principal/sylabus/fpdb/recursos/r73912.DOC
- True Business, C. G. (2010). *Herramientas de Business Intelligence*. Recuperado el 10 de Abril de 2012, de http://www.truebusiness.com.ec/index.php?option=com_content&view=article&id=17&Itemid=17
- Vaquiroy, J. E. (Septiembre de 2012). *Toma de decisiones*. (J. C. Valda, Editor) Obtenido de Grandes Pymes: <http://jcvalda.wordpress.com/2012/09/23/toma-de-decisiones-2/>
- Vitt, E., Luckevich, M., & Misner, S. (2003). *Business Intelligence: Técnicas de Análisis para la Toma de Decisiones Importantes*. ESPAÑA: McGraw-Hill / INteramericana de España S.A.
- Wrembel, R., & Concilia, C. (2007). *Datawarehouses ans OLAP:Concepts, Architectures and Solutions*. Usa, Hershey, Pennsylvania: IRM.