

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 PLANIFICACIÓN DE LA PRODUCCIÓN Y VENTAS DE LA INDUSTRIA PAPELERA CAICEDO MIÑO CIA. LTDA. DE LA CIUDAD DE AMBATO.

Trabajo de Investigación

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

Nombre del Autor: Ing. Juan Carlos Pérez Salinas

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 PLANIFICACIÓN DE LA PRODUCCIÓN Y VENTAS DE LA INDUSTRIA PAPELERA CAICEDO MIÑO CIA. LTDA. DE LA CIUDAD DE AMBATO.”

Presentada por: Ing. Juan Carlos Pérez Salinas 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 uso y 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. Edwin Hernando Buenaño Valencia
Miembro del Tribunal

Ing. Mg. Franklin Oswaldo Mayorga Mayorga
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 PLANIFICACIÓN DE LA PRODUCCIÓN Y VENTAS DE LA INDUSTRIA PAPELERA CAICEDO MIÑO CIA. LTDA. DE LA CIUDAD DE AMBATO.”**, nos corresponde exclusivamente a Ing. Juan Carlos Pérez Salinas, Autor e Ing. Mg. Elsa Pilar Urrutia Urrutia, Directora del Trabajo de Investigación; y el patrimonio intelectual del mismo a la Universidad Técnica de Ambato.

Ing. Juan Carlos Pérez Salinas

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. Juan Carlos Pérez Salinas

DEDICATORIA

Dedico este proyecto, mi carrera, mis logros y mi vida a Dios por ser quien me a través de sus bendiciones ha permitido encaminarme por la senda del progreso.

A mi amada esposa Carmen Elena, puntal fundamental de mi vida quien ha estado siempre a mi lado incondicionalmente apoyándome intelectual y sentimentalmente además de ser el motivo y la razón de mi esfuerzo y dedicación.

A mis queridos padres a mis Padres Alfonso Isaías Pérez y Gloria Magdalena Salinas que me han sabido inculcar valores como el respeto trabajo y dedicación, apoyándome en todos los proyectos de mi vida, a mis hermanos por compartir mis penas, alegrías y alentarme a seguir en el camino de la superación, primordialmente a mi querida hermana Dra. Gladys Cecilia por su apoyo, ejemplo de trabajo y superación constante.

Juan Carlos

AGRADECIMIENTO

Quiero expresar mis más sinceros agradecimientos a todas las personas que contribuyeron a la culminación de este proyecto, en especial a mi Tutora Ing. Mg. Pilar Urrutia, por su calidad humana, conocimiento, y muy acertada dirección.

Mi eterno agradecimiento al Ing. M.Sc. Oswaldo Paredes en su calidad de Decano y Presidente de Postgrado, por el apoyo moral e intelectual de forma desinteresada como es su característica en todos los actos que diariamente realiza en beneficio y engrandecimiento de nuestra querida Facultad.

Un agradecimiento especial a la Universidad Técnica de Ambato y a la Facultad de Ingeniería en Sistemas, Electrónica e Industrial y por su intermedio al noble cuerpo docente por compartir sus conocimientos.

Juan Carlos

ÍNDICE GENERAL

PRELIMINARES	PÁGINAS
Portada	i
Autoría de la Investigación	iii
Derechos de Autor	iv
Dedicatoria	v
Agradecimiento	vi
Índice General	vii
Índice de Figuras	x
Índice de Tablas	xiii
Resumen Ejecutivo	xiv
Introducción	xvi
CAPÍTULO I	
EL PROBLEMA	1
Planteamiento del problema	1
Contextualización	1
Árbol de Problemas	4
Análisis Crítico	5
Prognosis	6
Formulación del Problema	7
Preguntas Directrices	7
Delimitación de la Investigación	7
Delimitación Espacial	8
Delimitación Temporal	8
Unidades de Observación	8
Justificación	9
Objetivos	10
Objetivo General	10
Objetivos Específicos	10
CAPÍTULO II	
MARCO TEÓRICO	11
Antecedentes Investigativos	11
Filosófica	13
Tecnológica	13
Administrativa	13

Legal	14
Categorías Fundamentales	15
Red de Inclusiones Conceptuales	16
Constelación de Ideas de la Variable Independiente	17
Constelación de Ideas de la Variable Dependiente	18
Hipótesis	48
Variable Independiente	48
Variable Dependiente	48

CAPÍTULO III

METODOLOGÍA	49
Modalidades básicas de investigación	47
Investigación de Campo	48
Investigación Documental-bibliográfica	48
Tipos o niveles de Investigación	48
Asociación de Variables	24
Población y Muestra	48
Operacionalización de Variables	50
Técnicas e Instrumentos	52
Validez y Confiabilidad	52
Plan de Recolección de Información	52
Plan de Procesamiento de Información	53

CAPÍTULO IV

ANÁLISIS E INTERPRETACIÓN DE RESULTADOS	59
Verificación de la Hipótesis	68
Prueba de Hipótesis	72

CAPÍTULO V

CONCLUSIONES Y RECOMENDACIONES	77
Conclusiones	78
Recomendaciones	79

CAPÍTULO VI

LA PROPUESTA	80
Antecedentes de la Propuesta	81
Justificación	83

Objetivos	85
Objetivo General	85
Objetivos Específicos	85
Análisis de Factibilidad	86
Fundamentación	87
METODOLOGÍA	103
Análisis	104
Evaluación de la Organización	104
Evaluación de la Infraestructura de la Organización	104
Planeamiento del proyecto	106
Análisis de las fuentes de datos	106
Definición del plan	109
Diseño	110
Modelo Dimensional	110
Arquitectura	112
Diseño Proceso ETL	113
Extracción	115
Transformación y Carga	119
Construcción	124
Construcción cubos de información	133
Optimización para mejor desempeño de las consultas	137
Construcción DTS definitivo	140
Creación Job para Carga Automática	141
Diseño del Reporting Services	141
Implementación	145
Configuración consulta de datos	145
Reportes	150
Reportes en línea	159
Plan de capacitación a usuarios	163
CONCLUSIONES	168
RECOMENDACIONES	169
BIBLIOGRAFÍA	170
ANEXOS	174

ÍNDICE DE FIGURAS

Figura N.- 1.1. Árbol de Problemas	4
Figura N.- 2.1. Red de Inclusiones Conceptuales	18
Figura N.- 2.2. Constelación Ideas Variable Independiente	17
Figura N.- 2.3. Constelación Ideas Variable Dependiente	18
Figura N.- 2.4. Diagrama Data Mart	22
Figura N.- 2.5. Diagrama Cubo de Información	23
Figura N.- 2.6. Diagrama Metadatos	27
Figura N.- 4.1. Análisis gráfico de datos Pregunta 1	58
Figura N.- 4.2. Análisis gráfico de datos Pregunta 2	59
Figura N.- 4.3. Análisis gráfico de datos Pregunta 3	60
Figura N.- 4.4. Análisis gráfico de datos Pregunta 4	61
Figura N.- 4.5. Análisis gráfico de datos Pregunta 5	62
Figura N.- 4.6. Análisis gráfico de datos Pregunta 6	63
Figura N.- 4.7. Análisis gráfico de datos Pregunta 7	64
Figura N.- 4.8. Análisis gráfico de datos Pregunta 8	65
Figura N.- 6.1. Procesos Datawarehouse	79
Figura N.- 6.2. Procesos Construcción Datawarehouse	82
Figura N. 6.3. Metodología de Kimball	99
Figura N. 6.4. Metodología Hefesto	100
Figura N.- 6.5. Diagrama Metodología	103
Figura N.- 6.6. Diagrama Estructural INPAC	105
Figura N.- 6.7. Tablas usadas de la Base de Datos Transaccional	107
Figura N.- 6.8. Diagrama Tablas Base de Datos Transaccional	107
Figura N.- 6.9. Data Mart Cobranza	110
Figura N.- 6.10. Data Mart Pedidos	110
Figura N.- 6.11. Data Mart Producción	111
Figura N.- 6.12. Data Mart Ventas	111
Figura N.- 6.13. Arquitectura Datawarehouse	112
Figura N.- 6.14. Procesos ETL	114
Figura N.- 6.15. Base de Datos Fuente	115
Figura N.- 6.16. Propiedades Origen	116
Figura N.- 6.17. Propiedades Destino	116
Figura N.- 6.18. Propiedades SQL origen	117
Figura N.- 6.19. Selección Tablas Destino	117
Figura N.- 6.20. Propiedades Transform Data Task	118
Figura N.- 6.21. Sentencias de vaciado de datos	118
Figura N.- 6.22. Diseño de la Extracción	119
Figura N.- 6.23. Creación de las Bases de Datos	119
Figura N.- 6.24. Base Repositorio de Paso	120
Figura N.- 6.25. Diseño del DW_DWH	120
Figura N.- 6.26. Creación base resumen para cubos de información	121
Figura N.- 6.27. Procedimientos ETL creados	121
Figura N.- 6.28. Creación de procedimientos de carga	122

Figura N.- 6.29. Creación procedimientos carga de Dimensiones	122
Figura N.- 6.30. Diagrama de Procesos ETL	123
Figura N.- 6.31. Desarrollo Data Mart desde Analysis Services	124
Figura N.- 6.32. Conexión al servidor de base de datos	125
Figura N.- 6.33. Configuración Enlace de Datos	125
Figura N.- 6.34. Prueba de conexión enlace de datos	126
Figura N.- 6.35. Desarrollo dimensión de tiempo	127
Figura N.- 6.36. Creación esquema tipo estrella	127
Figura N.- 6.37. Selección tabla dimensión	128
Figura N.- 6.38. Selección del tipo de dimensión	128
Figura N.- 6.39. Creación de los niveles de tiempo para la dimensión	129
Figura N.- 6.40. Opciones avanzadas de la dimensión	129
Figura N.- 6.41. Fin del asistente para dimensión	130
Figura N.- 6.42. Creación dimensiones públicas	130
Figura N.- 6.43. Dimensiones Creadas	131
Figura N.- 6.44. Editor de Dimensiones	131
Figura N.- 6.45. Revisión preliminar de la dimensión	132
Figura N.- 6.46. Reconstrucción estructura de la dimensión	132
Figura N.- 6.47. Creación del cubo de información	133
Figura N.- 6.48. Selección Tabla de hechos	133
Figura N.- 6.49. Creación de las medidas	134
Figura N.- 6.50. Especificación de agregados	134
Figura N.- 6.51. Especificación tipo de dato	135
Figura N.- 6.52. Agregado de dimensiones públicas	135
Figura N.- 6.53. Selección de dimensiones para el cubo	136
Figura N.- 6.54. Diseño esquema estrella	136
Figura N.- 6.55. Cubo data Mart Ventas	137
Figura N.- 6.56. Diseño de agregaciones	137
Figura N.- 6.57. Selección tipo de almacenamiento de datos	138
Figura N.- 6.58. Configuración opciones de la agregación	138
Figura N.- 6.59. Procesar cubo de información	139
Figura N.- 6.60. Resumen de procesamiento del cubo	139
Figura N.- 6.61. Diagrama DTS final	140
Figura N.- 6.62. Creación del job para la carga	140
Figura N.- 6.63. Calendarización del job	141
Figura N.- 6.64. Reporting Services	141
Figura N.- 6.65. Selección Fuente de datos para el Reporting Services	142
Figura N.- 6.66. Query Reporting Services	142
Figura N.- 6.67. Tipo de Reporte Reporting Services	143
Figura N.- 6.68. Diseño de tabla Reporting Services	143
Figura N.- 6.69. Selección de estilo Reporting Services	144
Figura N.- 6.70. Ubicación para desplegar el informe Reporting Services	144
Figura N.- 6.71. Configuración consulta de datos	145
Figura N.- 6.72. Elección de origen de datos	145
Figura N.- 6.73. Creación de un nuevo origen de datos	146

Figura N.- 6.74. Conexión almacenamiento de datos multidimensional	146
Figura N.- 6.75. Conexión multidimensional	147
Figura N.- 6.76. Origen de datos desde un cubo	147
Figura N.- 6.77. Elegir origen de datos multidimensional	148
Figura N.- 6.78. Importación de datos a hoja Excel	148
Figura N.- 6.79. Datos importados en hoja Excel	149
Figura N.- 6.80. Información histórica Cobranzas por Vendedores	150
Figura N.- 6.81. Información histórica Recaudación por Región	151
Figura N.- 6.82. Información histórica de Pedidos por Vendedor	152
Figura N.- 6.83. Información histórica de Pedidos por Zona	153
Figura N.- 6.84. Información histórica de pedidos por producto	154
Figura N.- 6.85. Información histórica de la producción	155
Figura N.- 6.86. Información histórica de ventas por vendedor	156
Figura N.- 6.87. Información histórica de ventas por región	157
Figura N.- 6.88. Información histórica de las comisiones por vendedor	158
Figura N.- 6.89. Consultas En línea	159
Figura N.- 6.90. Consultas En línea Pedidos	159
Figura N.- 6.91. Consultas En línea Cliente	160
Figura N.- 6.92. Consultas En línea Producción	160
Figura N.- 6.93. Consultas En línea Ventas	163

ÍNDICE DE TABLAS

Tabla N.- 3.1. Muestreo no probabilístico	51
Tabla N.- 3.2. Operacionalización de la Variable Independiente	52
Tabla N.- 3.3. Operacionalización de la Variable Dependiente	53
Tabla N.- 3.4. Recolección de la Información	55
Tabla N.- 3.5. Formato Encuesta	57
Tabla N.- 3.6. Formato Entrevista	58
Tabla N.- 4.1. Análisis de datos Pregunta 1	60
Tabla N.- 4.2. Análisis de datos Pregunta 2	61
Tabla N.- 4.3. Análisis de datos Pregunta 3	62
Tabla N.- 4.4. Análisis de datos Pregunta 4	63
Tabla N.- 4.5. Análisis de datos Pregunta 5	64
Tabla N.- 4.6. Análisis de datos Pregunta 6	65
Tabla N.- 4.7. Análisis de datos Pregunta 7	66
Tabla N.- 4.8. Análisis de datos Pregunta 8	67
Tabla N.- 4.9. Estado Actual y Esperado de los Procesos	71
Tabla N.- 4.10. Análisis encuesta Pregunta 3	73
Tabla N.- 4.11. Análisis encuesta Pregunta 8	73
Tabla N.- 4.12. Frecuencias Observadas	74
Tabla N.- 4.13. Frecuencias Esperadas	75
Tabla N.- 4.14. Calculo de Ji Cuadrada	75
Tabla N.- 4.15. Procesos Datawarehouse	87
Tabla N.-6.1. Procesos Datawarehouse	96
Tabla N.- 6.2. Descripción Tablas de la Base de Datos Actual	109

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 PLANIFICACIÓN DE LA PRODUCCIÓN Y VENTAS DE LA INDUSTRIA PAPELERA CAICEDO MIÑO CIA. LTDA. DE LA CIUDAD DE AMBATO.”

Autor: Ing. Juan Carlos Pérez Salinas

Tutor: Ing. Mg. Elsa Pilar Urrutia Urrutia

Fecha: 19 de Septiembre del 2012

RESUMEN

El trabajo de investigación que se presenta a continuación investiga cómo influye el datawarehouse en el proceso de las áreas estratégicas de la Industria papelera Caicedo Miño que ha implementado dicha tecnología para tomar decisiones importantes para su futuro e indica mediante su aplicación sus beneficios.

Los principales objetivos de esta investigación se orientan hacia la construcción de un datawarehouse que está dirigido fundamentalmente a responder de manera rápida a las solicitudes de análisis de los directivos de las empresas. Un lugar importante en la construcción del datawarehouse lo ocupa el área de transformación de los datos, la cual varía en su complejidad en dependencia de los requerimientos del problema es así cómo en la actualidad se valora la importancia de contar con esta herramienta.

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 THE
PRODUCTION PLANNING AND SALES IN PAPER INDUSTRY CAICEDO
MIÑO CIA. LTDA. OF AMBATO CITY”

Author: Engineer Juan Carlos Pérez Salinas

Tutor: Engineer Mg. Elsa Pilar Urrutia Urrutia

Date: September 19, 2012

SUMMARY

The research presented below investigates how the datawarehouse influences in the process of the strategic areas of the paper industry Caicedo Miño has implemented the technology to make important decisions for their future through their application indicating their benefits.

The main objectives of this research are oriented towards building a data warehouse that is directed primarily to respond quickly to requests for analysis of business managers. An important place in building the data warehouse it occupies the area of data processing, which varies in complexity depending on the requirements of the problem is how it is now values the importance of this tool.

DESCRIPTORS

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

INTRODUCCIÓN

Desde que se inicia la era de la computadora, las organizaciones usan los datos desde los sistemas operacionales para atender sus necesidades de información. Algunas proporcionan acceso directo a la información contenida dentro de las aplicaciones operacionales. Otras, han extraído los datos desde sus bases de datos operacionales para combinarlos de varias formas no estructuradas, en su intento por atender a los usuarios en sus necesidades de información.

Actualmente y de la manera en que han evolucionado las computadoras, basadas en las tecnologías de información y sistemas. La mayoría de las organizaciones hacen lo posible por conseguir buena información y en el menor tiempo posible, pero el logro de ese objetivo depende fundamentalmente de su arquitectura actual, tanto de hardware como de software. El datawarehouse, es actualmente, el centro de atención de las grandes instituciones que manejan un alto volumen de información, porque provee un ambiente para que las organizaciones hagan un mejor uso de la información que es administrada por diversas aplicaciones operacionales.

Un datawarehouse es una colección de datos en la que se encuentra integrada la información de la Institución y que se usa como soporte para el proceso de toma de decisiones gerenciales u operacionales. El datawarehouse reúne los elementos de datos apropiados desde diversas fuentes de aplicación en un ambiente integral centralizado, simplificando el problema de acceso a la información y en consecuencia, acelera el proceso de análisis, consultas y el menor tiempo de uso de la información. En consecuencia la tecnología datawarehouse, puede hacer más práctica y fácil la explotación de datos para una mayor eficacia del negocio, que no se logra

cuando se usan sólo los datos que provienen de las aplicaciones operacionales, en los que la información se obtiene realizando procesos independientes y muchas veces complejos.

En un datawarehouse los datos extraídos son transformados para eliminar inconsistencias y resumir si es necesario y luego, cargados en el datawarehouse. El proceso de transformar, crear el detalle de tiempo variante, resumir y combinar los extractos de datos, ayudan a crear el ambiente para el acceso a la información institucional. Este nuevo enfoque ayuda a las personas individuales, en todos los niveles de la empresa, a efectuar su toma de decisiones con más responsabilidad lo que delimita su gran importancia.

Descripción de Capítulos que conforman la investigación

EL CAPÍTULO I - EL PROBLEMA:

Dentro de este capítulo está contenido el planteamiento del problema, su contextualización, interrogantes, justificación y objetivos con el fin de clarificar el contexto sobre el cual se va a desarrollar la presente investigación.

EL CAPÍTULO II - MARCO TEÓRICO:

En este capítulo se encuentran los antecedentes investigativos, la base teórica sobre la que se sustentan la tecnología datawarehouse, la fundamentación técnica, la formulación de la hipótesis, el señalamiento de las variables.

El CAPÍTULO III - METODOLOGÍA:

Está incluida en este capítulo la metodología, el tipo de investigación y recolección de información, que será usada en el análisis de datos, verificación de hipótesis, con un análisis de características, además se indica los resultados obtenidos.

El CAPÍTULO IV - ANÁLISIS E INTERPRETACIÓN DE RESULTADOS:

Comprende en este capítulo la descripción de la situación actual y diagnóstico, con un análisis de la recolección de la información de fuentes primarias usando como técnicas encuestas, entrevistas, el análisis y la organización de la información para matemáticamente ser cuantificada.

El CAPÍTULO V – CONCLUSIONES Y RECOMENDACIONES:

Contiene este capítulo las conclusiones y recomendaciones obtenidas luego de analizar los resultados tomando como guía los objetivos planteados en la investigación con los resultados del trabajo realizado.

El CAPÍTULO VI - LA PROPUESTA:

En este capítulo se encuentra la solución justificada, factible y fundamentada al problema, en base a un objetivo general y objetivos específicos, una metodología para la realización del proyecto.

CAPÍTULO I

EL PROBLEMA

Tema: “LA TECNOLOGÍA DATAWAREHOUSE Y SU INCIDENCIA EN LA PLANIFICACIÓN DE LA PRODUCCIÓN Y VENTAS DE LA INDUSTRIA PAPELERA CAICEDO MIÑO CIA. LTDA DE LA CIUDAD DE AMBATO.”

Planteamiento del Problema

Contextualización

Según ROBERT WREMBEL & CHRISTIAN CONCILIA (2007) “Datawarehouse no es un término nuevo si no una vieja rutina con un nombre nuevo. El almacenamiento de datos históricos y análisis de estos para tomar decisiones futuras ya era practicado por los aztecas y mayas en su increíble calendario solar. Incluso los egipcios atesoraban registros de las primaveras con amplios desbordamientos del Nilo, que les permitía saber si el año sería de una buena cosecha o no”.(pág. 12).

El estudio de datos relacionados con la gestión empresarial, empezó cuando todavía la computación no llegaba a dar respuesta a estos problemas. Los directivos estudiaban enormes informes elaborados por los departamentos comerciales y económicos compuestos de varias páginas de datos escrupulosamente resumidos. El avance de la computación ha hecho el trabajo un poco más fácil.

El uso de aplicaciones OLTP (En línea TransactionProcessing) ha traído consigo la recopilación muy rápida de datos que antes era casi imposible obtener, aunque haciendo uso en muchos casos de múltiples sistemas que usan SGBDR(Sistemas Gestores de Bases de Datos Relacionales) diferentes e incompatibles. Esto hace difícil el correlacionar los datos obtenidos desde estos diversos sistemas teniendo que volver al análisis impreso. Así esta nueva teoría viene a resolver un problema viejo usando una nueva técnica: OLAP (En línea AnalyticalProcessing) Procesamiento Analítico En Línea.

Un datawarehouse simplifica los procesos de toma de decisiones porque ofrece imágenes integradas de los datos. Facilita el proceso de comparación, proyección a futuro, relación con otros datos, muestra de indicadores, información consolidada, La tecnología Datawarehouse hace que el acceso a los datos sea fácil y rápido, permitiendo a los usuarios hacer sus propias consultas. Ayuda a mejorar el buen funcionamiento de los sistemas operacionales retroalimentando demandas para los sistemas transaccionales.

En la actualidad Industrial Papelera Caicedo Miño Cía. Ltda., carece de un datawarehouse que permita la gestión de la información de producción, que le permita el control y seguimiento unificado de la producción, compra de materia prima, suministros, tampoco cuenta con una base de datos que optimice estos

procesos, las personas que necesitan conocer del estado de la producción, así como el estado en que se encuentran provistas sus bodegas, tienen dificultades para saber cuánto y cuando pueden contar con estos productos.

Árbol de Problemas

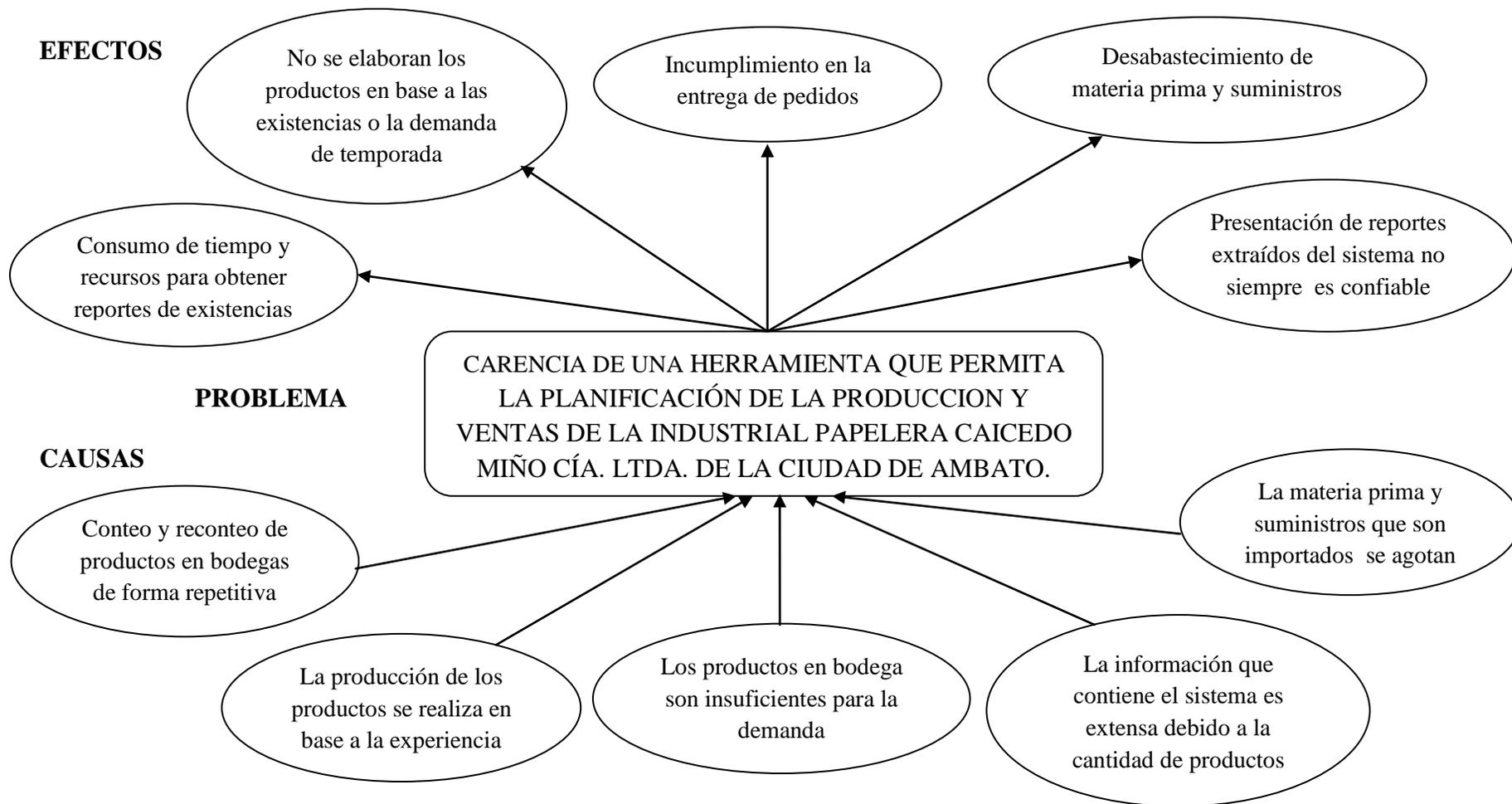


Figura No. 1.1: Árbol de Problemas
Elaborado por: Juan Carlos Pérez

Análisis Crítico

La carencia de un datawarehouse para la Planificación de la Producción y Ventas que aporte positivamente en la toma de decisiones para la planificación anual de la producción, así como de la proyección anual de ventas, ha creado un impacto negativo importante en las finanzas de Industrial Papelera Caicedo Miño, entre la serie de consecuencias tenemos el retraso en las ventas, entrega de pedidos puntualmente a los clientes, desabastecimiento de la materia prima y suministros necesarios para la producción, tomando en cuenta que las materias primas que se usan son importadas desde el extranjero, por lo que toma semanas o meses, además que se experimenta una pérdida excesiva del tiempo de producción.

Al no existir los medios necesarios para presentar de manera correcta la información de producción y ventas, así como la dificultad para obtener esta misma información de manera confiable y consolidada, tiene como consecuencia la pérdida de tiempo, pérdidas en las ventas y retrasos en la producción.

Por otra parte, el control de procesos que actualmente se realizan de manera manual se encuentra en registros impresos en papel, por lo que llevar un control de la información se vuelve un proceso complejo que lleva mucho tiempo, causando que se tomen decisiones de producción no acertadas.

Finalmente la incorrecta aplicación de los flujos de producción, ocasiona el desconocimiento del estado y ubicación de las órdenes de producción dentro de la cadena de elaboración de los productos.

Prognosis

De mantenerse la carencia de un datawarehouse para la Planificación de la Producción y Ventas en Industrial Papelera Caicedo Miño de la ciudad de Ambato, la empresa está en riesgo de que se generen grandes pérdidas de tiempo en la producción y acceso a la información de las ordenes de elaboración de los productos, provocando que la gestión administrativa no pueda desarrollarse de manera óptima y que las ventas se vean reducidas en su eficiencia.

De continuar con el control manual de la producción y venta, documentos secundarios y copias transitorias utilizadas eventualmente para la constancia de las ordenes de producción y venta, hará que en la empresa no existan garantías en la entrega de los productos requeridos por los clientes, y que se invierta tiempo excesivo en establecer el tiempo y el material necesario para la producción de estos productos.

De no dar atención al deficiente seguimiento de las ordenes de producción en cada uno 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.

De persistir en no llevar un control y registro automatizado de las órdenes de producción, se está incidiendo notablemente en la pérdida de tiempo y gastos administrativos, lo que de alguna forma conllevan a la inestabilidad económica y administrativa de la fábrica.

Formulación del problema

¿Cómo incide un Datawarehouse en la Planificación de la Producción y Ventas de la Industrial Papelera Caicedo Miño Cía. Ltda. de la Ciudad de Ambato?

Preguntas Directrices

- ¿Cuáles son los pasos a seguir para estructurar datawarehouse para la Planificación de la Producción y Ventas?
- ¿Cuáles son los problemas que se generan en la gestión de las órdenes de producción en Industrial Papelera Caicedo Miño Cía. Ltda.?
- ¿Cuál es el principal problema de la carencia de un datawarehouse para la Planificación de la Producción y Ventas en la IndustriaPapelera Caicedo Miño Cía. Ltda. de la Ciudad de Ambato?

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 desarrolla en los espacios físicos del departamento de sistemas de Industria Papelera Caicedo Miño Cía. Ltda. de la Ciudad de Ambato, ubicado en la av. Pasteur 10-26 y Grecia.

Delimitación Temporal

Esta investigación se enfoca en el periodo comprendido entre el 01 de enero y el 15 de Junio del 2011.

Unidades de Observación

- Gerencia
- Jefatura de Planta
- Jefatura de Sistemas
- Coordinador del área de Producción
- Supervisores de planta
- Bodega

Justificación

En cuanto a la relevancia del problema es necesario indicar que a nivel de empresas de manufactura y especialmente fábricas papeleras, todavía se manejan procesos de producción de forma manual, lo que no permite que el acceso a la información y ubicación de las ordenes de producción dentro de la cadena de fabricación se lo haga de manera automatizada, 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 el manejo de información para la gestión de la producción.

La realización de este proyecto es importante, ya que en la fábrica se mantiene un gran volumen de documentos que vigila el seguimiento y control de las ordenes de producción, que establecen las compras de materia prima y suministros en cada uno de los centros de producción, es por ello que se hace necesario agilizar una Herramienta para la Planificación de la Producción y Ventas el mismo que incluya una base de datos confiable que permita presentar información de primera mano, oportuna y precisa del estado de las ordenes de producción, materia prima y suministros, estableciendo de esta manera una búsqueda rápida y precisa de la orden de producción deseada por la jefatura de planta.

OBJETIVOS

Objetivo General

Determinar la Incidencia de la implementación de un Datawarehouse en la Optimización de la Planificación de la Producción y Ventas de la Industrial Papelera Caicedo Miño Cía. Ltda. de la Ciudad de Ambato.

Objetivos Específicos

- Identificar las causas de los problemas que se presentan por la carencia de un datawarehouse en el proceso de la Producción y Ventas.
- Determinar la funcionalidad del actual método de planificación de la Producción y Ventas.
- Plantear una solución al problema de carencia de un datawarehouse para la Planificación de la Producción y Ventas.

CAPÍTULO II

MARCO TEÓRICO

Antecedentes Investigativos

Realizando un recorrido por las principales bibliotecas de las Universidades que ofertan la carrera de Ingeniería en Sistemas se encuentra que en la Universidad Técnica de Ambato, específicamente en la Facultad de Ingeniería en Sistemas, no existen trabajos similares, o que se asemejen al tema en cuestión.

En la Biblioteca de la Escuela Politécnica Nacional existe un proyecto que desarrolla un datawarehouse con el tema: “Datawarehouse para el análisis académico de la Escuela Politécnica Nacional”, que sirve de apoyo a la toma de decisiones de nivel directivo de la Escuela Politécnica Nacional, desarrollado por: Valencia Arcos, Janeth del Carmen y Guevara Lenis, Jorge Eduardo, en el mes de Junio del 2007, publicado por la editorial de la Escuela Politécnica Nacional.

El trabajo incluye un resumen de los conceptos básicos relacionados con la teoría de datawarehouse y de información referente a los aspectos académicos de la Escuela. Además, realiza la consolidación de varias fuentes de datos provenientes de un conjunto de carreras tomadas como muestra para este proyecto, limpia y depura estos datos para posteriormente ser cargados a una base de datos que permita proporcionar diversas alternativas de visualización de información estadística e interactiva a los usuarios participantes en la toma de decisiones.

El propósito del proyecto es proveer una herramienta de apoyo al análisis académico mediante el estudio comparativo de la información histórica recopilada y de los índices involucrados en el proceso.

Tesis desarrollada por Abad Rivadeneira, Esteban Mauricio, de la Escuela Politécnica Nacional de la Ciudad de Quito, en el año 2004, publicada por la Editorial de la EPN, con el tema: “Sistema de gestión de inventarios y ventas utilizando tecnología de datawarehouse”, apoyan a la presente tesis.

La empresa moderna, especialmente la de gran tamaño, resulta en lo operativo rehén de la complejidad. Las opciones para la toma de decisiones operativas de la compañía son casi infinitas, por lo que se requieren herramientas cada día más potentes y soportadas por más y mejor información. Este proyecto emplea la información de ventas e inventarios de una organización para analizarlas por medio de un dataMart que permite tomar decisiones oportunas de una fuente de información confiable y segura, publicándola en un portal corporativo. Además, este proyecto contribuye a la compra de inventarios en base a un sugerido de compra que toma el comportamiento histórico de ventas e inventarios.

Filosófica

Para realizar el trabajo de grado el investigador asume los principios del paradigma crítico-propositivo porque cuestiona la manera de hacer investigación y por el contrario plantea una propuesta de solución al problema investigado basado en la existencia de múltiples realidades socialmente construidas.

Este enfoque privilegia la interpretación, comprensión y explicación de los fenómenos sociales en un aspecto de totalidad. La investigación está comprendida con los seres humanos y su crecimiento en comunidad de manera solidaria y equitativa, es por tal motivo que propicia la participación de los actores de la sociedad en calidad de protagonistas.

Tecnológica

Tecnológicamente se considera que es un aporte en beneficio de las industrias de la producción específicamente en la las industrias que se dedican a la manufactura papelera y concretamente sobre el sujetodel problema investigado.

Administrativa

Industrial Papelera Caicedo Miño Cía. Ltda., es una empresa privada que goza con capacidad para realizar los actos que fueren necesarios para el

cumplimiento de sus fines, en la forma y condiciones que determinan la Constitución y las leyes de la república del Ecuador.

Su estructura organizacional se divide en 5 niveles de gestión: Directivo, Ejecutivo; Jefatura de planta, Supervisión y Producción. Dentro del nivel directivo encontramos (Junta de accionistas); el nivel ejecutivo lo conforma la Gerencia, la Presidencia, Recursos Humanos; el nivel de la Jefatura de planta lo integra el jefe de planta; el nivel de Supervisión lo conforman los Supervisores, Financiera, Secretaria General; dentro de los Niveles de Producción encontramos a los Coordinadoras de Producción.

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

Industrial Papelera Caicedo Miño Cía. Ltda. está conformada por una sola dependencia física: La planta de producción y ventas ubicada en las calles Av. Pasteur y Grecia (Sector Cashapamba).

Legal

Para realizar el trabajo de grado se fundamenta en artículo 5 del Reglamento de Control y Administración de Bienes de Industrial Papelera Caicedo Miño S.A., que dice: “Los materiales y suministros que requieran las diferentes unidades que conforman la estructura orgánica, deben ser solicitados a la Junta de Accionistas,

utilizando para el efecto un oficio el mismo que llevará la firma del Jefe de la Unidad solicitante. El oficio de solicitud de bienes se elaborará en original y una copia, y su distribución será la siguiente, Original: Secretaria– Archivo, 1 copia: Unidad Solicitante”. (Ver Anexo 1)

En la Sección octava Ciencia, tecnología, innovación y saberes ancestrales de la constitución de la República del Ecuador que reza en su Art. 385, inciso 3: “El sistema nacional de ciencia, tecnología, innovación y saberes ancestrales, en el marco del respeto al ambiente, la naturaleza, la vida, las culturas y la soberanía, tendrá como finalidad: 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”. (Ver Anexo 2) Fuente: (Contraloría General del Estado, 2010).

Categorías Fundamentales

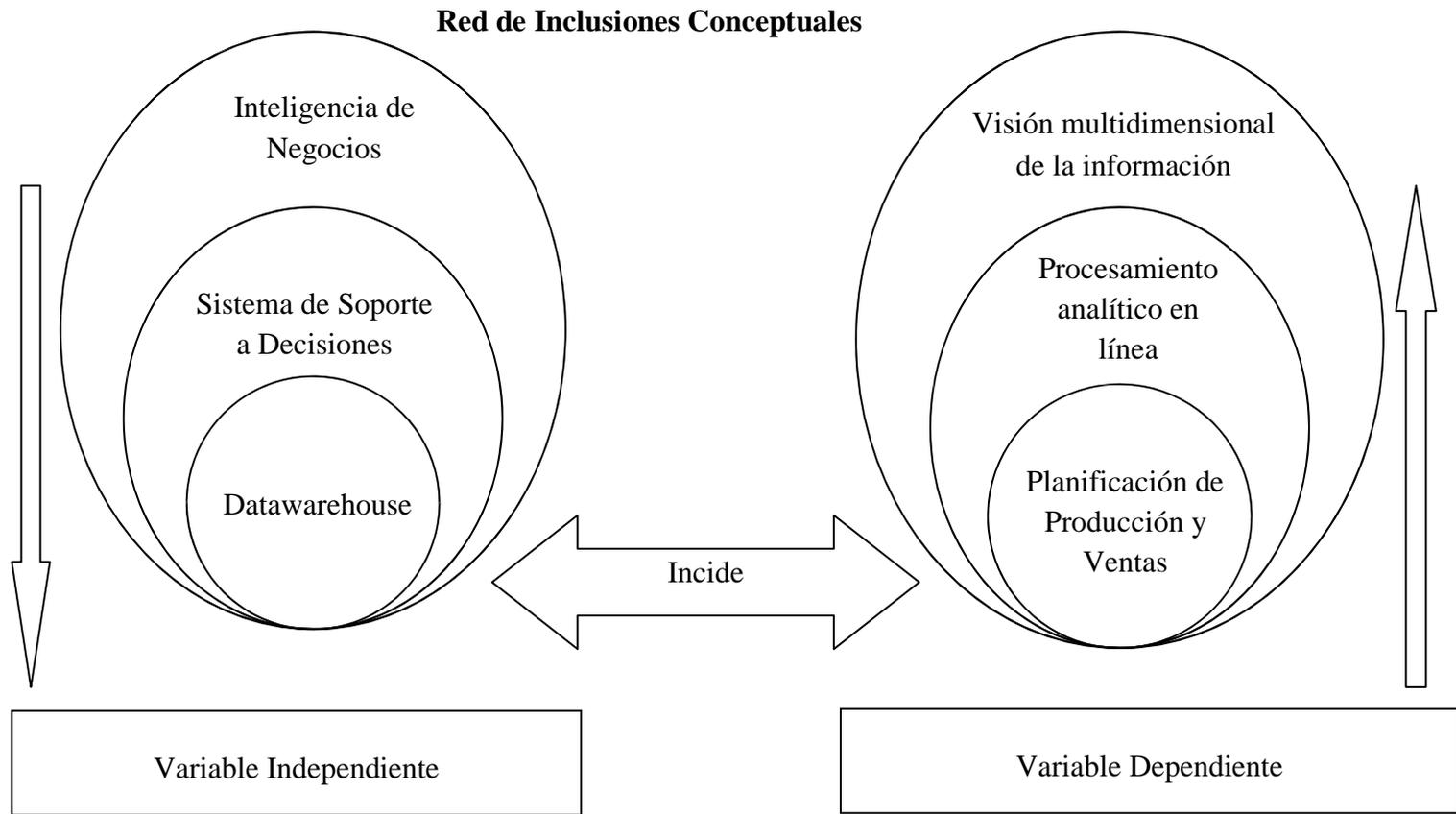


Figura No. 2.1: Red de Inclusiones Conceptuales
Elaborado por: Juan Carlos Pérez

Constelación de Ideas de la Variable Independiente

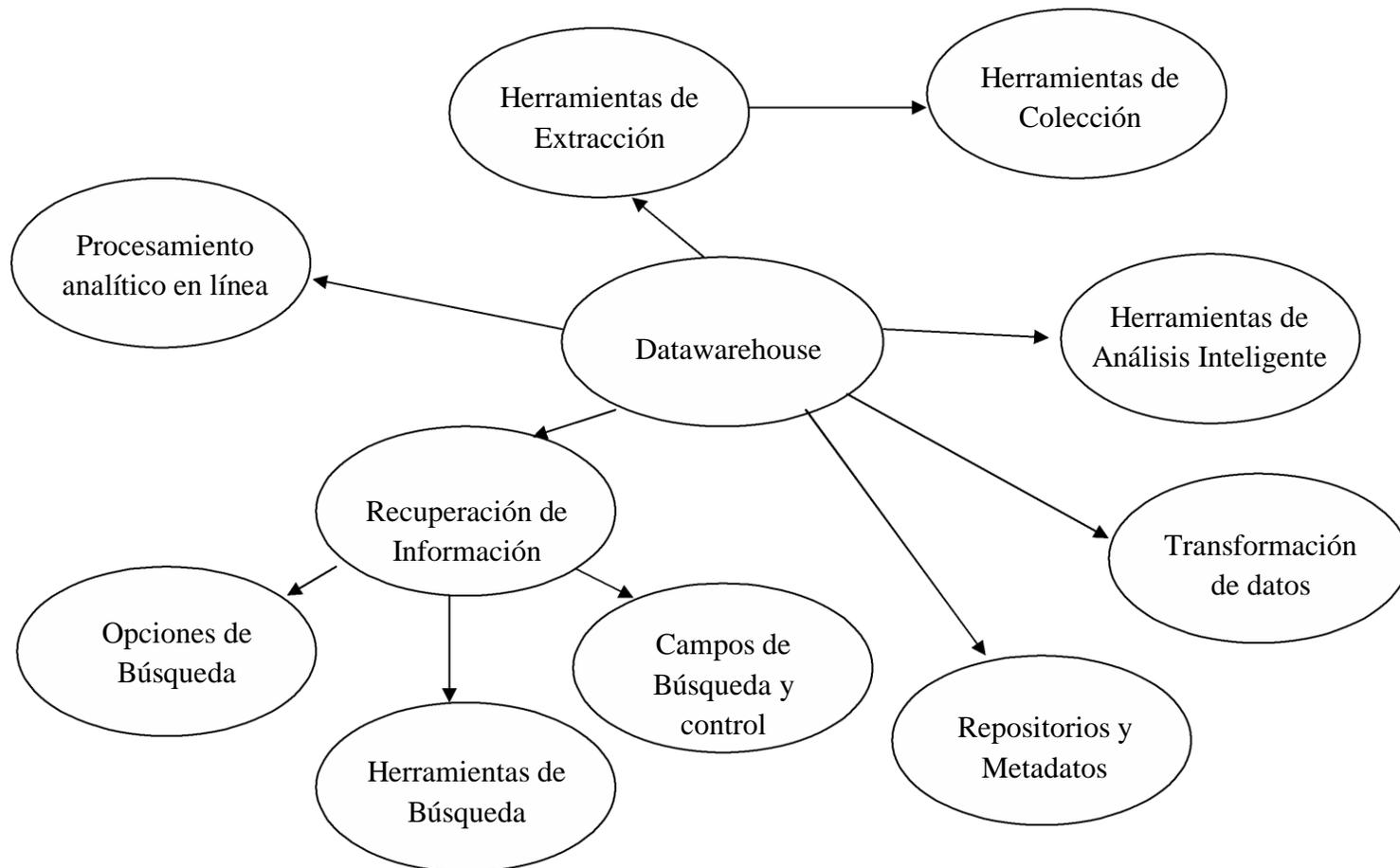


Figura No. 2.2: Constelación Ideas Variable Independiente
Elaborado por: Juan Carlos Pérez

Constelación de Ideas de la Variable Dependiente

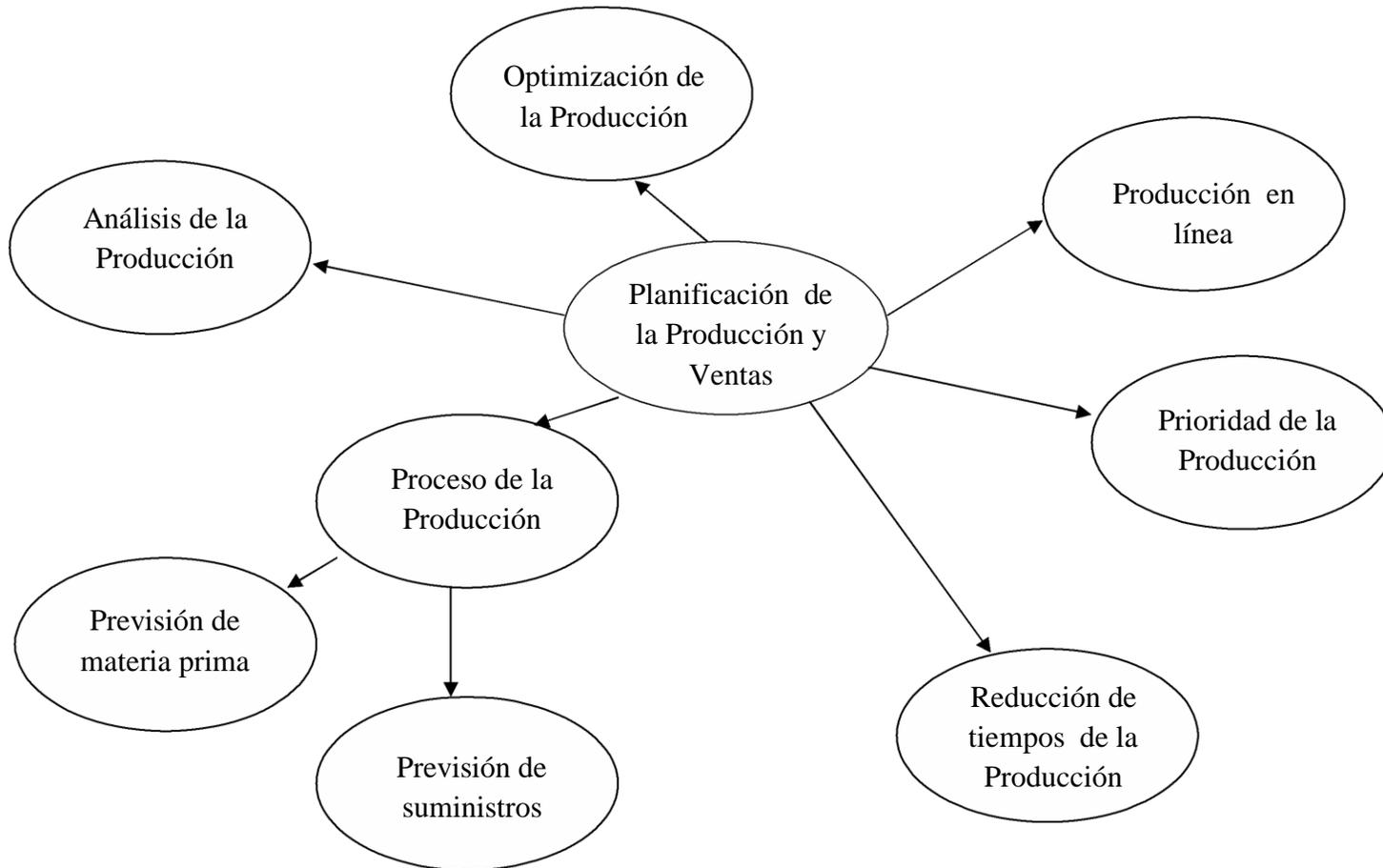


Figura No. 2.3: Constelación Ideas Variable Dependiente
Elaborado por: Juan Carlos Pérez

CATEGORÍAS VARIABLE INDEPENDIENTE

Inteligencia de Negocios

Se denomina inteligencia empresarial, inteligencia de negocios o BI (del inglés *business intelligence*) al conjunto de estrategias y herramientas enfocadas a la administración y creación de conocimiento mediante el análisis de datos existentes en una organización o empresa. (Robles, 2010)

El término inteligencia empresarial se refiere al uso de datos en una empresa para facilitar la toma de decisiones. Abarca la comprensión del funcionamiento actual de la empresa, bien como la anticipación de acontecimientos futuros, con el objetivo de ofrecer conocimientos para respaldar las decisiones empresariales.

Las herramientas de inteligencia se basan en la utilización de un sistema de información de inteligencia que se forma con distintos datos extraídos de los datos de producción, con información relacionada con la empresa o sus ámbitos y con datos económicos.

Mediante las herramientas y técnicas ELT (extraer, cargar y transformar), o actualmente ETL (extraer, transformar y cargar) se extraen los datos de distintas fuentes, se depuran y preparan (homogeneización de los datos) para luego cargarlos en un almacén de datos.

La vida o el periodo de éxito de un software de inteligencia de negocios dependerá únicamente del nivel de éxito del cual haga en beneficio de la empresa que lo usa, si esta empresa es capaz de incrementar su nivel financiero, administrativo y sus decisiones mejoran el accionar de la empresa, la inteligencia de negocios usada estará presente por mucho tiempo, de lo contrario será sustituido por otro que aporte mejores resultados y más precisos.

Por último, las herramientas de inteligencia analítica posibilitan el modelado de las representaciones con base en consultas para crear un cuadro de mando integral que sirve de base para la presentación de informes.

Características

Este conjunto de herramientas y metodologías tienen en común las siguientes características:

- ***Accesibilidad a la información.*** Los datos son la fuente principal de este concepto. Lo primero que deben garantizar este tipo de herramientas y técnicas será el acceso de los usuarios a los datos con independencia de la procedencia de estos.
- ***Apoyo en la toma de decisiones.*** Se busca ir más allá en la presentación de la información, de manera que los usuarios tengan acceso a herramientas de análisis que les permitan seleccionar y manipular sólo aquellos datos que les interesen.
- ***Orientación al usuario final.*** Se busca independencia entre los conocimientos técnicos de los usuarios y su capacidad para utilizar estas herramientas.

Niveles de Realización de la inteligencia de negocios

De acuerdo a su nivel de complejidad se pueden clasificar las soluciones de Business Intelligence en:

- Reportes
- Reportes predefinidos
- Reportes a la medida
- Consultas ("Query") / Cubos OLAP (*On-Line Analytic Processing*).
- Alertas
- Análisis
- Análisis estadístico
- Pronósticos ("Forecasting")
- Modelado Predictivo o Minería de datos ("Data Mining")
- Optimización

Sistemas de Soporte a Decisiones

Durante todo el transcurso de nuestra vida, continuamente tenemos que decidir, por una cosa o la otra, hago esto o lo otro. Y para las empresas es la misma situación, actualmente todo se ha acelerado, la tecnología avanza drásticamente segundo a segundo, el mercado se vuelve muy dinámico e inestable, surge la economía digital, hay más competencia, la información que necesitamos ha aumentado en grandes cantidades, y el tiempo que se tiene para decidir es cada vez menor. (Lara Treviño, 2012)

Es por eso que las organizaciones necesitan herramientas acordes al momento que se está viviendo, para que puedan salir adelante, e ir cambiando en sincronía con el ambiente que las rodea. Una de estas herramientas surge gracias a la tecnología, y es conocida como Sistemas de Soporte a la Decisión o DSS por sus siglas en inglés, que vienen a apoyar un proceso fundamental que es la toma de decisiones.

El concepto de sistema de soporte a las decisiones (DSS por sus siglas en inglés Decision Support System) es muy amplio, debido a que hay muchos enfoques para la toma de decisiones y debido a la extensa gama de ámbitos en los cuales se toman. Estos sistemas de apoyo son del tipo OLAP o de minería de datos, que proporcionan información y soporte para tomar una decisión.

Un DSS puede adoptar muchas formas diferentes. En general, podemos decir que un DSS es un sistema informático utilizado para servir de apoyo, más que automatizar, el proceso de toma de decisiones. La decisión es una elección entre alternativas basadas en estimaciones de los valores de esas alternativas. El apoyo a una decisión significa ayudar a las personas que trabajan solas o en grupo a reunir inteligencia, generar alternativas y tomar decisiones. Apoyar el proceso de toma de decisión implica el apoyo a la estimación, la evaluación y/o la comparación de alternativas. En la práctica, las referencias a DSS suelen ser referencias a aplicaciones informáticas que realizan una función de apoyo.(Lara Treviño, 2012)

Función y características

Los DSS son herramientas de mucha utilidad en Inteligencia empresarial (*Business Intelligence*), permiten realizar el análisis de las diferentes variables de negocio para apoyar el proceso de toma de decisiones de los directivos:

- Permite extraer y manipular información de una manera flexible.
- Ayuda en decisiones no estructuradas.
- Permite al usuario definir interactivamente qué información necesita y cómo combinarla.
- Suele incluir herramientas de simulación, modelización, etc.
- Puede combinar información de los sistemas transaccionales internos de la empresa con los de otra empresa externa.
- Su principal característica es la capacidad de análisis multidimensional (OLAP) que permite profundizar en la información hasta llegar a un alto nivel de detalle, analizar datos desde diferentes perspectivas, realizar proyecciones de información para pronosticar lo que puede ocurrir en el futuro, análisis de tendencias, análisis prospectivo, etc.
- Un DSS da soporte a las personas que tienen que tomar decisiones en cualquier nivel de gestión, ya sean individuos o grupos, tanto en situaciones semiestructuradas como en no estructuradas, a través de la combinación del juicio humano e información objetiva:
- Soporta varias decisiones interdependientes o secuenciales.
- Ofrece ayuda en todas las fases del proceso de toma de decisiones - inteligencia, diseño, selección, e implementación- así como también en una variedad de procesos y estilos de toma de decisiones.
- Es adaptable por el usuario en el tiempo para lidiar con condiciones cambiantes.
- Genera aprendizaje, dando como resultado nuevas demandas y refinamiento de la aplicación, que a su vez da como resultado un aprendizaje adicional.
- Generalmente utiliza modelos cuantitativos (estándar o hechos a la medida).
- Los DSS avanzados están equipados con un componente de administración del conocimiento que permite una solución eficaz y eficiente de problemas muy complejos.

- Puede ser implantado para su uso en Web, en entornos de escritorio o en dispositivos móviles (PDA).
- Permite la ejecución fácil de los análisis de sensibilidad.

Datawarehouse

Definición de Bill Inmon: Bill Inmon fue uno de los primeros autores en escribir sobre el tema de los almacenes de datos, define un datawarehouse (almacén de datos) en términos de las características del repositorio de datos:

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

Inmon defiende una metodología descendente (estrategias de procesamiento de información) a la hora de diseñar un almacén de datos, ya que de esta forma se considerarán mejor todos los datos corporativos. En esta metodología los **Data Marts** se crearán después de haber terminado el datawarehouse completo de la organización.(Inmon, 1992)

Definición de Ralph Kimball: Ralph Kimball es otro conocido autor en el tema de los datawarehouse, define un almacén de datos como: "una copia de las transacciones de datos específicamente estructurada para la consulta y el análisis". También fue Kimball quien determinó que un datawarehouse no era más que: "la unión de todos los Data Marts de una entidad". Defiende por tanto una metodología ascendente a la hora de diseñar un almacén de datos. (Kimball & Margy Ross, 2008)

En el contexto de la informática, un almacén de datos (del inglés *data warehouse*) es una colección de datos orientada a un determinado ámbito (empresa, organización, etc.), integrado, no volátil y variable en el tiempo, que ayuda a la toma de decisiones en la entidad en la que se utiliza. Se trata, sobre todo, de un expediente completo de una organización, más allá de la información transaccional y operacional, almacenado en una base de datos diseñada para favorecer el análisis y la divulgación eficiente de datos (especialmente OLAP, *procesamiento analítico en línea*). El almacenamiento de los datos no debe usarse con datos de uso actual. Los almacenes de datos contienen a menudo grandes cantidades de información que se subdividen a veces en unidades lógicas más pequeñas dependiendo del subsistema de la entidad del que procedan o para el que sea necesario.

Función de un Datawarehouse

En un datawarehouse lo que se quiere es contener datos que son necesarios o útiles para una organización, es decir, que se utiliza como un repositorio de datos para posteriormente transformarlos en información útil para el usuario. Un datawarehouse debe entregar la información correcta a la gente indicada en el momento óptimo y en el formato adecuado. El almacén de datos da respuesta a las necesidades de usuarios expertos, utilizando Sistema de Soporte de Decisiones (DSS), Sistemas de

información ejecutiva (EIS) o herramientas para hacer consultas o informes. Los usuarios finales pueden hacer fácilmente consultas sobre sus almacenes de datos sin tocar o afectar la operación del sistema.

En el funcionamiento de un datawarehouse son muy importantes las siguientes ideas:

- Integración de los datos provenientes de bases de datos distribuidas por las diferentes unidades de la organización y que con frecuencia tendrán diferentes estructuras (fuentes heterogéneas). Se debe facilitar una descripción global y un análisis comprensivo de toda la organización en el datawarehouse.
- Separación de los datos usados en operaciones diarias de los datos usados en el datawarehouse para los propósitos de divulgación, de ayuda en la toma de decisiones, para el análisis y para operaciones de control. Ambos tipos de datos no deben coincidir en la misma base de datos, ya que obedecen a objetivos muy distintos y podrían entorpecerse entre sí.
- Periódicamente, se importan datos al datawarehouse de los distintos sistemas de planeamiento de recursos de la entidad (Planificación de Recursos Empresariales) y de otros sistemas de software relacionados con el negocio para la transformación posterior. Es práctica común normalizar los datos antes de combinarlos en el almacén de datos mediante herramientas de extracción, transformación y carga (ETL). Estas herramientas leen los datos primarios (a menudo bases de datos OLTP de un negocio), realizan el proceso de transformación al almacén de datos (filtración, adaptación, cambios de formato, etc.) y escriben en el almacén.

Data Mart

Un Data Mart 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.(Inmon, 1992)

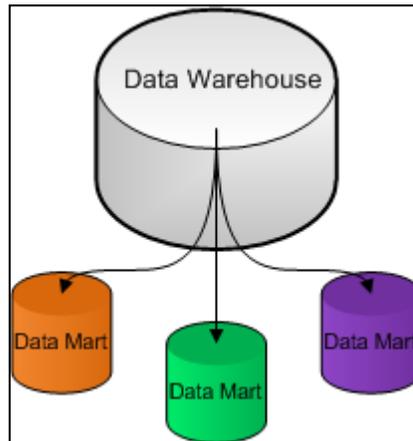


Figura No. 2.4: Diagrama Data Mart

Fuente: <http://anabuigues.com/2010/04/19/data-Mart-y-data-warehouse/>

El Data Mart 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 DataMining 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 Data Marts son pequeños datawarehouse centrados en un tema o un área de negocio específico dentro de una organización. Los Data Marts son subconjuntos de datos de un datawarehouse para áreas específicas. Entre las características de un Data Mart destacan:

- Usuarios limitados.
- Área específica.
- Tiene un propósito específico.
- Tiene una función de apoyo.

Cubos de Información

Un cubo **OLAP**, *OnLine Analytical Processing* o procesamiento Analítico En Línea, término acuñado por Edgar Frank Codd de *EF Codd & Associates*, encargado por Arbor Software (en la actualidad Hyperion Solutions), es una base de datos multidimensional, en la cual el almacenamiento físico de los datos se realiza en un vector multidimensional. Los cubos **OLAP** se pueden considerar como una ampliación de las dos dimensiones de una hoja de cálculo. (Kimball & Margy Ross, 2008).

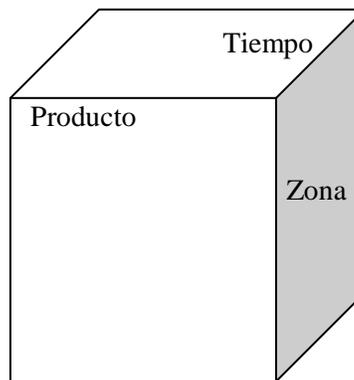


Figura No. 2.5: Diagrama Cubo de Información
Elaborado por: Juan Carlos Pérez

Los cubos de información o cubos OLAP funcionan como los cubos de rompecabezas en los juegos, en el juego se trata de armar los colores y en el datawarehouse se trata de organizar los datos por tablas o relaciones; los primeros (el juego) tienen 3 dimensiones, los cubos OLAP tienen un número indefinido de dimensiones, razón por la cual también reciben el nombre de hipercubos. Un cubo OLAP contendrá datos de una determinada variable que se desea analizar, proporcionando una vista lógica de los datos provistos por el sistema de información hacia el datawarehouse, esta vista estará dispuesta según unas dimensiones y podrá contener información calculada. El análisis de los datos está basado en las dimensiones del hipercubo, por lo tanto, se trata de un análisis multidimensional.

A la información de un cubo puede acceder el ejecutivo mediante "tablas dinámicas" en una hoja de cálculo o a través de programas personalizados. Las tablas dinámicas le permiten manipular las vistas (cruces, filtrados, organización, totales) de la información con mucha facilidad. Las diferentes operaciones que se pueden realizar con cubos de información se producen con mucha rapidez. Llevando estos conceptos a un datawarehouse, éste es una colección de datos que está formada por

“dimensiones” y “variables”, entendiendo como dimensiones a aquellos elementos que participan en el análisis y variables a los valores que se desean analizar.

Dimensiones

Las dimensiones de un cubo son atributos relativos a las variables, son las perspectivas de análisis de las variables (forman parte de la tabla de dimensiones). Son catálogos de información complementaria necesaria para la presentación de los datos a los usuarios, como por ejemplo: descripciones, nombres, zonas, rangos de tiempo, etc. Es decir, la información general complementaria a cada uno de los registros de la tabla de hechos.

Elementos que Integran un Datawarehouse

Metadatos

El término “metadatos” no tiene una definición única. Según la definición más difundida de metadatos es que son “datos sobre datos”. También hay muchas declaraciones como “informaciones sobre datos¹”, “datos sobre informaciones²” e “informaciones sobre informaciones³”.

¹Tom Sheldon. Linktionary. Entrada «Metadata». 2001, visitado 29 mayo de 2012

²A. Steinacker, A. Ghavam, R. Steinmetz. Metadata Standards for Web-Based Resources, Visitado 29 mayo 2012

³Ralph Swick. Metadata Activity Statement.2002, visto 29 mayo de 2012

Los metadatos han cobrado gran relevancia en el mundo de Internet, por la necesidad de utilizar los metadatos para la clasificación de la enorme cantidad de datos. Además de la clasificación los metadatos pueden ayudar en las búsquedas.

Podemos también considerar los metadatos, en las áreas de telecomunicaciones e informática, como información no relevante para el usuario final pero sí de suma importancia para el sistema que maneja la data. Los metadatos son enviados junto a la información cuando se realiza alguna petición o actualización de la misma.

En el campo biológico los metadatos se han convertido en una herramienta fundamental para el descubrimiento de datos e información. En este contexto se pueden definir los metadatos como «una descripción estandarizada de las características de un conjunto de datos» con esto se incluye la descripción del contexto en el cual los datos fueron coleccionados y además se refiere al uso de estándares para describirlos

Uno de los componentes más importantes de la arquitectura de un almacén de datos son los metadatos. Se define comúnmente como "datos acerca de los datos", en el sentido de que se trata de datos que describen cuál es la estructura de los datos que se van a almacenar y cómo se relacionan.

El metadato documenta, entre otras cosas, qué tablas existen en una base de datos, qué columnas posee cada una de las tablas y qué tipo de datos se pueden almacenar. Los datos son de interés para el usuario final, el metadato es de interés para los programas que tienen que manejar estos datos. Sin embargo, el rol que

cumple el metadato en un entorno de almacén de datos es muy diferente al rol que cumple en los ambientes operacionales. En el ámbito de los datawarehouse el metadato juega un papel fundamental, su función consiste en recoger todas las definiciones de la organización y el concepto de los datos en el almacén de datos, debe contener toda la información concerniente a:

- Tablas
- Columnas de tablas
- Relaciones entre tablas
- Jerarquías y Dimensiones de datos
- Entidades y Relaciones
- Funciones ETL (extracción, transformación y carga)

Los procesos de extracción, transformación y carga (ETL) son importantes ya que son la forma en que los datos se guardan en un almacén de datos (o en cualquier base de datos). Implican las siguientes operaciones:

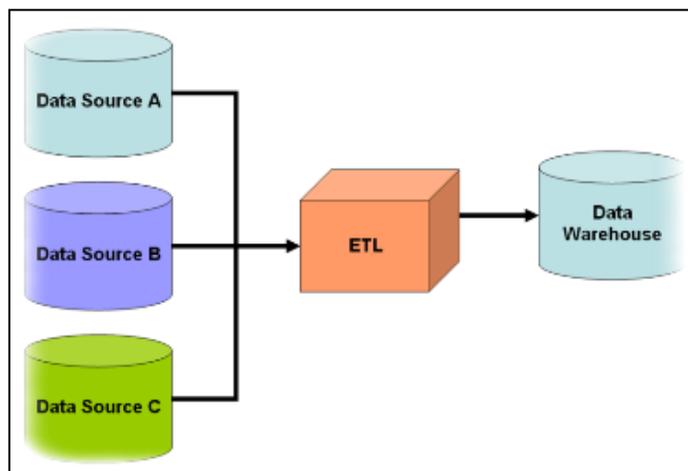


Figura No. 2.6: Diagrama Metadatos

Fuente: <http://www.irishtechjobs.net/2011/08/jobopportunity-senior-informatica-etl.html>

Extracción: La primera parte del proceso ETL consiste en extraer los datos desde los sistemas de origen. La mayoría de los proyectos de almacenamiento de datos fusionan datos provenientes de diferentes sistemas de origen. Cada sistema separado puede usar una organización diferente de los datos o formatos distintos. Los formatos de las fuentes normalmente se encuentran en bases de datos relacionales o ficheros planos, pero pueden incluir bases de datos no relacionales u otras estructuras diferentes. La extracción convierte los datos a un formato preparado para iniciar el proceso de transformación.

Una parte intrínseca del proceso de extracción es la de analizar los datos extraídos, de lo que resulta un chequeo que verifica si los datos cumplen la pauta o estructura que se esperaba. De no ser así los datos son rechazados.

Un requerimiento importante que se debe exigir a la tarea de extracción es que ésta cause un impacto mínimo en el sistema origen. Si los datos a extraer son muchos, el sistema de origen se podría ralentizar e incluso colapsar, provocando que éste no pueda utilizarse con normalidad para su uso cotidiano. Por esta razón, en sistemas grandes las operaciones de extracción suelen programarse en horarios o días donde este impacto sea nulo o mínimo.

Transformación: La fase de transformación aplica una serie de reglas de negocio o funciones sobre los datos extraídos para convertirlos en datos que serán cargados. Algunas fuentes de datos requerirán alguna pequeña manipulación de los datos. No obstante en otros casos pueden ser necesarias aplicar algunas de las siguientes transformaciones:

- Seleccionar sólo ciertas columnas para su carga (por ejemplo, que las columnas con valores nulos no se carguen).
- Traducir códigos (por ejemplo, si la fuente almacena una "H" para Hombre y "M" para Mujer pero el destino tiene que guardar "1" para Hombre y "2" para Mujer).
- Codificar valores libres (por ejemplo, convertir "Hombre" en "H" o "Sr" en "1").
- Obtener nuevos valores calculados (por ejemplo, total_venta = cantidad * precio).
- Unir datos de múltiples fuentes (por ejemplo, búsquedas, combinaciones, etc.).
- Calcular totales de múltiples filas de datos (por ejemplo, ventas totales de cada región).
- Generación de campos clave en el destino.
- Transponer o pivotar (girando múltiples columnas en filas o viceversa).
- Dividir una columna en varias (por ejemplo, columna "Nombre: García, Miguel"; pasar a dos columnas "Nombre: Miguel" y "Apellido: García").
- La aplicación de cualquier forma, simple o compleja, de validación de datos, y la consiguiente aplicación de la acción que en cada caso se requiera:
 - Datos OK: Entregar datos a la siguiente etapa (Carga).
 - Datos erróneos: Ejecutar políticas de tratamiento de excepciones (por ejemplo, rechazar el registro completo, dar al campo erróneo un valor nulo)

Carga: La fase de carga es el momento en el cual los datos de la fase anterior (transformación) son cargados en el sistema de destino. Dependiendo de los requerimientos de la organización, este proceso puede abarcar una amplia variedad de

acciones diferentes. En algunas bases de datos se sobrescribe la información antigua con nuevos datos. Los datawarehouse mantienen un historial de los registros de manera que se pueda hacer una auditoría de los mismos y disponer de un rastro de toda la historia de un valor a lo largo del tiempo.

Middleware

Middleware es un término genérico que se utiliza para referirse a todo tipo de software de conectividad que ofrece servicios u operaciones que hacen posible el funcionamiento de aplicaciones distribuidas sobre plataformas heterogéneas. Estos servicios funcionan como una capa de abstracción de software distribuida, que se sitúa entre las capas de aplicaciones y las capas inferiores (sistema operativo y red).

El *middleware* puede verse como una capa API, que sirve como base a los programadores para que puedan desarrollar aplicaciones que trabajen en diferentes entornos sin preocuparse de los protocolos de red y comunicaciones en que se ejecutarán. De esta manera se ofrece una mejor relación costo/rendimiento que pasa por el desarrollo de aplicaciones más complejas, en menos tiempo.(Inmon, 1992)

La función del middleware en el contexto de los datawarehouse es la de asegurar la conectividad entre todos los componentes de la arquitectura de un almacén de datos.

Ventajas y Desventajas de un Datawarehouse

Ventajas

Hay muchas ventajas por las que es recomendable usar un almacén de datos. Algunas de ellas son:

- Los almacenes de datos hacen más fácil el acceso a una gran variedad de datos a los usuarios finales.
- Facilitan el funcionamiento de las aplicaciones de los sistemas de apoyo a la decisión tales como informes de tendencia', *por ejemplo*: obtener los ítems con la mayoría de las ventas en un área en particular dentro de los últimos dos años;*informes de excepción*,informes que muestran los resultados reales frente a los objetivos planteados a priori.
- Los almacenes de datos pueden trabajar en conjunto y, por lo tanto, aumentar el valor operacional de las aplicaciones empresariales, en especial la gestión de relaciones con clientes.

Desventajas

Utilizar almacenes de datos también plantea algunos inconvenientes, algunos de ellos son:

- A lo largo de su vida los almacenes de datos pueden suponer altos costos. El almacén de datos no suele ser estático. Los costos de mantenimiento son elevados.
- Los almacenes de datos se pueden quedar obsoletos relativamente pronto.
- A veces, ante una petición de información estos devuelven una información subóptima, que también supone una pérdida para la organización.
- A menudo existe una delgada línea entre los almacenes de datos y los sistemas operacionales. Hay que determinar qué funcionalidades de estos se pueden aprovechar y cuáles se deben implementar en el datawarehouse, resultaría costoso implementar operaciones no necesarias o dejar de implementar alguna que sí vaya a necesitarse.

CATEGORÍAS VARIABLE DEPENDIENTE

Planificación de la Producción y Ventas

La idea no es planificar en el aire con estimaciones dadas de datos secundarios. Ahora está en plena producción la empresa y hay que ser eficientes. Se necesita contar con datos reales primarios.(García, 2010)

Presupuesto o pronóstico de venta: Ya se menciona, el punto de partida es el presupuesto o pronóstico de ventas. De hecho si se cumplen de alguna forma las premisas indicadas al tratar el proyecto o plan de negocio, allí siempre está el presupuesto o pronóstico de ventas. Debe existir, es lo básico para el planificador. Bien o mal, intuitivo, en secreto, formal, con investigación de mercado basado en estadísticas de demanda, calculado con algún método de aplicación disponible. El pronóstico de ventas está como punto de partida.

Estándar de producción: Este dato es muy importante, en él están basados los costos. Se deben controlar los pasos realizados detectando las diferencias entre lo real gastado y los estándares de producción para cada artículo en cada uno de los puntos en cuestión: Materia prima, insumos, materiales.

La ingeniería de producción puede diseñar otro proceso para fabricar lo mismo, se deberán re calcular los costos considerando todos los elementos que inciden directamente en el bien.

Capacidades de producción: Esto es limitativo. Más de ello no se puede hacer. Puede ser dado en horas, turnos, días, semanas, meses. Se puede hacer cálculos expresando combinaciones: tiempo de maquina es HM (horas maquinas), tiempo de trabajo de la mano de obra es HH (horas hombre). Son todos datos muy importantes en los cálculos industriales.

El plan de producción

La idea que la mercadería esté disponible en el momento de venta o pedido a un costo menor de producción. Es cumplir con las previsiones de ventas. Ahora si de entrada se ve que las previsiones son mayores que las posibilidades de producción en tiempo y forma, se deberá informar al sector ventas para que disminuya sus pretensiones y considerar la posibilidad de proponer un plan de inversión para el próximo ejercicio.

Los datos y puntos a controlar: Los datos a controlar son: Ventas (o usado), Producción, inventario de existencias (PT e Interproceso).

El control del interproceso es también llamado control de producción en proceso, tiene puntos de control. Las planillas de control se llaman de control acumulado y en ello el dato a leer inmediatamente es el inventario de existencias actual. Según lo visto en reaprovisionamiento de materiales, es el que se compara con el punto de pedido.

$$\text{Inventario de existencias actual} = \text{Inventario de existencias anterior} + \text{producción} - \text{ventas}$$

Por supuesto el cálculo es para cada artículo, en función del inventario de existencias actual se irán asignando los pedidos. Si se trata del producto terminado. El pedido se hará a la producción. Si se trata de materia prima, insumos, materiales el pedido se hará a compras.

Programar la producción: Para armar la planilla deberemos conocer el periodo a programar de cada producto. Puede ser diaria, semanal, decenal, quincenal.

En analista programador de la producción debe conocer los procesos y contar con la información actualizada. El debe hacer producir con la formula o, de la forma diseñada con todos las características necesarias como ser cantidades, unidades de medidas usadas.

Procesamiento Analítico en Línea

OLAP

Es el acrónimo en inglés de procesamiento analítico en línea (On-Line Analytical Processing). Es una solución utilizada en el campo de la Inteligencia de Negocios (Business Intelligence), la cual consiste en consultas a estructuras multidimensionales (o Cubos OLAP) que contienen datos resumidos de grandes Bases de Datos o Sistemas Transaccionales (OLTP). Se usa en informes de negocios de ventas, márketing, informes de dirección, minería de datos y áreas similares.(Mailvaganam, 2007)

La razón de usar OLAP para las consultas es la velocidad de respuesta. Una base de datos relacional almacena entidades en tablas discretas si han sido normalizadas. Esta estructura es buena en un sistema OLTP pero para las complejas consultas multitabla es relativamente lenta. Un modelo mejor para búsquedas, aunque peor desde el punto de vista operativo, es una base de datos multidimensional. La principal característica que potencia a OLAP, es que es lo más rápido a la hora de hacer selects, en contraposición con OLTP que es la mejor opción para INSERTS, UPDATES Y DELETES. Existen algunas clasificaciones entre las implementaciones OLAP. La clasificación está hecha sobre la base de en qué tipo de motor son almacenados los datos:

OLTP

OLTP es la sigla en inglés de Procesamiento de Transacciones En Línea (OnLine Transaction Processing) es un tipo de sistemas que facilitan y administran aplicaciones transaccionales, usualmente para entrada de datos y recuperación y procesamiento de transacciones (gestor transaccional). Los paquetes de software para OLTP se basan en la arquitectura cliente-servidor ya que suelen ser utilizados por empresas con una red informática distribuida (Mailvaganam, 2007).

El término puede parecer ambiguo, ya que puede entenderse "*transacción*" en el contexto de las "*transacciones computacionales*" o de las "*transacciones en bases de datos*". También podría entenderse en términos de transacciones de negocios o comerciales. OLTP también se ha utilizado para referirse a la transformación en la que el sistema responde de inmediato a las peticiones del usuario. Un cajero automático de un banco es un ejemplo de una aplicación de procesamiento de transacciones comerciales.

OLAP

OLAP es el acrónimo en inglés de procesamiento analítico en línea (*On-Line Analytical Processing*). Es una solución utilizada en el campo de la llamada Inteligencia empresarial (o *Business Intelligence*) cuyo objetivo es agilizar la consulta de grandes cantidades de datos. Para ello utiliza estructuras multidimensionales (o Cubos OLAP) que contienen datos resumidos de grandes Bases de datos o Sistemas Transaccionales (OLTP). Se usa en informes de negocios de ventas, marketing, informes de dirección, minería de datos y áreas similares (Mailvaganam, 2007).

La razón de usar OLAP para las consultas es la rapidez de respuesta. Una base de datos relacional almacena entidades en tablas discretas si han sido normalizadas. Esta estructura es buena en un sistema OLTP pero para las complejas consultas multitabla es relativamente lenta. Un modelo mejor para búsquedas (aunque peor desde el punto de vista operativo) es una base de datos multidimensional.

La principal característica que potencia a OLAP, es que es lo más rápido a la hora de ejecutar sentencias SQL de tipo SELECT, en contraposición con OLTP que es la mejor opción para operaciones de tipo INSERT, UPDATE Y DELETE

ROLAP

Implementación OLAP que almacena los datos en un motor relacional. Típicamente, los datos son detallados, evitando las agregaciones y las tablas se encuentran desnormalizadas. Los esquemas más comunes sobre los que se trabaja son

estrella ó copo de nieve, aunque es posible trabajar sobre cualquier base de datos relacional. La arquitectura está compuesta por un servidor de banco de datos relacional y el motor OLAP se encuentra en un servidor dedicado. La principal ventaja de esta arquitectura es que permite el análisis de una enorme cantidad de datos(Mailvaganam, 2007).

Esquema Estrella:En las bases de datos usadas para *data warehousing*, un esquema en estrella es un modelo de datos que tiene una tabla de hechos (o *tabla fact*) que contiene los datos para el análisis, rodeada de las tablas de dimensiones. Este aspecto, de tabla de hechos (o central) más grande rodeada de radios o tablas más pequeñas es lo que asemeja a una estrella, dándole nombre a este tipo de construcciones.

Las tablas de dimensiones tendrán siempre una clave primaria simple, mientras que en la tabla de hechos, la clave principal estará compuesta por las claves principales de las tablas dimensionales.

Esquema Copo de Nieve:En las bases de datos utilizadas en (*data warehousing*), un esquema en copo de nieve es una estructura algo más compleja que el esquema en estrella. Se da cuando alguna de las dimensiones se implementa con más de una tabla de datos. La finalidad es normalizar las tablas y así reducir el espacio de almacenamiento al eliminar la redundancia de datos; pero tiene la contrapartida de generar peores rendimientos al tener que crear más tablas de dimensiones y más relaciones entre las tablas (*JOINS*) lo que tiene un impacto directo sobre el rendimiento.

MOLAP

Esta implementación OLAP almacena los datos en una base de datos multidimensional. Para optimizar los tiempos de respuesta, el resumen de la información es usualmente calculado por adelantado. Estos valores pre calculados o agregaciones son la base de las ganancias de desempeño de este sistema. Algunos sistemas utilizan técnicas de compresión de datos para disminuir el espacio de almacenamiento en disco debido a los valores pre calculados(Mailvaganam, 2007).

HOLAP (Hybrid OLAP)

Almacena algunos datos en un motor relacional y otros en una base de datos multidimensional.

VISIÓN MULTIDIMENSIONAL DE LA INFORMACIÓN

La información es un factor crítico para el éxito empresarial, ésta es cada día más abundante y diversa, procedente de múltiples fuentes, que llega en diferentes formatos, que hay que recoger, ordenar, explotar, y manipular para obtener un valor agregado, forma parte de la estrategia competitiva de las organizaciones. (Lara Treviño, 2012)

Si la información de una empresa no es administrada adecuadamente y no está disponible para su uso en el momento adecuado, puede perder todo valor ante el

proceso de toma de decisiones. Esto hace evidente la necesidad de procesos y herramientas que faciliten el manejo e interpretación de los datos que arrojan los sistemas de operación y producción. Para entender un DSS es necesario entender también lo que es la toma de decisiones.

La Toma de Decisiones.

La toma de decisiones es una parte esencial en la existencia y supervivencia de las organizaciones, y por desgracia en muchas ocasiones no es una actividad sencilla y rápida, además de que para tomar una buena decisión están involucrados otros factores como por ejemplo, Información y el juicio de la persona.

Conforme van pasando los años esto se vuelve más y más complejo, debido a la dinámica de las empresas y el ambiente en el que están inmersos, y conforme pase el tiempo esto se irá incrementando. De ahí surge la necesidad de estar siempre actualizados en cuanto a información relevante para la compañía, ya no solo interna, sino externa, y por otro lado contar con las herramientas necesarias que apoyen el proceso de la toma de decisión. Algunas razones por las que ahora se vuelve más complejo y difícil, son las siguientes:

- El número de alternativas disponibles es ahora mayor debido al surgimiento de la tecnologías y los sistemas de comunicación
- El costo de los errores administrativos puede ser muy grande por la complejidad y la magnitud de las operaciones y las reacciones que estos errores pueden causar.

- La información necesaria para tomar decisiones puede ser difícil de acceder y se ha incrementado exponencialmente.
- Las decisiones deben ser tomadas rápidamente.

Para comprender mejor la toma de decisión es importante conocer las etapas que están involucradas en este proceso. Rivera (2001) y Turban (2001) las dividen y describen de la siguiente manera:

- **Inteligencia:** Es la etapa en la que el individuo recopila información para identificar los problemas que ocurren en la organización.
- **Diseño:** Es cuando el individuo concibe las posibles alternativas al problema en la organización.
- **Selección:** El individuo selecciona una alternativa entre las posibles soluciones.

Ventajas del uso de un DSS

Las ventajas que se listan a continuación, son comentadas por diferentes autores como son (Bobadilla, 2001), (Elizondo, 2002) y (Turban, 2001).

- Una alta calidad en la toma de decisiones.
- Mayor comunicación.
- Reducción de costos.
- Mayor productividad.
- Ahorro de tiempo en la toma de decisiones.
- Mayor satisfacción del cliente y los empleados.

- Analizan y procesan grandes volúmenes de información, necesaria para la correcta toma de decisiones.
- Rápido acceso a la información.
- Algunos de estos sistemas ya vienen preparados para poner en marcha rápidamente sistemas integrados de medición de objetivos u otros indicadores clave, por áreas estratégicas o áreas de responsabilidad y su control a través de avisos que concentren la atención en los puntos débiles de la organización y permitan actuar directamente sobre ellos.
- Presentan a través de una interfaz amigable y de fácil manejo, una visión clave de la organización según los factores críticos definidos en su creación.
- Permiten el análisis de datos en línea y en tiempo real, a través de múltiples perspectivas: geográfica, productos, distribución, etc.; buscar patrones estadísticos significativos: relaciones entre datos, segmentación, etc., construir sistemas de información sinóptica y orientada a la alta dirección o la distribución electrónica de información a través de servidores de documentos.

Ejemplos de áreas de aplicación

Algunos ejemplos de aplicaciones más específicas según Sterling (2003) son:

- Investigación y planeación de mercado
- Planeación operacional y estratégica
- Soporte a ventas

Hipótesis

“El datawarehouse incide significativamente en la Planificación de Producción y Ventas de Industria Papelera Caicedo Miño Cía. Ltda. de la Ciudad de Ambato.”

Variable Independiente: Datawarehouse

Variable Dependiente: Planificación de la Producción y Ventas

CAPÍTULO III

METODOLOGÍA

Investigación de Campo

La investigación se considera de campo debido a la aplicación de encuestas a las personas que comprenden de cada uno de los departamentos sobre su problemática actual, permitiendo determinar los requerimientos para la propuesta y de esta forma determinar el alcance de la misma en el desarrollo de la propuesta.

Investigación documental – bibliográfica

La investigación es de naturaleza documental - bibliográfica debido a propósitos de fortalecimiento de la investigación se recurre a diferentes fuentes obtenidas de libros especializados, artículos, páginas web publicadas en internet, tesis desarrolladas en otras Universidades, y otros, como se cita en la bibliografía.

Tipos y 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

Para esta investigación se ha considerado que se va a trabajar con 1 jefe de planta, 2 supervisores, 1 coordinador del área de sistemas, 1 funcionario de ventas y empleados de distintos departamentos.

POBLACIONES	FRECUENCIAS	PORCENTAJES
Jefe de planta	1	10%
Supervisores	1	10%
Coordinador área de sistemas	1	10%
Departamento de Ventas	2	20%
Departamento de Producción	5	50%
TOTAL	10	100%

Tabla N.-3.1. Muestreo no probabilístico
Elaborado por: Ing. Juan Carlos Pérez

En virtud de que ninguna de las poblaciones a ser investigadas pasa de 100 elementos se trabajará con la totalidad del universo sin que sea necesario sacarmuestras representativas.

Operacionalización de Variables

Operacionalización de la Variable Independiente

Variable Independiente: Datawarehouse

CONCEPTUALIZACIÓN	DIMENSIONES	INDICADORES	ÍTEMS BÁSICOS	TÉCNICAS E INSTRUMENTOS
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. 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 Business Intelligence.	Integrar y depurar información Desorganización en la información, lentitud en los reportes	Organizar, acceder y mostrar: La información de la base de datos del sistema de producción. En una sola consulta: Producto, estado de la producción, suministro, materia prima, tiempos.	¿Actualmente como accede y administra la información sobre el proceso de producción /ventas? ¿Cómo se presentan la información de la base de datos mediante los reportes de producción/ ventas?	Entrevista Guía de la entrevista Encuesta Cuestionario

Tabla N.- 3.2.Operacionalización de la Variable Independiente
Elaborado por: Ing. Juan Carlos Pérez

Técnicas e Instrumentos

Encuesta

Dirigido a los funcionarios de los departamentos de producción así como al departamento de ventas para recabar información del manejo de las bases de datos en la fábrica.

Entrevista

Enfocado al departamento de producción, departamento de ventas y jefatura de planta para 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

La información recolectada en base a las encuestas y entrevistas realizadas en campo, estas se transformó de acuerdo a los procedimientos siguientes:

- Revisión de la información.
- Tabulación de la información.
- Estudio de los resultados.

Aplicadas y analizadas las técnicas a continuación se efectuó la tabulación de esta información, lo que permitió realizar una estructuración de conclusiones y recomendaciones, las que organizadas en una propuesta permitan la solución lógica y factible a la problemática planteada.

Para concretar el plan de recolección de la información, conviene contestar las siguientes preguntas:

N.-	PREGUNTAS	EXPLICACIÓN
1	¿Para qué?	Para alcanzar los objetivos de la investigación
2	¿De qué personas u objetos?	Sujetos: Jefatura, funcionario departamento de ventas y departamento de producción.
3	¿Sobre qué aspectos?	Indicadores (matriz de Operacionalización de variables)
4	¿Quién, quienes?	Investigador
5	¿Cuándo?	Poner fecha tentativa eje: enero 2012
6	¿Dónde?	Oficinas: Jefatura de planta, Producción y Ventas
7	¿Cuántas veces?	10
8	¿Qué técnicas de recolección?	Encuesta Entrevista
9	¿Con qué?	Cuestionario Guía de Entrevista
10	¿En qué situación?	Durante las jornadas de trabajo, y previa cita.

Tabla N.- 3.4.Recolección de la Información
Elaborado por: Ing. Juan Carlos Pérez

Plande Procesamiento de Información

- Revisión crítica de la información recogida; es decir, limpieza de la información defectuosa: contradictoria, incompleta, no pertinente, etc.
- Repetición de la recolección, en ciertos casos individuales, para corregir fallas de contestación.
- Tabulación o cuadros según variables de cada hipótesis
- Manejo de información (reajuste de cuadros con casillas vacías o condatos tan reducidos cuantitativamente, que no influyen significativamente en los análisis).
- Estudio estadístico de datos para presentación de resultados.
- La presentación de datos puede hacerse siguiendo los siguientes procedimientos:
 - Representación escrita
 - Representación tabular
 - Representación gráfica

FORMATO GUÍA PARA LA ENCUESTA

Guíade encuesta a funcionarios de losdistintos departamentos de la empresa
Industria Papelera Caicedo Miño.

Departamento: _____ Fecha _____			
Sírvase marcar en el paréntesis con una X de acuerdo a su criterio			
N.	PREGUNTAS	RESPUESTAS	COD.
1	¿Dispone usted información detallada de sus ordenes de producción / ventas?	<ul style="list-style-type: none"> • Mucho • Poco • Nada 	1. () 2. () 3. ()
2	¿Con que frecuencia se retrasa la entrega de sus ordenes de producción / ventas?	<ul style="list-style-type: none"> • Siempre • Frecuentemente • Rara vez 	1. () 2. () 3. ()
3	¿Considera usted que las herramientas utilizadas para la planificación de la producción y ventas es el más recomendado?	<ul style="list-style-type: none"> • Si • No 	1. () 2. ()
4	¿Considera la forma en la que se realizan los informes de producción y ventas es el más acertada?	<ul style="list-style-type: none"> • Si • No • Desconozco 	1. () 2. () 3. ()
5	¿Las herramientas informáticas existentes colaboran con la planificación de la producción y ventas en la empresa?	<ul style="list-style-type: none"> • Siempre • Frecuentemente • Rara vez 	1. () 2. () 3. ()
6	¿Los informes actuales de producción y ventas ayudan a tomar las decisiones en las planificaciones futuras a la empresa?	<ul style="list-style-type: none"> • Siempre • Frecuentemente • Rara vez 	1. () 2. () 3. ()
7	¿El tiempo y recursos que toma realizar un informe detallado de la producción / ventas es?	<ul style="list-style-type: none"> • Alto • Medio • Baio 	1. () 2. () 3. ()
8	¿Cree Usted que una herramienta que procese los datos ágilmente ayudara a la producción y ventas de la empresa?	<ul style="list-style-type: none"> • Si • No 	1. () 2. ()

Tabla N.- 3.5.Formato Encuesta
Elaborado por: Ing. Juan Carlos Pérez

FORMATO GUÍA PARA LA ENTREVISTA

Guíade entrevista aplicada afuncionarios de los distintos departamentos de la empresa Industria Papelera Caicedo Miño.

Sírvase escribir en el recuadro de la derecha correspondiente de acuerdo al criterio del encuestado

Departamento: _____ Fecha: _____	
Entrevistado: _____ Cargo: _____	
Pregunta	Respuesta
1	¿Actualmente como accede y administra la información sobre el proceso de producción /ventas?
2	¿Cuenta con un informe de inventario / ventas confiable?
3	¿Cómo obtiene información del estado de la producción?
4	¿Qué método utiliza para la provisión anual de la materia prima - suministros?

Tabla N.- 3.6.Formato Entrevista
Elaborado por: Ing. Juan Carlos Pérez

CAPÍTULO IV

Análisis e interpretación de resultados

Luego de aplicar las técnicas e Instrumentos de recolección de la información en los departamentos involucrados en la presente investigación se realiza la respectiva tabulación e interpretación de los datos obtenidos de la aplicación de las encuestas y entrevistas realizadas en todos los departamentos involucrados en la problemática, misma que arroja los resultados que a continuación se muestran:

PREGUNTA 1: ¿Dispone Usted información detallada de sus órdenes de producción/ventas?

Respuesta	Encuestados	Porcentaje
Mucho	1	11%
Poco	7	78%
Nada	1	11%
Total:	9	100%

Tabla N.-4.1. Análisis de Datos Pregunta 1

Elaborado por: Ing. Juan Carlos Pérez

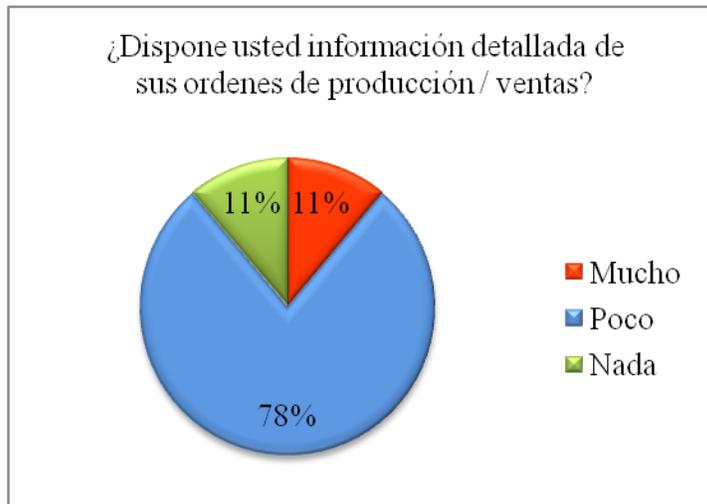


Figura N.-4.1. Análisis gráfico de datos Pregunta 1

Elaborado por: Juan Carlos Pérez

En esta pregunta podemos observar que las 9 personas encuestadas que corresponde el 100%, el 11% cuenta con mucha información, el 78% responde a que tiene una poca disponibilidad de sus órdenes de producción / ventas de forma detallada, mientras que el 11% restante definitivamente no cuenta con esta información.

PREGUNTA 2: ¿Con que frecuencia se retrasa la entrega de sus órdenes de producción/ventas?

Respuesta	Encuestados	Porcentaje
Siempre	0	0%
Frecuentemente	6	75%
Rara vez	2	25%
Total:	8	100%

Tabla N.- 4.2. Análisis de Datos Pregunta 2

Elaborado por: Ing. Juan Carlos Pérez

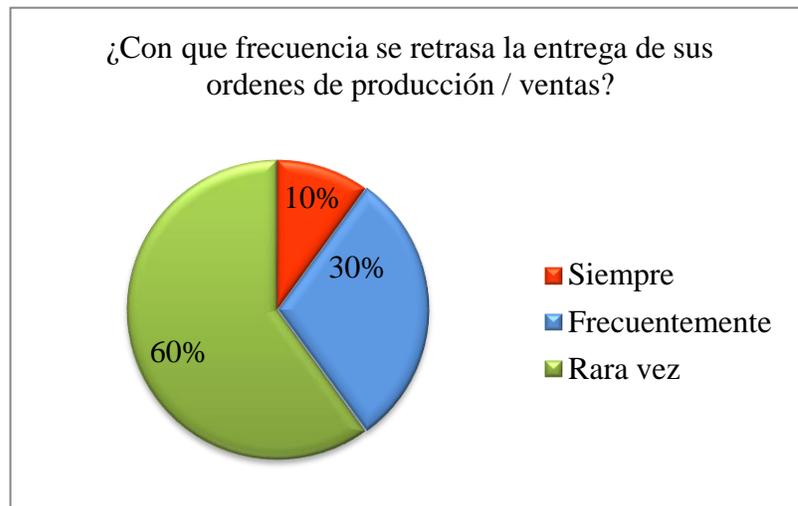


Figura N.-4.2. Análisis gráfico de datos Pregunta 2

Elaborado por: Juan Carlos Pérez

De un total de 8 personas encuestadas es decir el 100% sobre la frecuencia con la que se retrasan en la entrega de las ordenes de producción / ventas, se obtuvo que el 75% frecuentemente se retrasa en la entrega tanto de las ordenes de producción como en las ventas, finalmente un 25% rara vez tiene este problema.

PREGUNTA 3: ¿Considera Usted que las herramientas utilizadas para la planificación de la producción y ventas es el más recomendado?

Respuesta	Encuestados	Porcentaje
Si	1	10%
No	9	90%
Total:	10	100%

Tabla N.- 4.3: Análisis de Datos Pregunta 3

Elaborado por: Juan Carlos Pérez

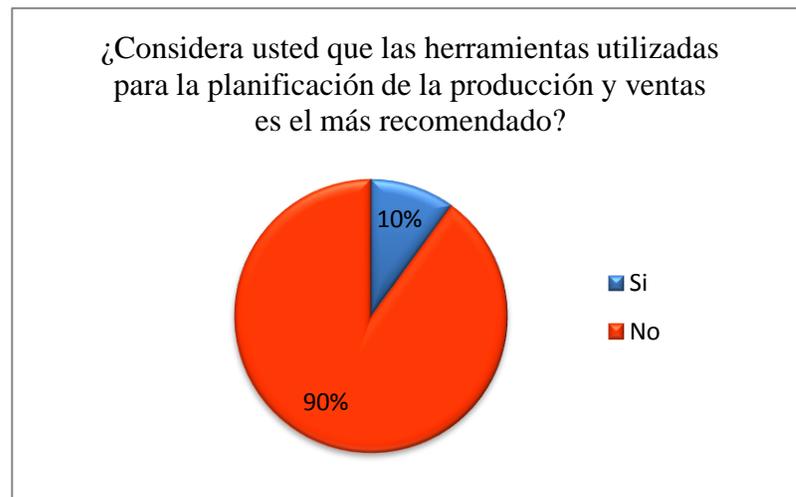


Figura N.-4.3. Análisis gráfico de datos Pregunta 3

Elaborado por: Juan Carlos Pérez

Esta pregunta en la que participaron todos los departamentos involucrados es decir 10 personas en cuanto a que si se considera que las herramientas utilizadas para la planificación de la producción y ventas es el más recomendado un 90% piensa que no, y un 10% ha expresado que sí.

PREGUNTA 4: ¿Considera la forma en la que se realizan los informes de producción y ventas es la más acertada?

Respuesta	Encuestados	Porcentaje
Si	2	20%
No	6	60%
Desconozco	2	20%
Total:	10	100%

Tabla N.- 4.4. Análisis de Datos Pregunta 4

Elaborado por: Juan Carlos Pérez



Figura N.-4.4. Análisis gráfico de datos Pregunta 4

Elaborado por: Juan Carlos Pérez

En cuanto a las consideraciones sobre si la forma de realizar actualmente los informes de producción y ventas es la más acertada, 10 personas que corresponde el 100%, dieron su punto de vista, como resultado 60% piensa que no, debido a los resultados y la forma en que lo obtienen, 20% piensa que si pero tomando en cuenta que son informes que no influyen en las planificaciones y un 20% desconoce esta posibilidad.

PREGUNTA 5: ¿Las herramientas informáticas existentes colaboran con la planificación de la producción y ventas en la empresa?

Respuesta	Encuestados	Porcentaje
Siempre	1	10%
Frecuentemente	3	30%
Rara vez	6	60%
Total:	10	100%

Tabla N.-4.5. Análisis de Datos Pregunta 5

Elaborado por: Juan Carlos Pérez

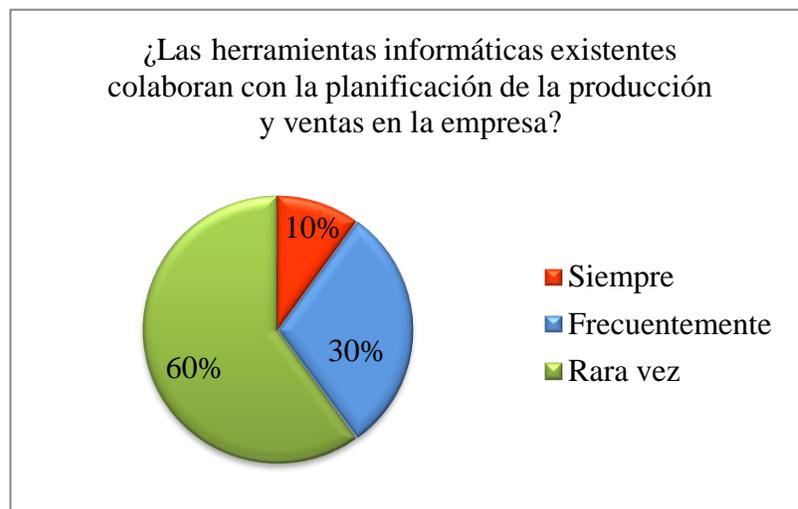


Figura N.-4.5. Análisis gráfico de datos Pregunta 5

Elaborado por: Juan Carlos Pérez

Se encuestó si las herramientas informáticas que se usan actualmente para realizar la planificación de la producción y ventas de la empresa, de las 10 personas que es el 100%, usan directa o indirectamente estas herramientas dijeron un 60% rara vez les colaboran las herramientas en la planificación, 30% frecuentemente les ayuda y un 10% siempre utiliza las herramientas para sus planificaciones.

PREGUNTA 6: ¿Los informes actuales de producción y ventas ayudan a tomar las decisiones en las planificaciones futuras a la empresa?

Respuesta	Encuestados	Porcentaje
Siempre	1	11%
Frecuentemente	2	22%
Rara vez	6	67%
Total:	9	100%

Tabla N.-4.6. Análisis de Datos Pregunta 6

Elaborado por: Juan Carlos Pérez

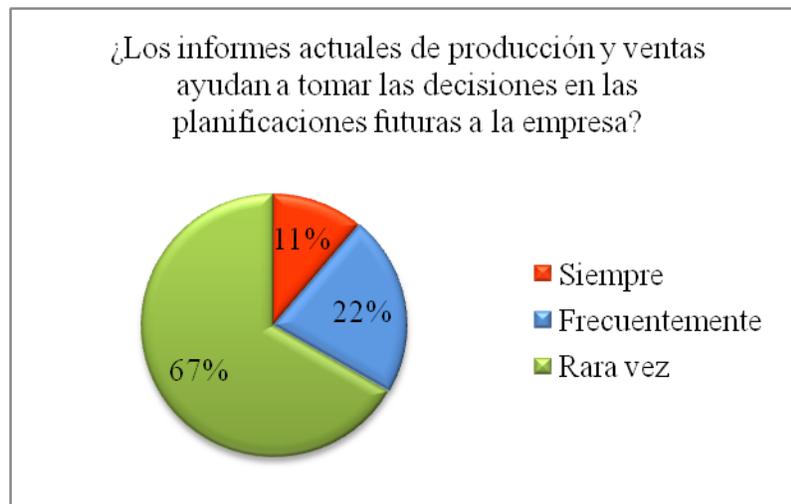


Figura N.-4.6. Análisis gráfico de datos Pregunta 6

Elaborado por: Juan Carlos Pérez

Las 9 personas que son las encargadas de las planificaciones futuras de la producción y ventas de la empresa y corresponden al 100%, un 67% se pronunció que rara vez se basa en estos informes, un 22% piensa que frecuentemente ayudan los informes en la toma de decisiones futuras y un 11% siempre se ayuda de dicha información.

PREGUNTA 7: ¿El tiempo y recursos que toma realizar un informe detallado de la producción/ventas es?

Respuesta	Encuestados	Porcentaje
Bajo	0	0%
Medio	0	0%
Alto	10	100%
Total:	10	100%

Tabla N.-4.7. Análisis de Datos Pregunta7

Elaborado por: Juan Carlos Pérez

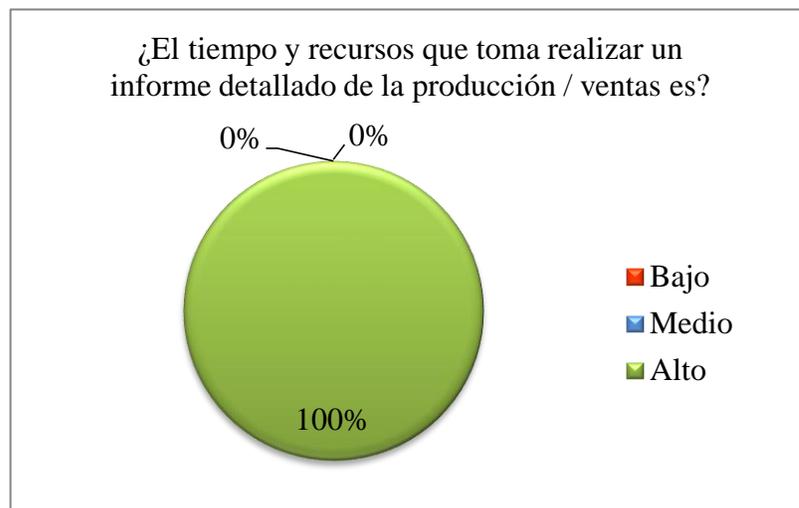


Figura N.-4.7. Análisis gráfico de datos Pregunta 7

Elaborado por: Juan Carlos Pérez

El 100% de los involucrados en la elaboración de informes de producción y ventas fueron consultados sobre la cantidad de tiempo que invierten en realizar esta tarea, opinando de forma unánime que hacen una alta inversión de tiempo y recursos, considerando este tema lo más complicado en sus actividades diarias dentro de cada uno de sus departamentos.

PREGUNTA 8: ¿Cree Usted que una herramienta que procese los datos ágilmente ayudará a la producción y ventas de la empresa?

Respuesta	Encuestados	Porcentaje
Si	9	90%
No	1	10%
Total:	10	100%

Tabla N.-4.8. Análisis de Datos Pregunta8

Elaborado por: Juan Carlos Pérez

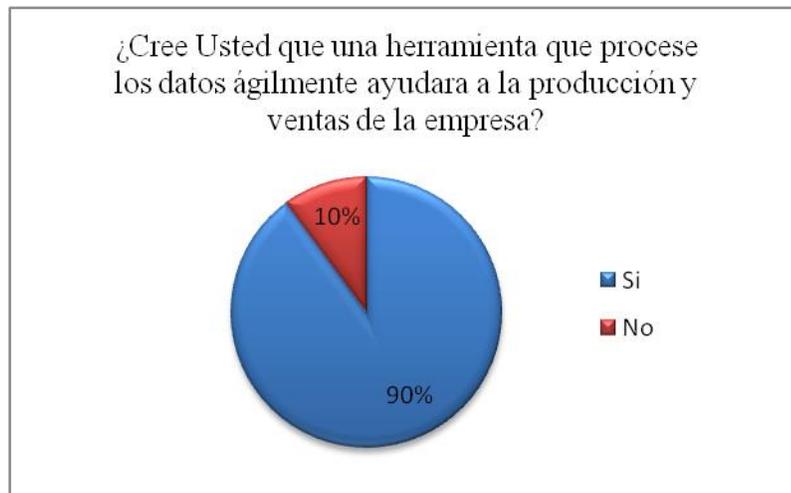


Figura N.-4.8. Análisis gráfico de datos Pregunta 8

Elaborado por: Juan Carlos Pérez

10 Personas el 100% consultadas en cuanto a que si se considera que la implementación de una nueva herramienta que procese de una mejor manera la información de la base de datos de forma ágil para que de esta forma ayude a la toma de decisiones a los departamentos de producción y ventas, un 90% piensa que si ayudaría, y un 10% ha expresado que no cree que una herramienta solucione sus problemas.

VERIFICACIÓN DE LA HIPÓTESIS

Primeramente se procede al plantear de la hipótesis para ser verificada.

Formulación de la Hipótesis:

- H_0 = Hipótesis nula
- H_1 = Hipótesis alterna

H_0 = la aplicación de un datawarehouse **SI** incide significativamente en la planificación de producción y ventas de Industrial Papelera Caicedo Miño.

H_1 = la aplicación de un datawarehouse **NO** incide significativamente en la planificación de producción y ventas de Industrial Papelera Caicedo Miño.

En esta investigación se utiliza los métodos siguientes para la verificación de la hipótesis.

METODO LÓGICO DEDUCTIVO

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

Estado actual de la información en Industrial Papelera Caicedo Miño S. A. de la Ciudad de Ambato.

NOMBRE DEL PROCESO	ESTADO ACTUAL DEL PROCESO	FUNCIONAMIENTO ESPERADO DEL PROCESO
Cobranzas	Para poder obtener esta información es un trabajo arduo para el departamento de cobranzas, debiéndose consumir muchos recursos y tiempo.	El departamento de ventas cuenta con un flujo de información rápida, precisa y actualizada de los valores recaudados a los clientes de la empresa.
Recaudaciones por regiones	Las recaudaciones económicas que se realizan por regiones del país son muy importantes para hacer hincapié en mejorar las cobranzas en dichas regiones, pero no se tiene esta información de una forma ágil.	Lo óptimo es que se obtenga información del progreso de las cobranzas de cada región del país, mejorando de esta forma la liquidez de la empresa.
Pedidos por vendedor	Cada vendedor realiza un reporte de los pedidos los mismos que son ingresados al sistema, otros se encuentran en papel pero al momento de obtener información de estos pedidos se acumula y no es fácil para el departamento obtener esta información.	Se espera que se cuente con información ágil y precisa de los pedidos realizados por cada vendedor, para de esta manera no causar el retraso en la entrega de los pedidos de los clientes, por los diferentes vendedores.

Pedidos por zona	En nuestro país el año lectivo es diferente entre zonas como son costa, sierra y oriente, donde por la naturaleza del negocio que es la producción de útiles escolares, se depende de esta información para poder planificar la producción para las diferentes temporadas lectivas. Pero como no se cuenta con esta información se realizan las programaciones de producción en base a costumbre.	La información de los pedidos esté siempre actualizada y se pueda hacer consultas inmediatamente para de esta manera planificar la producción de los distintos productos que la industria produce, de una manera eficiente.
Pedidos por producto	Cada producto es vendido en mayor o menor cantidad, dependiendo de la temporada en la que se encuentre, lo que causa que se deba producir en cantidades siempre diferentes, como no se cuenta con los pedidos inmediatamente, se pierde la oportunidad de producir con prioridades a los productos más pedidos.	Se espera que la cantidad que se requiere de cada producto se cuente de forma actualizada, para de esta forma mandar a producir con la antelación del caso y no retrasar las entregas.

Producción	En base a la información obtenida un alto porcentaje del personal a cargo del departamento de producción, no cuenta con históricos detallados de la producción en la empresa.	Lo más óptimo sería que se tenga históricos de producción entre años anteriores, los pedidos actuales, para de esta manera planificar las ordenes de producción de los diferentes pedidos.
Ventas por vendedor	No se puede mejorar el desempeño de cada vendedor por qué no se tiene una información detallada en base a estadísticas de ventas por cada uno de los vendedores.	Contar con información de las ventas por vendedora para de esta manera solicitar un mejor desempeño a cada vendedor.
Consultas en línea	No se cuentan con información en línea de la información de la producción y ventas.	Disponer de información en línea de la información de producción y ventas.

Tabla 4.9: Estado Actual y Esperado de los Procesos
Elaborado por: Ing. Juan Carlos Pérez

Una vez conocido el estado actual de los procesos que se realizan en la empresa se puede *lógicamente* deducir que la hipótesis H_1 que indica que el desarrollo de un datawarehouse incide positivamente en la Planificación de la Producción y Ventas de la Industrial Papelera Caicedo Miño 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

Las pruebas de Ji cuadrada es una prueba estadística para evaluar hipótesis acerca de la relación entre dos variables categóricas.

PRUEBA DE HIPÓTESIS

Elección de la prueba estadística

Para la verificación de la hipótesis se eligió la prueba Ji Cuadrada, cuya fórmula es la siguiente:

$$X^2 = \sum \left(\frac{(fo - fe)^2}{fe} \right)$$

Simbología:

- X^2 = Ji cuadrada
- fo = Frecuencia observada
- fe = Frecuencia esperada

Para la realización de la matriz de tabulación cruzada se toma en cuenta 2 preguntas de la encuesta realizada a los departamentos de producción y ventas de la siguiente manera:

Pregunta 3: ¿Considera usted que las herramientas utilizadas para la planificación de la producción y ventas es el más recomendado?

Respuesta	Encuestados	Porcentaje
Si	1	10%
No	9	90%
Total:	10	100%

Tabla N.-4.10. Análisis encuesta Pregunta 3
Elaborado por: Juan Carlos Pérez

Pregunta 8: ¿Cree Usted que una herramienta que procese los datos ágilmente ayudara a la producción y ventas de la empresa?

Respuesta	Encuestados	Porcentaje
Si	9	90%
No	1	10%
Total:	10	100%

Tabla N.-4.11. Análisis encuesta Pregunta 8
Elaborado por: Juan Carlos Pérez

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 Observada (fo)

Parámetros	Alternativas		Total
	Si	No	
las herramientas utilizadas para la planificación de la producción y ventas	1	9	10
herramienta que procese los datos ágilmente ayudara a la producción y ventas	9	1	10
Total	10	10	20

Tabla N.-4.12. Frecuencias Observadas
Elaborado por: Juan Carlos Pérez

Frecuencias Esperadas

Para calcular la frecuencia esperada se utiliza la siguiente fórmula:

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

Parámetros	Alternativas	
	Si	No
herramientas utilizadas para la planificación de la producción y ventas	5	5
herramienta que procese los datos ágilmente ayudara a la producción y ventas	5	5

Tabla N.-4.13.Frecuencias Esperadas
Elaborado por: Juan Carlos Pérez

Cálculo de Ji Cuadrada

$x^2 = \sum \left(\frac{(f_o - f_e)^2}{f_e} \right)$	f_o	f_e	$f_o - f_e$	$(f_o - f_e)^2$	$\frac{(f_o - f_e)^2}{f_e}$
Herramientas utilizadas /SI	1	5	- 4	16	3.2
Herramientas utilizadas /NO	9	5	4	16	3.2
herramienta que procese los datos ágilmente /SI	9	5	4	16	3.2
herramienta que procese los datos ágilmente /NO	1	5	- 4	16	3.2
					$x^2 = 12.8$

Tabla N.-4.14.Calculo de Ji Cuadrada
Elaborado por: Juan Carlos Pérez

Decisión

El valor de $x_t^2 = 2.2141 < x_c^2 = 12.8$ (Ver Anexo 3)

En base a la verificación de la hipótesis alterna se determina que un datawarehouse incide significativamente en la Planificación de Producción y Ventas de Industria Papelera Caicedo Miño Cía. Ltda. de la Ciudad de Ambato y por consiguiente se rechaza la hipótesis nula.

CAPÍTULO V

CONCLUSIONES Y RECOMENDACIONES

Conclusiones

- Uno de los problemas frecuentes detectados es la poca de información detallada del estado en que se encuentran sus procesos de producción, así como las ventas, acarreando retrasos en la producción y ventas, principalmente debido a que las herramientas utilizadas para realizar esta labor no son las recomendadas.
- Del análisis realizado, la forma en que se realizan los informes por parte del personal de los departamentos de producción y ventas no es la más acertada debido a que las herramientas informáticas existentes en la empresa rara vez colaboran para poder obtener de forma ágil dichos informes desde la base de datos, lo que causa que actualmente se invierta una alta cantidad de tiempo y recursos en su elaboración.
- Del estudio efectuado los actuales informes del estado de la producción y ventas no ayudan a tomar oportunas decisiones a los directivos de la empresa,

por lo que la utilización de una herramienta que procese los datos de una forma más ágil y oportuna ayudará a realizar una mejor planificación.

- El 90 % del personal que fue consultado en cuanto a que si se considera que la implementación de una nueva herramienta que procese de una mejor manera la información de la base de datos de forma ágil para que de esta forma ayude a la toma de decisiones a los departamentos de producción y ventas piensa que la implementación de una nueva herramienta seria una solución muy importante para la mejora de sus problemas al momento de realizar sus informes departamentales y por ende realizar mejores planificaciones.

Recomendaciones

- Es recomendable que los departamentos de Producción y Ventas disponga de información detallada del estado de sus respectivos procesos para evitar los frecuentes problemas que actualmente tiene la empresa para de esta forma reducir los retrasos en la producción y de esta manera se cumplan con las entregas de producto terminado de tal forma que el departamento de ventas pueda planificar de mejor manera sus entregas disminuyendo sus retrasos.
- Se recomienda proveer a los Directivos de la empresa de información detallada del estado en que se encuentra su producción así como el estado de sus ventas con el objetivo de que sus decisiones, estrategias y planificaciones sean las más acertadas beneficiando económicamente a la empresa y mejorando su competitividad dentro del mercado.
- Se debe incluir recomendablemente una herramienta informática que provea la información de las bases de datos de forma transparente para los usuarios de los diferentes departamentos de la empresa, de manera que estos de una forma sencilla puedan contar con esta información en el momento que la situación lo requiera colaborando positivamente en la planificación de sus procesos optimizando su tiempo y recursos.
- Se recomienda la implementación de una herramienta Datawarehouse ya que esta ayuda a procesar y contener los datos que son necesarios o útiles para cada departamento, entregando la información correcta a la gente indicada, de forma rápida y oportuna, que favorezca el análisis y la divulgación eficiente de los registros de la Base de Datos.

CAPÍTULO VI

LA PROPUESTA

TEMA:

DESARROLLO DE UN DATAWAREHOUSE PARA LA PLANIFICACIÓN DE LA PRODUCCIÓN Y VENTAS DE LA INDUSTRIA PAPELERA CAICEDO MIÑO CIA. LTDA. DE LA CIUDAD DE AMBATO.

Datos Informativos

Campo: Sistemas

Área: Gestión de Producción

Institución: Industria Papelera Caicedo Miño Cía. Ltda.

Dirección: Av. Pasteur 10-26 y Grecia.

Ciudad: Ambato

Antecedentes de la Propuesta

En la actualidad Industrial Papelera Caicedo Miño Cía. Ltda., carece de un datawarehouse que le permita el control y seguimiento unificado de la producción, compra de materia prima, suministros, cuenta con una base de datos que no ayuda a la optimización de estos procesos, las personas que necesitan conocer del estado de la producción, así como el estado en que se encuentran provistas sus bodegas, tienen dificultades para saber con agilidad cuánto y cuando pueden contar con estos productos.

El conteo y recuento de los productos en las bodegas de forma repetitiva causa que se consuma tiempo y recursos para obtener información de las existencias de los productos que se encuentran en bodega, así también no existe una herramienta que ayude a la planificación de la producción por lo que las órdenes de producción se realizan de acuerdo a la experiencia de producción de años pasados, lo que causa que no se elaboren los productos en base a las existencias o la demanda del mercado.

Ciertos productos en bodega son insuficientes para la demanda, mientras que otros son excesivos y sin embargo se siguen produciendo, lo que conlleva al frecuente retraso en la entrega de pedidos, el inventario de existencias de los suministros y materia prima que se utiliza para la producción llega al país por medio de importaciones, debido a la mala planificación de la producción se agotan inesperadamente causando desabastecimiento, ya que unos productos consumen más materia prima que otros.

La información que contiene el sistema es extensa debido a la cantidad de ítems que se producen, la elaboración de cualquier informe es un trabajo manual prolongado y tedioso que consume una alta inversión de tiempo y recursos, tanto que por lo general los reportes que se realizan no son 100% confiables.

Justificación

En cuanto a la relevancia del problema es necesario indicar que a nivel de empresas de manufactura y en este caso específico Industria Papelera Caicedo Miño, todavía se manejan procesos de producción de forma manual, lo que no permite que el acceso a la información y ubicación de las ordenes de producción dentro de la cadena de fabricación se lo haga de manera automatizada, 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, ya que en la fábrica se mantiene un gran volumen de documentos que vigila el seguimiento y control de las ordenes de producción, que establecen las compras de materia prima y suministros en cada uno de los centros de producción, es por ello que se hace necesario automatizar el manejo de la información para la Planificación de la Producción y Ventas que permita presentar información de primera mano, oportuna y precisa del estado de las ordenes de producción, materia prima y suministros, estableciendo de esta manera una búsqueda rápida y precisa de las ordenes de producción deseada por la jefatura de planta.

Con la información oportuna y confiable, la gerencia, 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 existencias de materia prima y suministros en las bodegas. Además los resultados obtenidos con la automatización del manejo de información de los procesos, permitirá que la jefatura de planta pueda tomar las mejores decisiones y realizar buenas planificaciones de producción y ventas.

OBJETIVOS

Objetivo General

Desarrollar un Datawarehouse para la Planificación de la Producción y Ventas de la Industrial Papelera Caicedo Miño Cía. Ltda. de la Ciudad de Ambato.

Objetivos Específicos

- Establecer las necesidades específicas que Industrial Papelera Caicedo Miño requiera para la optimización de sus procesos de producción y ventas.
- Determinar la secuencia a seguir para Desarrollar un Datawarehouse para la Planificación de la Producción y Ventas.
- Construir un banco de información que facilite la adquisición de reportes del estado de la producción y ventas de la organización.
- Fijar políticas de capacitación a los usuarios para la correcta explotación de la herramienta datawarehouse.

Análisis de Factibilidad

Luego del estudio realizado se concluye que es **factible** realizar este proyecto, para brindar confiabilidad de la información de los procesos de producción y ventas que la Industria Papelera Caicedo Miño realiza, adicionalmente cuenta con la infraestructura lógica y física además de la aprobación del desembolso de los recursos económicos presupuestados.

Factibilidad Técnica

Se cuenta con la tecnología necesaria para el desarrollo del proyecto, ya que se tiene el hardware adecuado, como se muestra en la evaluación de la infraestructura, lo que permitirá el buen desempeño del datawarehouse, en cuanto a software, se cuenta con las licencias de cada una de las herramientas utilizadas, por lo que técnicamente es factible.

Factibilidad Operacional

La empresa cuenta con la aceptación y predisposición al cambio por parte de la junta de accionistas, gerencia, personal administrativo y operativo, pero sobre todo con la base de datos que contiene los datos históricos muy necesaria para establecer un buen diseño de un Datawarehouse para la Planificación de la Producción y Ventas con el propósito de mejorar la situación actual de la fábrica.

Factibilidad Económica

Económicamente no se necesita de ninguna inversión, ya que para el proyecto se dispone de todos los recursos tecnológicos y el operacional, puesto que la persona que ejecuta el proyecto trabaja para en la misma 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.

FUNDAMENTACIÓN

Metodología Datawarehouse

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. 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 Business Intelligence.

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

Arquitectura Típica de un Datawarehouse

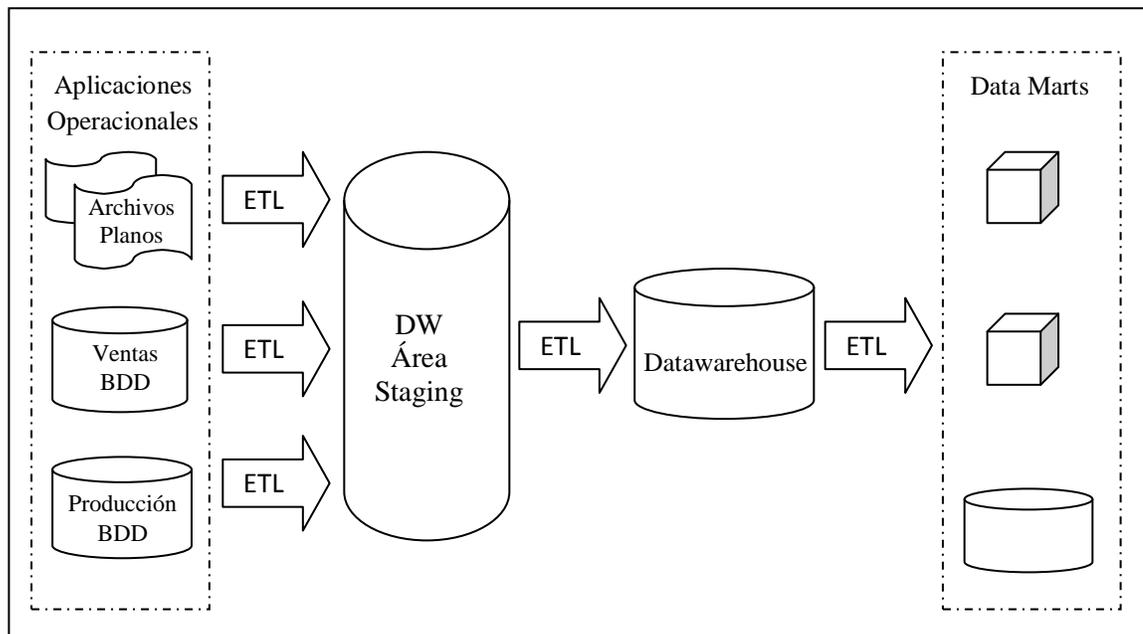


Figura N.-6.1. Procesos Datawarehouse
Elaborado por: Juan Carlos Pérez

El término Datawarehouse fue acuñado por primera vez por Bill Inmon, y se traduce literalmente como *almacén de datos*. No obstante, y como cabe suponer, es mucho más que eso. Según definió el propio Bill Inmon, un datawarehouse se caracteriza por ser(Inmon, 1992):

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

- **Temático:** sólo los datos necesarios para el proceso de generación del conocimiento del negocio se integran desde el entorno operacional. Los datos se organizan por temas para facilitar su acceso y entendimiento por parte de los usuarios finales. Por ejemplo, todos los datos sobre clientes pueden ser consolidados en una única tabla del datawarehouse. De esta forma, las peticiones de información sobre clientes serán más fáciles de responder dado que toda la información reside en el mismo lugar.
- **Histórico:** el tiempo es parte implícita de la información contenida en un datawarehouse. En los sistemas operacionales, los datos siempre reflejan el estado de la actividad del negocio en el momento presente. Por el contrario, la información almacenada en el datawarehouse sirve, entre otras cosas, para realizar análisis de tendencias. Por lo tanto, el datawarehouse se carga con los distintos valores que toma una variable en el tiempo para permitir comparaciones.
- **No volátil:** el almacén de información de un datawarehouse existe para ser leído, pero no modificado. La información es por tanto permanente, significando la actualización del datawarehouse la incorporación de los últimos valores que tomaron las distintas variables contenidas en él sin ningún tipo de acción sobre lo que ya existía.

Otra característica del datawarehouse es que contiene metadatos, es decir, datos sobre los datos. Los metadatos permiten saber la procedencia de la información, su periodicidad de refresco, su fiabilidad, forma de cálculo... etc.

Metadatos: contiene datos relativos a los datos, concepto que se ha venido asociando al término de metadatos. Los metadatos permiten mantener información de la procedencia de la información, la periodicidad de refresco, su fiabilidad, forma de cálculo, etc., relativa a los datos de nuestro almacén

Los metadatos serán los que permiten simplificar y automatizar la obtención de la información desde los sistemas operacionales a los sistemas informacionales.

Los objetivos que deben cumplir los metadatos, según el colectivo al que va dirigido, son:

- Dar soporte al usuario final, ayudándole a acceder al datawarehouse con su propio lenguaje de negocio, indicando qué información hay y qué significado tiene. Ayudar a construir consultas, informes y análisis, mediante herramientas.
- Dar soporte a los responsables técnicos del datawarehouse en aspectos de auditoría, gestión de la información histórica, administración del datawarehouse, elaboración de programas de extracción de la información, especificación de las interfaces para la realimentación a los sistemas operacionales de los resultados obtenidos... etc.

Por último, destacar que para comprender íntegramente el concepto de datawarehouse, es importante entender cuál es el proceso de construcción del mismo, denominado ETL (Extracción, Transformación y Carga), a partir de los sistemas operacionales de una compañía:

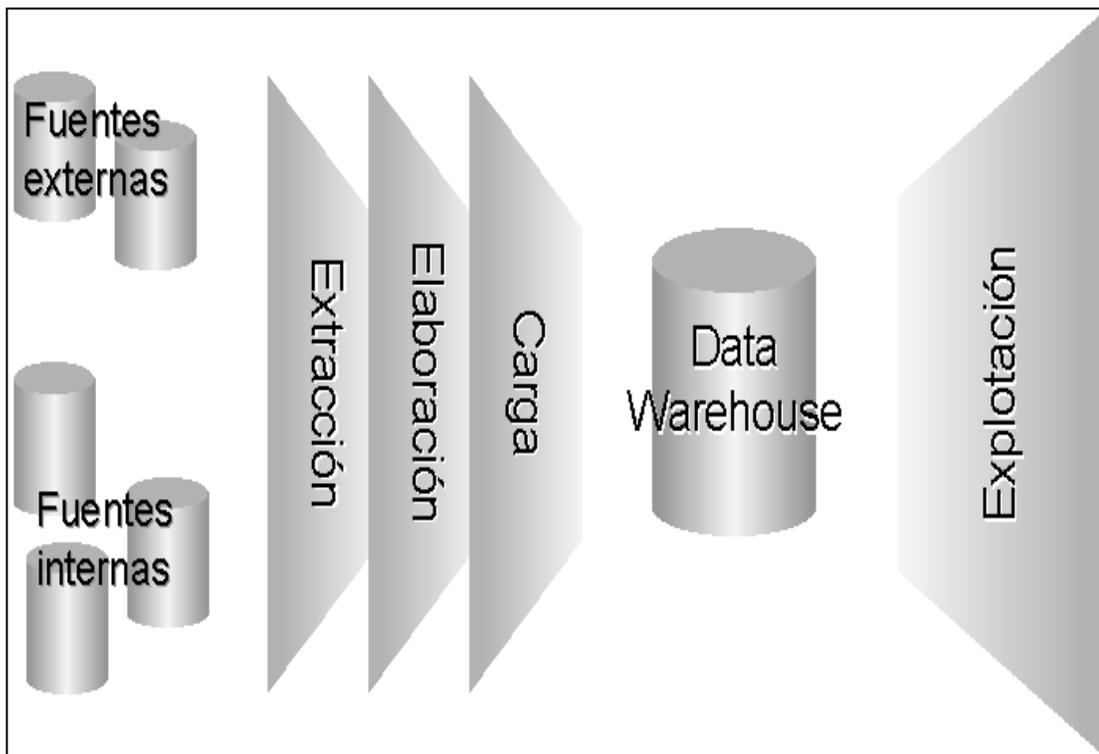


Figura N.-6.2. Proceso Construcción Datawarehouse
Elaborado por: Juan Carlos Pérez

E (Extraction): Obtención de información de las distintas fuentes tanto internas como externas.

T (Transformation): Filtrado, limpieza, depuración, homogeneización y agrupación de la información.

L (Load): Organización y actualización de los datos y los metadatos en la base de datos.

Explotación: Extracción y análisis de la información en los distintos niveles de agrupación.

Data Marts: Una vez contando con la base de información empresarial integrada y, a partir de esta, se crean subconjuntos de datos con el propósito de ayudar

aque un área específica dentro del negocio pueda tomar mejores decisiones. Los datos existentes en este contexto pueden ser resumidos, agrupados, explorados y reportados de múltiples formas para que diversos grupos de usuarios realicen la explotación de los mismos.

Es un modelo multidimensional basado en tecnología OLAP, incluyendo variables claves y los indicadores claves para el proceso de toma de decisiones.

Algunas ventajas de la construcción del Data Mart:

- Son más simples de implementar que un Datawarehouse
- Pequeños conjuntos de datos y, en consecuencia, menor necesidad de recursos
- Se encuentran más rápidamente las necesidades de las Unidades de Negocio
- Consultas sql más rápidos por menor volumen de datos Como desventaja se puede decir que, en algunos casos, añaden tiempo al proceso de actualización.

En síntesis, son pequeños datawarehouse centrados en un tema o un área de negocio específico. En muchos casos, los datawarehouse comienzan siendo Data Marts con el objetivo de minimizar los riesgos para luego ir ampliando su espectro gradualmente.

OLTP (On-Line Transaction Processing): Son aplicaciones que definen el comportamiento habitual de un entorno operacional de gestión y ejecutan las operaciones del día a día. Algunas de las características más comunes de este tipo de transacciones podrían ser:

- Altas/Bajas/Modificaciones
- Consultas rápidas, escuetas y predecibles
- Poco volumen de información e información disgregada

- Transacciones rápidas
- Gran nivel de concurrencia
- Modo de actualización on-line
- Baja redundancia de datos

Algunos ejemplos de este tipo de aplicaciones son:

- Compras
- Ventas
- Inventario
- Sueldos

Consolidación: Es la parte del proceso de Datawarehouse que se encarga de producir el cambio de los sistemas OLTP a las Bases de Datos OLAP. Consolidan datos de aplicaciones no integradas, resumen datos desgregados y los transforman. Este proceso está compuesto por tres pasos:

Validación de Consistencia de los datos

- Comprueba la validez de los datos en el entorno operacional
- Inconsistencia entre distintas aplicaciones dentro del sistema

Mecanismos de Consolidación

Refresco de datos: Volcado completo de los datos procedentes del sistema operacional entre el Cliente y el Servidor. Actúa como traductor entre distintas tecnologías. Permite que dos o más sistemas trabajen juntos aunque no estén preparados para ello.

Algunas de sus características más relevantes son:

- Un mismo middleware puede poseer más de una máquina virtual para soportar diferentes entornos de desarrollo
- Gestiona las comunicaciones con el Datawarehouse
- Controla la concurrencia y controla los procesos Batch
- Posee diversos controladores de Bases de Datos para acceder a las distintas fuentes, por ejemplo, Oracle, Sybase, AS400, etc.

Ejemplos:

- Monitores de procesamiento de transacciones
- Convertidores de datos
- Replicación de datos
- Controladores de comunicación

Factores técnicos

- Mecanismo de transporte
- Tiempos de carga
- Reformato de datos

Middleware: Es un software que reside físicamente en un Cliente y en un Servidor de Comunicaciones, localizado

OLAP (On-Line Analytical Process): Son aplicaciones que se encargan de analizar datos del negocio para generar información táctica y estratégica que sirve de soporte para la toma de decisiones. Mientras que las transacciones OLTP utilizan bases

de Datos Relacionales u otro tipo de archivos, OLAP logra su máxima eficiencia y flexibilidad operando sobre Bases de datos Multidimensionales.

Podemos nombrar las siguientes características como las más sobresalientes de estas aplicaciones:

- Estructura de datos transparente al usuario
- Solo Consulta, trabajan sobre la información operacional generada por los sistemas OLTP
- Consultas sobre grandes volúmenes de datos no predecibles
- Información histórica
- Modo de actualización Batch
- Alta redundancia de datos para facilitar la generación de consultas y obtener buenos tiempos de respuesta
- Poderoso Back-end analítico para múltiples aplicaciones de usuarios Trabaja con resúmenes de miles de registros condensados en una sola respuesta

Aplicaciones OLTP y OLAP: Mientras que las aplicaciones OLTP se caracterizan por estar actualizadas constantemente por varios usuarios a través de transacciones operacionales sobre datos individuales, las aplicaciones OLAP son utilizadas por personal de niveles ejecutivos que requieren datos con alto grado de agregación y desde distintas perspectivas (dimensiones), como ser: totales de venta por región, por producto, por período de tiempo, ..., etc.

OLTP	OLAP
Atomizado	Sumarizado
Datos Históricos	Datos Actuales
Un registro a la vez	Muchos registros a la vez
Orientado a la información operativa	Orientado a la información estratégica
Consultas simples predefinidas	Consultas ad-hoc
Volumen de datos acotados	Grandes volúmenes de datos

Tabla N.-6.1. Procesos Datawarehouse
Elaborado por: Juan Carlos Pérez

Una de las claves del éxito en la construcción de un datawarehouse es el desarrollo de forma gradual, seleccionando a un departamento usuario como piloto y expandiendo progresivamente el almacén de datos a los demás usuarios. Por ello es importante elegir este usuario inicial o piloto, siendo importante que sea un departamento con pocos usuarios, en el que la necesidad de este tipo de sistemas es muy alta y se pueda obtener y medir resultados a corto plazo.

Data Mining: Su relación con el datawarehouse y las técnicas de Data Mining son el resultado de un largo proceso de investigación y desarrollo de productos orientados al almacenamiento, extracción y análisis de datos. Esta evolución comenzó cuando los datos de negocios fueron almacenados por primera vez en computadoras, y continuó con mejoras en el acceso a los datos, y más recientemente con tecnologías generadas para permitir a los usuarios navegar a través de los datos en tiempo real.

Principales aportaciones de un Datawarehouse

Proporciona una herramienta para la toma de decisiones en cualquier área funcional, basándose en información integrada y global del negocio.(Galarraga, 2000):

- Facilita la aplicación de técnicas estadísticas de análisis y modelización para encontrar relaciones ocultas entre los datos del almacén; obteniendo un valor añadido para el negocio de dicha información.
- Proporciona la capacidad de aprender de los datos del pasado y de predecir situaciones futuras en diversos escenarios.
- Simplifica dentro de la empresa la implantación de sistemas de gestión integral de la relación con el cliente.
- Supone una optimización tecnológica y económica en entornos de Centro de Información, estadística o de generación de informes con retornos de la inversión espectaculares.

Tipos de Metodologías Datawarehouse

Metodología de Kimball

Esta metodología denominada Ciclo de Vida Dimensional del Negocio (Business Dimensional Lifecycle). Este ciclo de vida del proyecto de DW, está basado en cuatro principios básicos(Rivadera):

- *Centrarse en el negocio:* Hay que concentrarse en la identificación de los requerimientos del negocio y su valor asociado, y usar estos esfuerzos para desarrollar relaciones sólidas con el negocio, agudizando el análisis del mismo y la competencia consultiva de los implementadores.
- *Construir una infraestructura de información adecuada:* Diseñar una base de información única, integrada, fácil de usar, de alto rendimiento donde se reflejará la amplia gama de requerimientos de negocio identificados en la empresa.
- *Realizar entregas en incrementos significativos:* crear el almacén de datos (DW) en incrementos entregables en plazos de 6 a 12 meses. Hay que usar el valor de negocio de cada elemento identificado para determinar el orden de aplicación de los incrementos. En esto la metodología se parece a las metodologías ágiles de construcción de software.
- *Ofrecer la solución completa:* proporcionar todos los elementos necesarios para entregar valor a los usuarios de negocios. Para comenzar, esto significa tener un almacén de datos sólido, bien diseñado, con calidad probada, y accesible. También se deberá entregar herramientas de consulta ad hoc,

aplicaciones para informes y análisis avanzado, capacitación, soporte, sitio web y documentación.

La construcción de una solución de DW es sumamente compleja, y Kimball nos propone una metodología que nos ayuda a simplificar esa complejidad. Las tareas de esta metodología (ciclo de vida) se muestran en la siguiente figura:

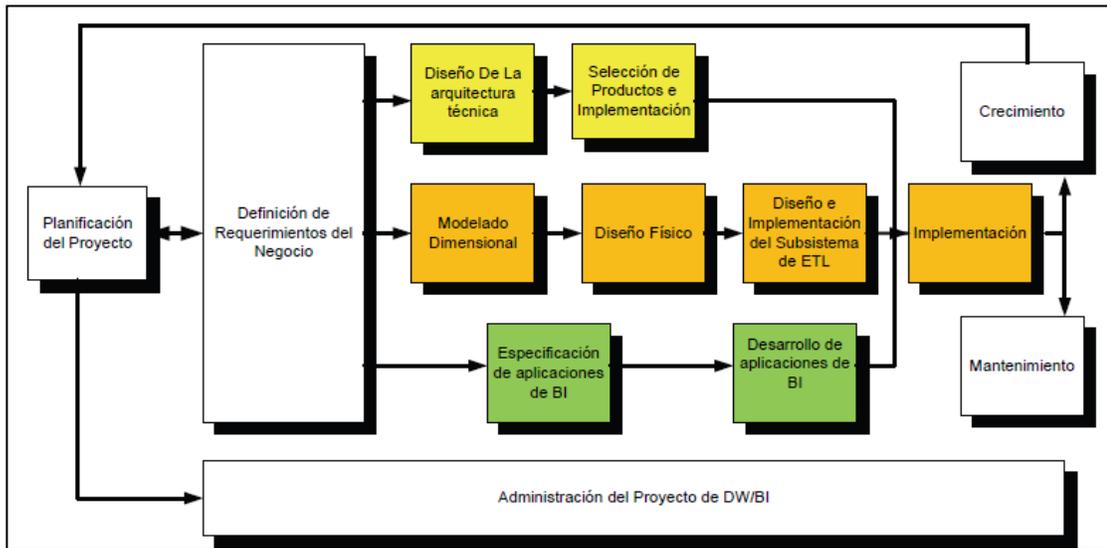


Figura N. 6.3. Metodología de Kimball
Fuente: (Rivadera)

Metodología Hefesto.

La metodología está orientada a la construcción de Datawarehouse para Análisis Dimensional (OLAP) y comprende las siguientes fases, que podemos ver en el gráfico siguiente(Espinosa, 2009):



Figura N. 6.4. Metodología Hefesto
Fuente: (Espinosa, 2009)

Análisis de requerimientos:

- Identificar preguntas para las que queremos tener respuesta y los objetivos que se quieren conseguir con el nuevo sistema.
- Analizar las preguntas para determinar las perspectivas de análisis y los indicadores de negocio.
- Diseñar el modelo conceptual, que incluirá las perspectivas e indicadores identificados. A través del modelo se podrán alcanzar claramente cuáles son

los alcances del proyecto, y será un punto de partida con alto nivel de definición para su exposición a los usuarios y responsables.

Análisis de los sistemas transaccionales:

- Determinación de indicadores: identificar el origen de los indicadores en los sistemas transaccionales y determinar la forma de su cálculo.
- Correspondencias: establecer correspondencias entre los elementos definidos en el modelo conceptual y las fuentes de datos existentes en los OLTP (sistemas transaccionales).
- Definición del nivel de granularidad: nivel de detalle de los datos a obtener para cada dimensión de análisis.
- Modelo conceptual ampliado con los campos identificados para cada perspectiva.

Modelo lógico del ETL:

- Tipo de modelo lógico del DW: selección del tipo de esquema que utilizaremos (estrella, copo de nieve, etc.).
- Tabla de dimensiones: Construcción de las tablas de dimensiones para cada una de las perspectivas de análisis considerada.
- Tablas de Hechos: definición de las tablas de hechos que contendrá la información a partir de los cuales construiremos los indicadores de análisis.
- Uniones: relaciones entre las tablas de dimensiones y las tablas de hechos.

Procesos ETL: análisis, definición y desarrollo de todos aquellos procesos necesarios para la extracción, transformación y carga de datos desde los sistemas origen para llenar el Datawarehouse.

Performance y mantenimiento del DW: ajustes en el diseño del DW y mantenimiento en el tiempo.

Metodología “Rapid Warehousing Methodology”. (SAS Institute, 2001)

Es una metodología propuesta por el Instituto SAS⁴ (EEUU) es una metodología que compacta todas las áreas tomadas en cuenta por (Kimball,1998), y las plantea en 5 fases:

- Definición de objetivos
- Definición de requerimientos
- Diseño y modelización
- Implementación
- Implantación o Gestión del proyecto.

La propuesta de este proyecto se fundamenta 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.

⁴SAS es la compañía líder de software de Business Analytics y servicios y el mayor proveedor independiente del mercado de Business Intelligence.

Metodología

La metodología propuesta para la realización del presente proyecto consta de cuatro etapas definidas en: análisis, diseño, implementación e implantación. En adelante se describirán a detalle cada uno de estas etapas.

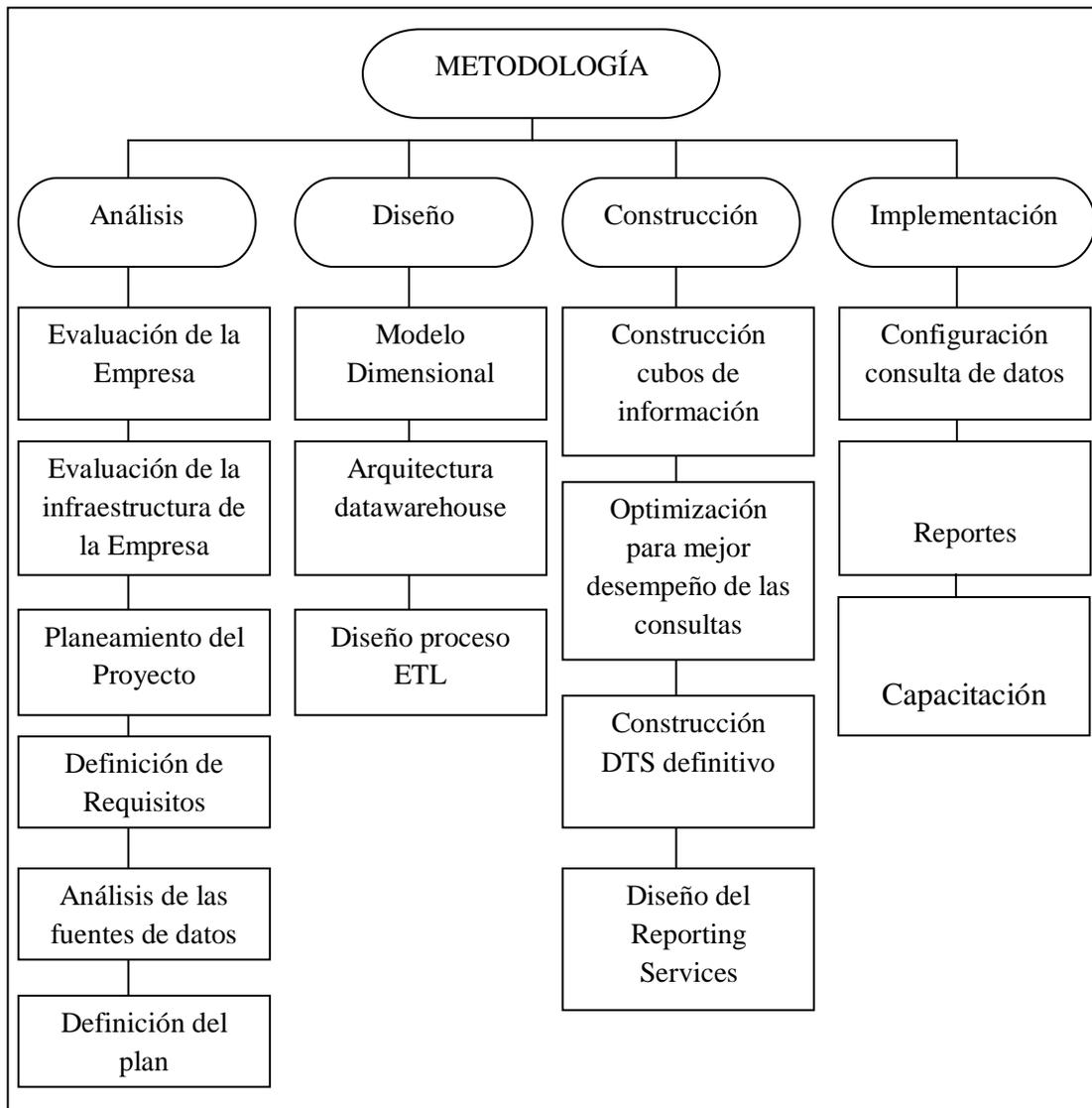


Figura N.- 6.5. Diagrama Metodología
Elaborado por: Ing. Juan Carlos Pérez

Análisis

Se realiza una evaluación a la organización de su infraestructura. En esta etapa se define la planeación del proyecto lo que incluye sus requisitos, herramientas a utilizarse y un análisis de datos.

Evaluación de la Organización

En esta etapa se define los problemas de la organización y se propone una solución, cada aplicación es justificada, define los beneficios, plantea la solución a los problemas de la organización y define las ventajas que obtendrá la empresa con la aplicación de las soluciones propuestas.

Evaluación de la Infraestructura de la Organización

Debido a que las aplicaciones del presente proyecto afectan a toda la organización, es necesaria la disponibilidad de una adecuada infraestructura para soportar dichas aplicaciones, estos componentes existen actualmente:

- ***Infraestructura técnica:*** dentro de esta infraestructura incluyen el hardware, software, base de datos, sistemas operativos, sistema de redes, repositorios metadatos y sistemas utilitarios.

Hardware

- Servidor: SUPERMICRO
- Modelo: X7DWU
- Numero de Procesadores: 8
- Memoria RAM: 8 Gb.
- Disco Duro: 1 Tb.

Software

- Sistema Operativo: Windows 2008 Server
- Motor de Base de Datos: SQL Server 2008
- Utilitarios: Analysis Services 2000, DTS SQL 2000, Microsoft Excel.

Organigrama Estructural de Industrial Papelera Caicedo Miño actualmente operativo.

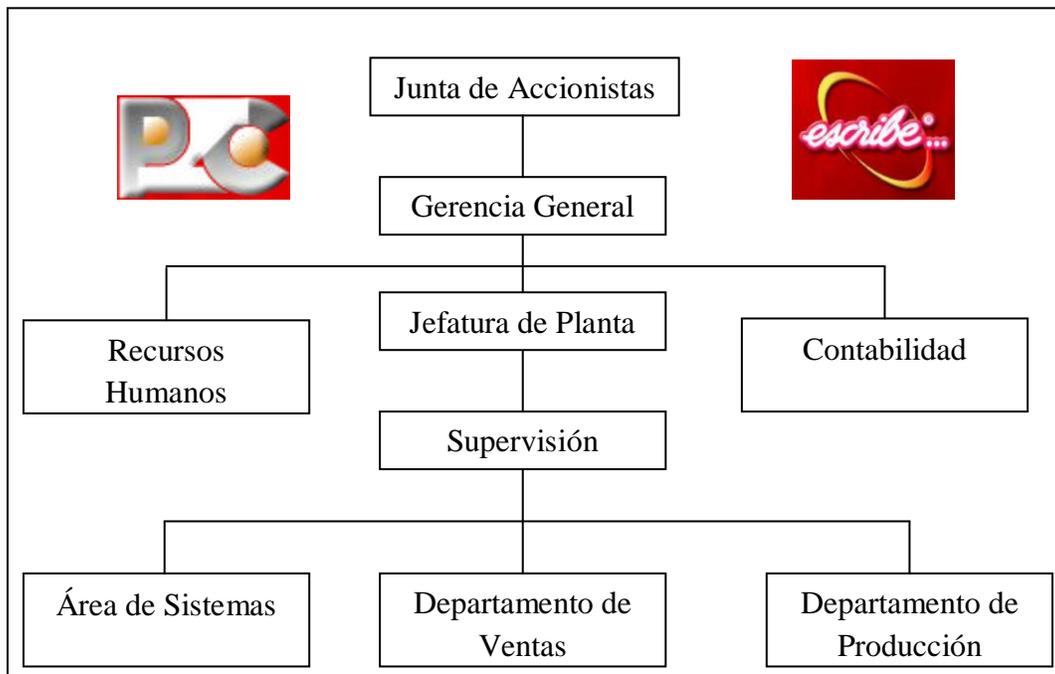


Figura N.- 6.6. Diagrama Estructural INPAC
Elaborado por: Ing. Juan Carlos Pérez

Planteamiento del proyecto

Luego del estudio efectuado dentro de la organización a través de las encuestas y entrevistas realizadas a los distintos departamentos se determinó que la falta de información ágil y confiable incide negativamente en la planificación de la producción y ventas de la empresa, motivo por el cual se estipuló que para mitigar esta problemática se plantean los siguientes requerimientos:

- Información histórica de las cobranzas realizadas por los vendedores
- información histórica de total recaudado en una determinada región
- información histórica de pedidos por vendedor
- información histórica de pedidos por zona
- información histórica de pedidos por producto
- información histórica de la producción
- información histórica de ventas por vendedor
- información histórica de ventas por región
- información histórica de las comisiones por vendedor
- Consultas en línea

Análisis de las fuentes de datos

Una vez determinados los requerimientos por parte de los usuarios de la empresa, se verifica que existan los datos en el sistema transaccional actual que nos permita generar la información solicitada, para lo cual se detalla las tablas de la base de datos transaccional.

Las tablas que intervienen en el proyecto según los requerimientos son las siguientes:

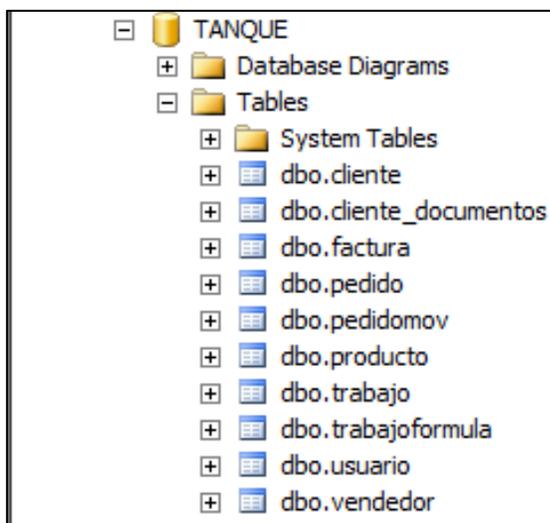


Figura N.- 6.7. Tablas usadas de la Base de Datos Transaccional
Elaborado por: Ing. Juan Carlos Pérez

trabajo	pedidomov	trabajoformula	pedido
id_trabajo	id_pedido	id_trabajo	id_pedido
fecha	id_producto	peso	id_cliente
descripcion	cantidad	ancho	fecha
fechafin	valor_unitario	largo	subtotal
numhojas	valor_total	id_producto	por_descuento
gramaje	id_unidad		val_descuento
peso	despacho		iva
bobinas	desuento		total
ancho	iva		egreso1
corte	valor_real		observacion
lgresma	secuencial		id_vend
lgpeso	valor_unitario_ori		status
resorden			fecha_despachar
teorico			comision
division			val_descuentoef
pastas			numeroproveedor
bobina			fn
numpastas			tel_otro
numdivision			dir_otro
bobres			nom_otro
por_desperdido			dir_otro
numdivision2			nuc_otro
in_muerto			raz_otro

cliente_documentos	factura
tipo	id_factura
numero	id_cliente
id_cliente	fecha
iva	subtotal
total	por_descuento
devolucion	val_descuento
abono	iva
	total
	egreso1
	egreso2
	egreso3
	tipo
	numfaprov
	retencion

Figura N.- 6.8. Diagrama Tablas Base de Datos Transaccional
Elaborado por: Ing. Juan Carlos Pérez

En el siguiente cuadro se realiza una descripción del contenido de cada uno de las tablas que son utilizadas para la construcción del Datawarehouse, cabe destacar que la base de datos se encuentra en un motor de base de datos SQL Server 2005.

NOMBRE	DESCRIPCION	AREA
dbo.ciente	Contiene la información detallada de todos los clientes de la empresa.	General
dbo.ciente_documentos	Se encuentra la información documental del los todos los clientes de la empresa.	Ventas
dbo.factura	Contiene el detalle histórico de los registros de las facturas de todos los clientes de la empresa.	Ventas
dbo.pedido	En esta tabla se registran el histórico de los pedidos realizados por parte de todos los clientes de la empresa.	Producción, Ventas
dbo.pedidomov	Todos los movimientos de la empresa se registran en forma detallada en esta tabla de la base de datos.	Producción, Ventas
dbo.producto	Cada producto tiene sus respectivas características, precios, costos, etc. estos se registran de forma detallada en esta tabla	Producción
dbo.trabajo	Las ordenes de trabajo se registran en un histórico, los mismos que se almacenan en esta tabla.	Producción
dbo.trabajoformula	Para la realización de cada producto incluye la aplicación de una fórmula para la obtención del mismo, estos se registran en esta tabla	Producción

dbo.usuario	Los usuarios del sistema, se registran en esta tabla.	General
dbo.vendedor	Cada vendedor tiene información detallada de sus movimientos, estos se registran en esta tabla.	Ventas

Tabla N.- 6.2. Descripción Tablas de la Base de Datos Actual
Elaborado por: Ing. Juan Carlos Pérez

Definición del plan

Extraer de los datos fuente desde la BDD operacional, el cual su origen está alojada en SQL server 2008 con el nombre TR_BASE1.

1. Crear un repositorio llamado TANQUE para la carga de estos datos.
2. Crear el datawarehouse DW_DWH, lo que incluye: Tabla de Hechos, Store Procedures para la carga de las tablas y filtrado de datos. Con el fin de llenar los datos requeridos desde el repositorio TANQUE.
3. Crear un repositorio llamado DW_Comercial para el alojamiento de la información (metadatos).
4. Crear Data Marts a partir del DW_DWH según los requerimientos analizados.
5. crear procesos ETL para la carga de datos absoluta
6. Automatizar el proceso ETL cada Data Mart
7. Elaborar los reportes que muestre la información requerida para la planificación de la producción y ventas en las herramientas como Excel e Internet Explorer.

Diseño

Modelo Dimensional

DataMart Cobranzas: El DataMart Cobranzas, contiene la información de las cobranzas realizadas a los clientes por los vendedores de la empresa en un periodo de tiempo.

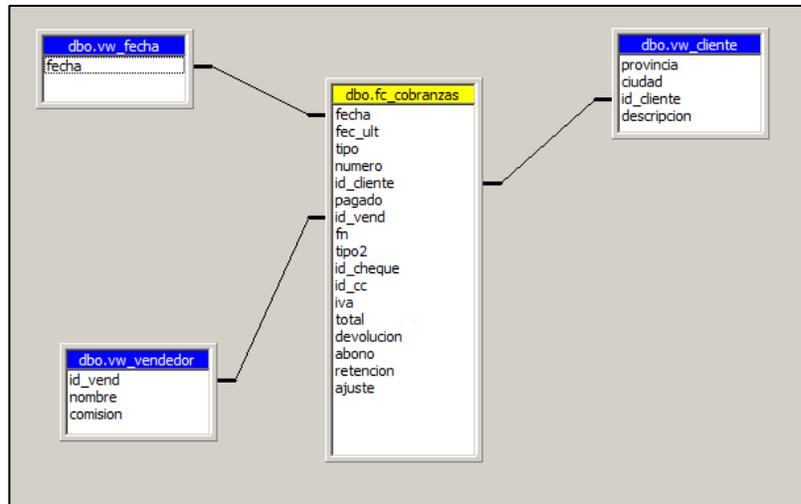


Figura N.-6.9. Data Mart Cobranza
Elaborado por: Juan Carlos Pérez

DataMart Pedidos: Contiene la información detallada de la cantidad de productos que un cliente ha pedido a un vendedor en una determinada fecha.

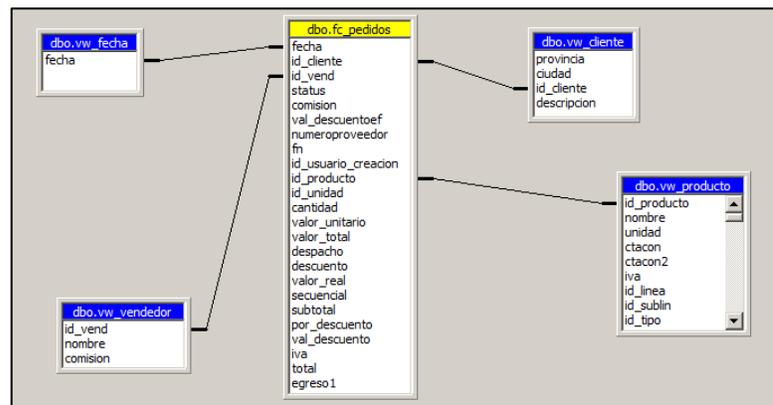


Figura N.-6.10. Data Mart Pedidos
Elaborado por: Juan Carlos Pérez

DataMartProducción: La información en detalle de cada uno de los productos en el proceso de producción se encuentra en este DataMart.

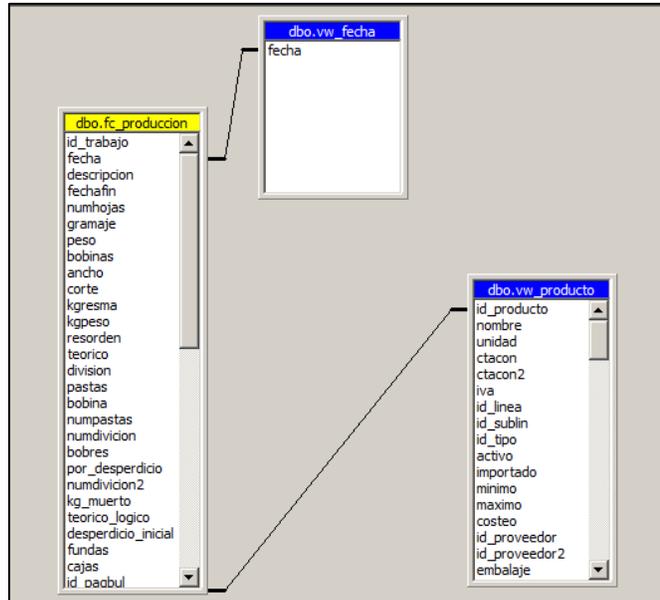


Figura N.-6.11.Data MartProducción
Elaborado por: Juan Carlos Pérez

DataMart Ventas: Este cubo de información contiene los registros de lo que se ha vendido a los clientes por parte de cada vendedor.

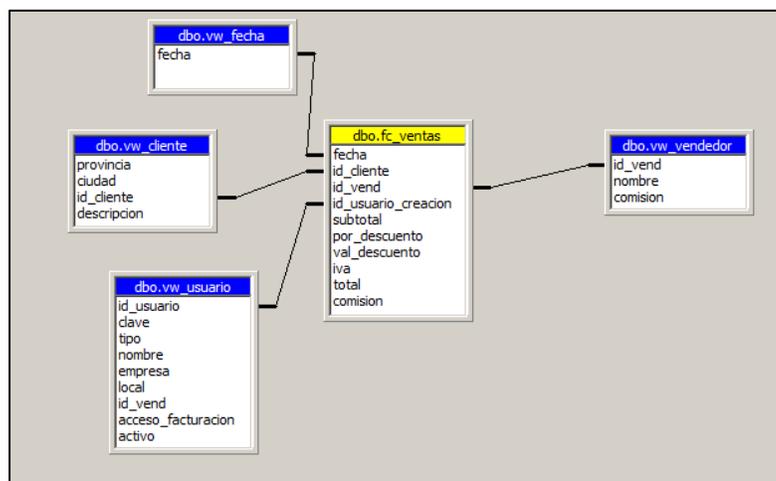


Figura N.- 6.12.Data Mart Ventas
Elaborado por: Juan Carlos Pérez

Arquitectura

En el diseño la arquitectura del datawarehouse se plantea los diferentes entornos por los que pasan los datos en su camino hacia su Data Mart o cubo de destino. Dada la cantidad de transformaciones que se realizan, además de cumplir su función de soporte a los requerimientos analíticos, realiza una función de integración de datos que van a conformar el data Mart y que van a ser consultados por los usuarios de la empresa mediante reportes. Nuestro proyecto se suma al enfoque presentado por (Kimball,2001) que describe una arquitectura de comunicación entre los diferentes data Marts basada en Dimensiones conformadas y definición conformada de hechos.

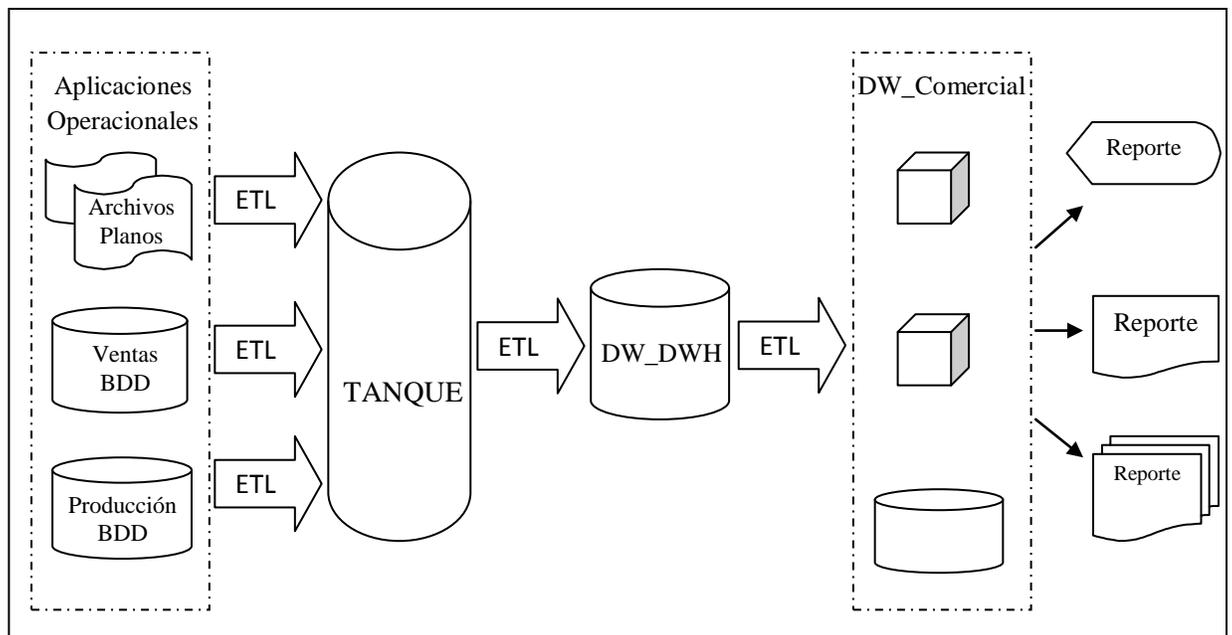


Figura N.- 6.13.Arquitectura Datawarehouse
Elaborado por: Juan Carlos Pérez

Donde las aplicaciones operacionales alimentadas por los datos del día a día son extraídas hacia nuestra zona de extracción y transformación de datos TANQUE

por medio de los procesos ETL, para luego enviarlas hacia el datawarehouse DW_DWH, mismo que provee los datos para nuestros cubos de información del servidor de presentación, desde donde se realizan los reportes que se requieren por los usuarios.

Diseño de Procesos ETL

Luego de realizado el análisis de nuestras fuentes de datos, y hemos definido el conjunto de dimensiones y tablas de hecho que forma nuestro datawarehouse, se estructura formalmente los procesos que nos permite alimentar nuestro datawarehouse desde las fuentes de datos. Se define los procesos que nos permitan mapear los datos desde los sistemas fuentes correspondientes hacia el data warehouse, El primer proceso ETL definirá el proceso de extracción, transformación y carga inicial, es decir, el proceso con el cual alimentamos inicialmente.

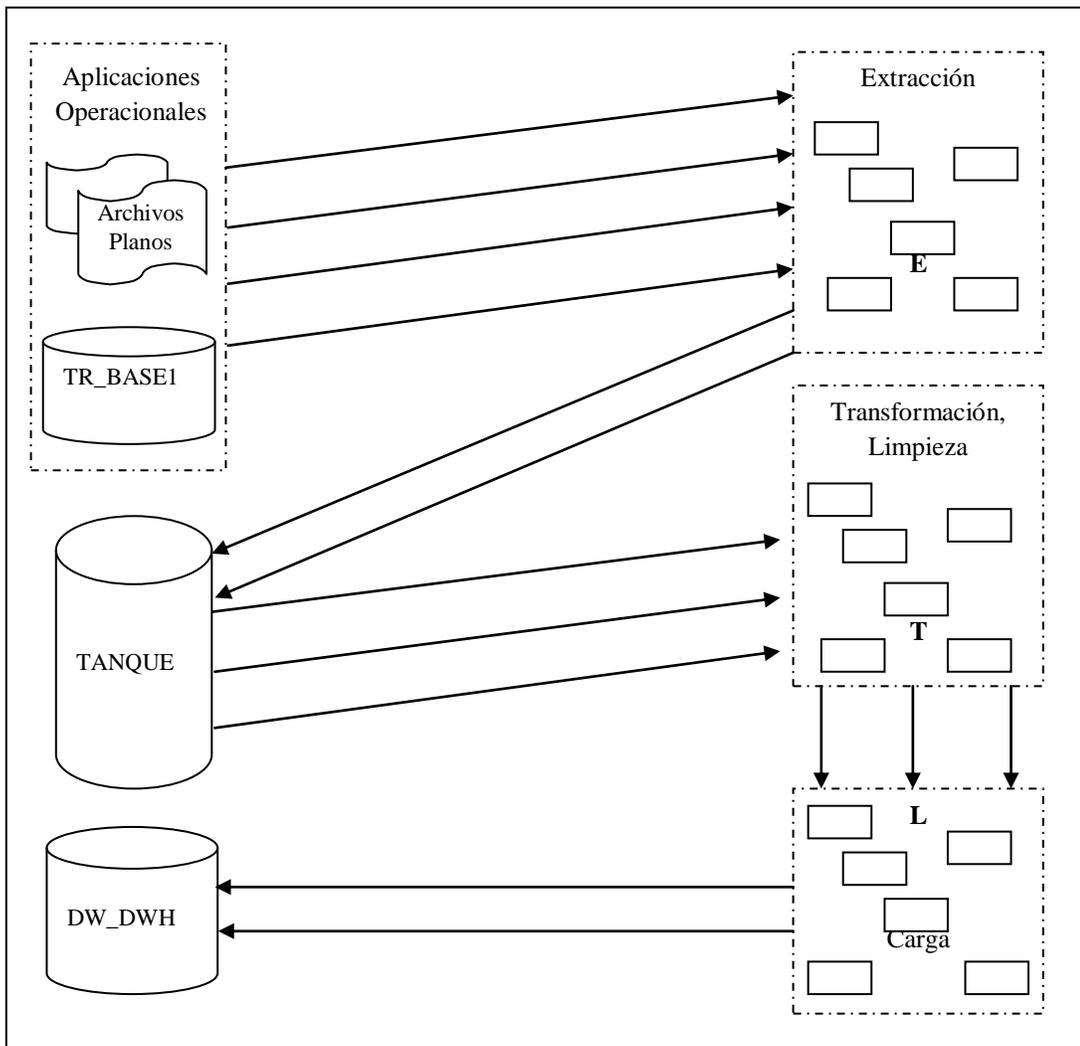


Figura N.-6.14. Procesos ETL
 Elaborado por: Juan Carlos Pérez

Extracción

Fuentes: La fuente del proyecto se encuentra en un motor de base de datos relacional SQL Server 2008 denominada TR_BASE1.

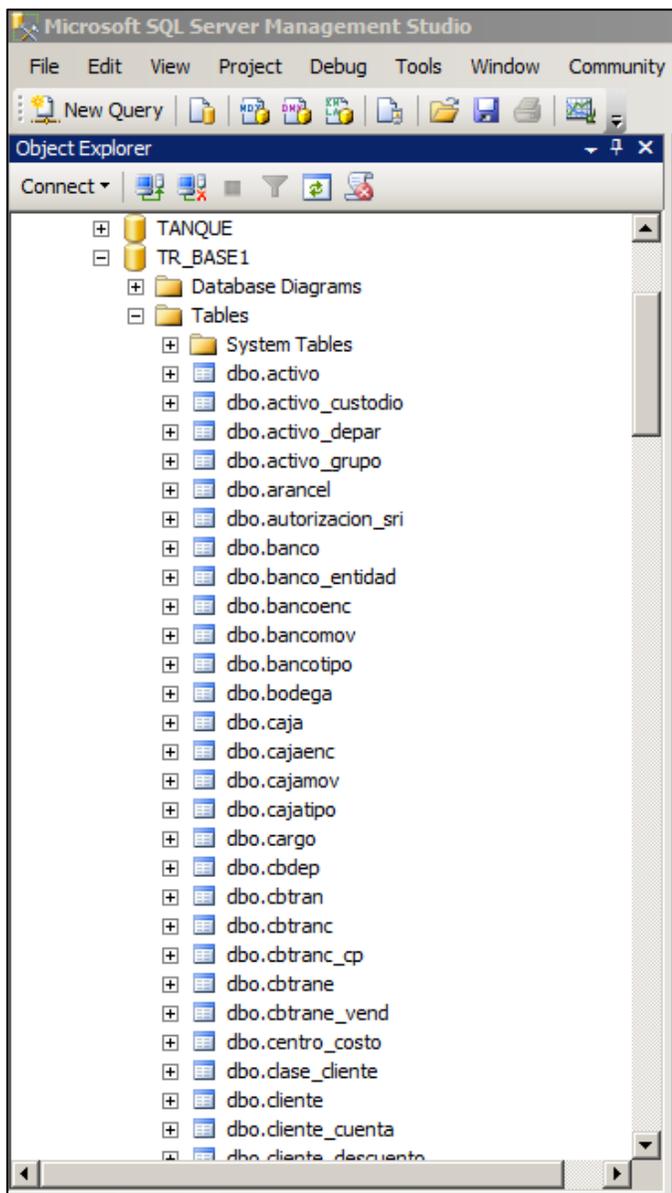


Figura N.-6.15. Base de Datos Fuente
Elaborado por: Juan Carlos Pérez

Origen: Se especifica una nueva conexión o una existente en caso de haberlo al origen de datos, se apunta a la base de datos origen.

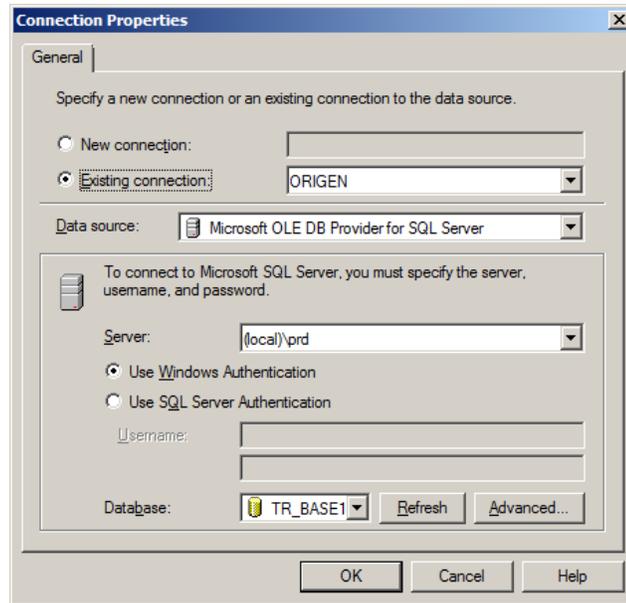


Figura N.-6.16. Propiedades Origen
Elaborado por: Juan Carlos Pérez

Destino: Especificar una conexión a la base de datos Destino.

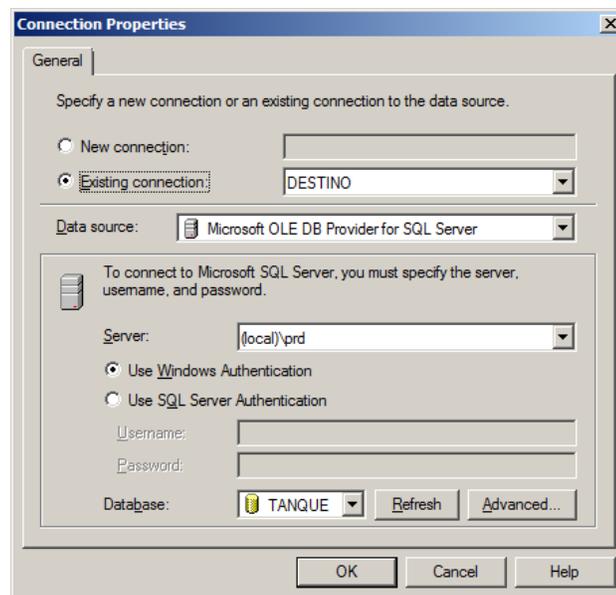


Figura N.-6.17.Propiedades Destino
Elaborado por: Juan Carlos Pérez

Se ingresa el nombre de la actividad, insertando la consulta que se desea obtener.

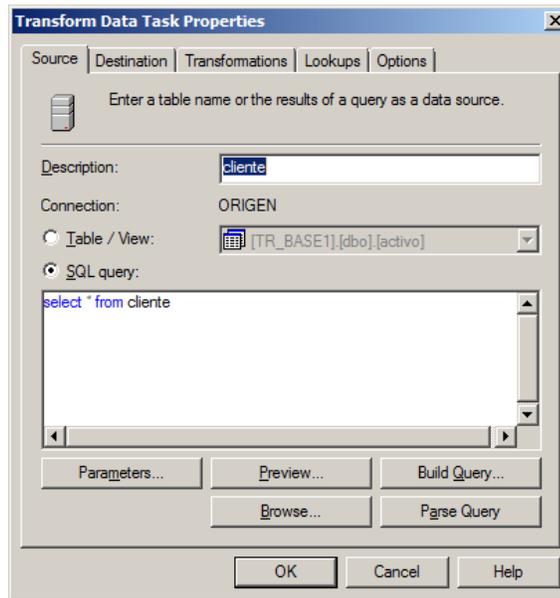


Figura N. 6.18.Propiedades SQL origen
Elaborado por: Juan Carlos Pérez

Para el destino se almacenan los resultados para la transformación, tomando en cuenta los tipos de datos.

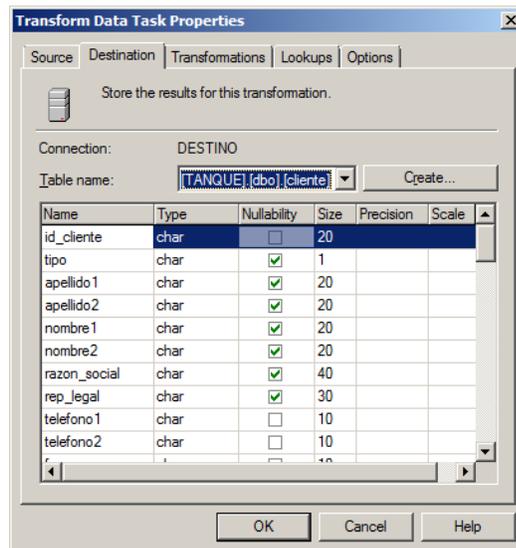


Figura N.-6.19. Selección Tablas Destino
Elaborado por: Juan Carlos Pérez

Se define las transformaciones entre el origen y el destino de los datos, en este caso se corresponden.

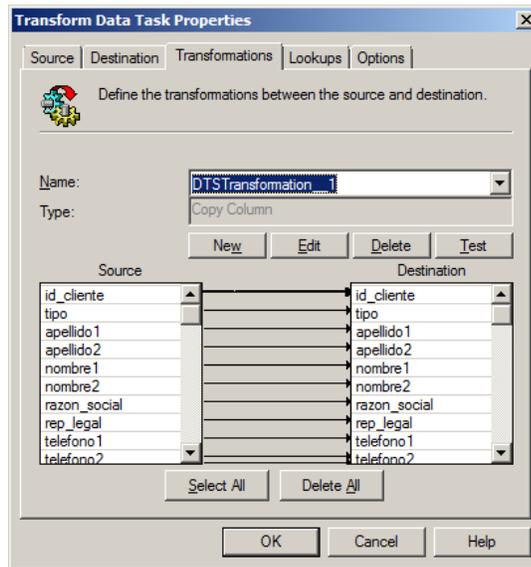


Figura N.-6.20. Propiedades Transform Data Task
Elaborado por: Juan Carlos Pérez

Se configura las sentencias de vaciado de datos, para de esta forma evitar duplicar la información.

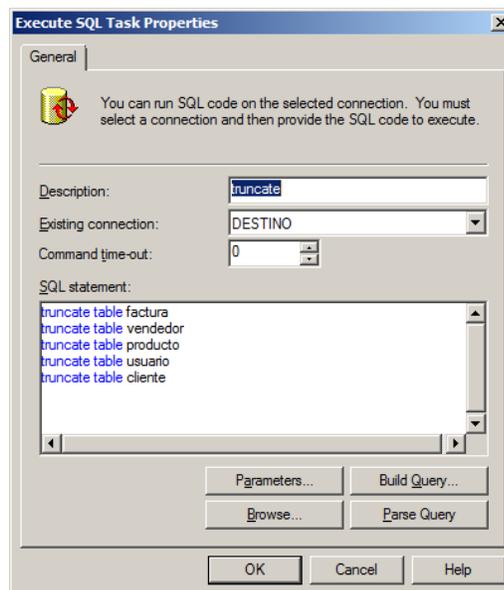


Figura N.-6.21.Sentencias de vaciado de datos
Elaborado por: Juan Carlos Pérez

Diseño de la extracción: Se diseña la extracción de los datos en la herramienta SQL Server Enterprise Manager

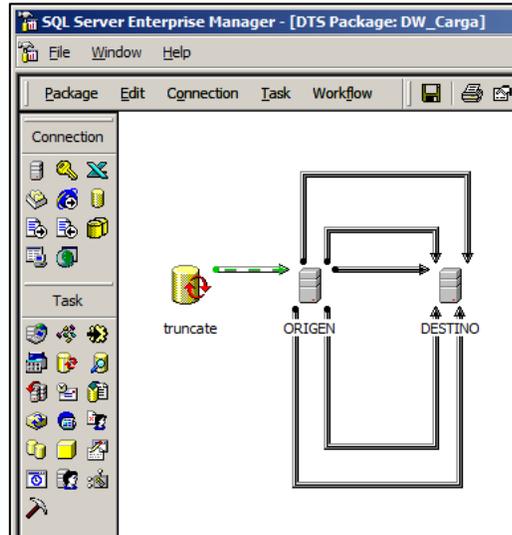


Figura N.-6.22.Diseño de la Extracción
Elaborado por: Juan Carlos Pérez

Transformación y Carga

Creación de las bases de datos:

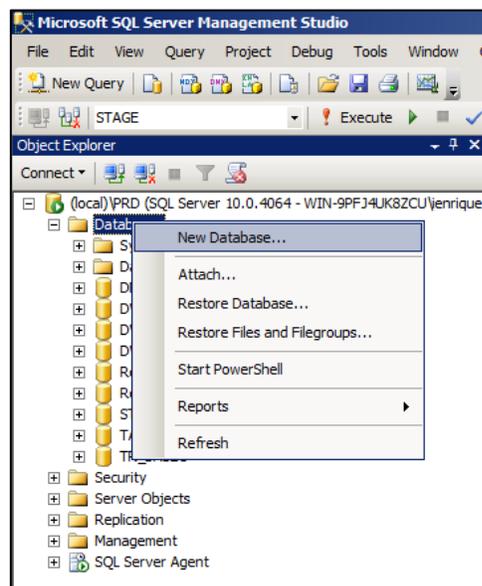


Figura N.-6.23.Creación de las Bases de Datos
Elaborado por: Juan Carlos Pérez

Base repositorio de paso: Se crea la base de datos que servirá como repositorio de paso el mismo que la denominamos TANQUE

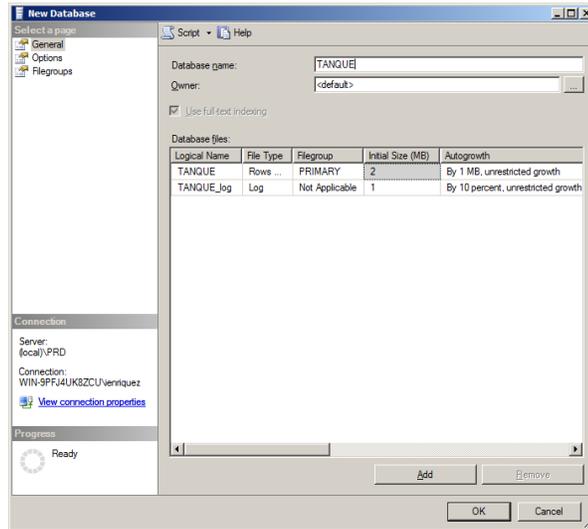


Figura N.-6.24.Base Repositorio de Paso
Elaborado por: Juan Carlos Pérez

Diseño del DW_DWH: Se crea la base de datos DW_DWH con contiene el datawarehouse.

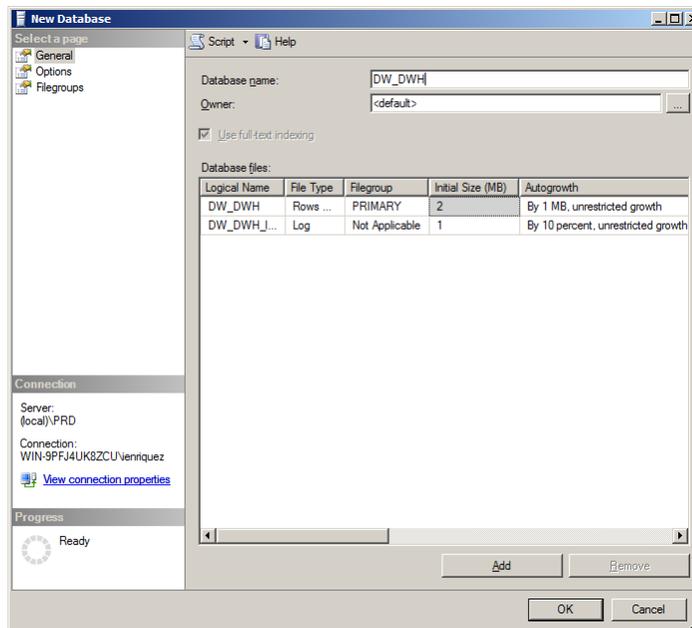


Figura N.-6.25.Diseño del DW_DWH
Elaborado por: Juan Carlos Pérez

Creación de la base resumen para los cubos de información:

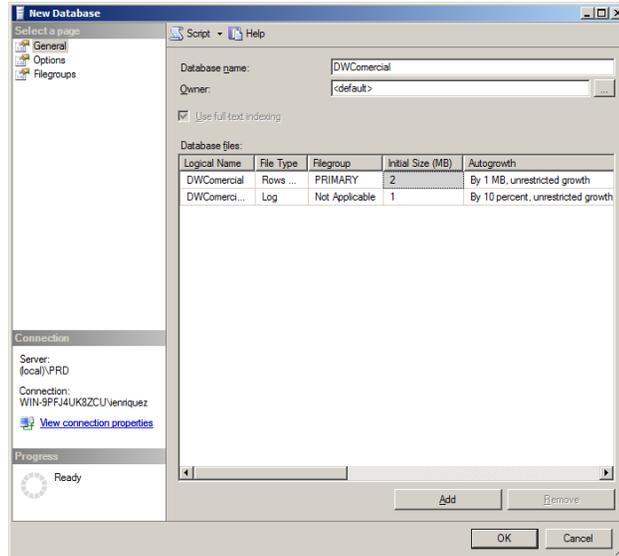


Figura N.-6.26.Creación base resumen para cubos de información
Elaborado por: Juan Carlos Pérez

Procedimientos de carga ETL: Se crean los procedimientos de extracción, transformación y carga

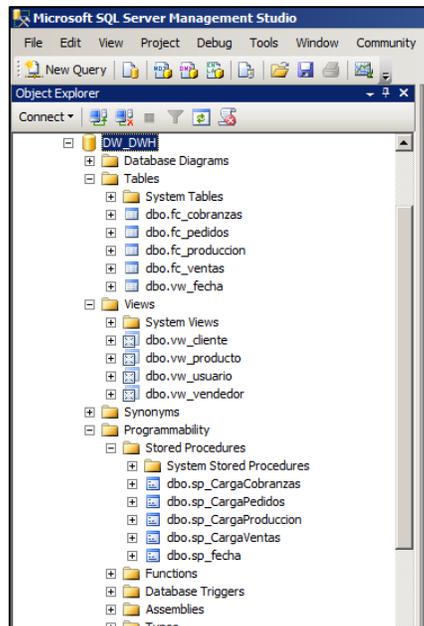
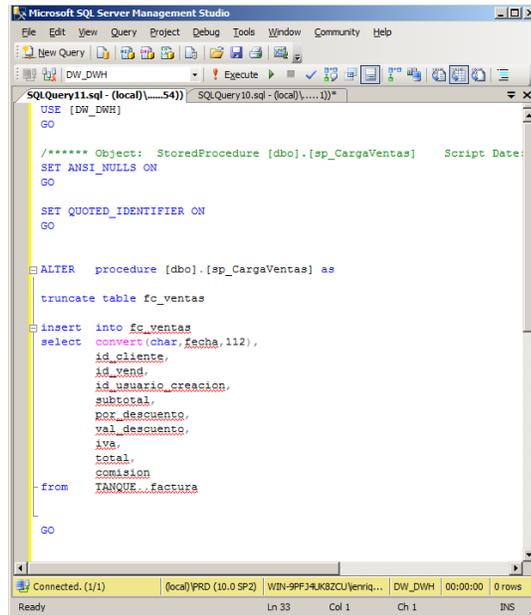


Figura N.-6.27.Procedimientos ETL creados
Elaborado por: Juan Carlos Pérez

Creación de procedimientos de carga: Se definen los procedures para que actúen de acuerdo a nuestras necesidades.



```
USE [DW_DWH]
GO

/***** Object: StoredProcedure [dbo].[sp_CargaVentas]   Script Date:
SET ANSI_NULLS ON
GO

SET QUOTED_IDENTIFIER ON
GO

ALTER procedure [dbo].[sp_CargaVentas] as

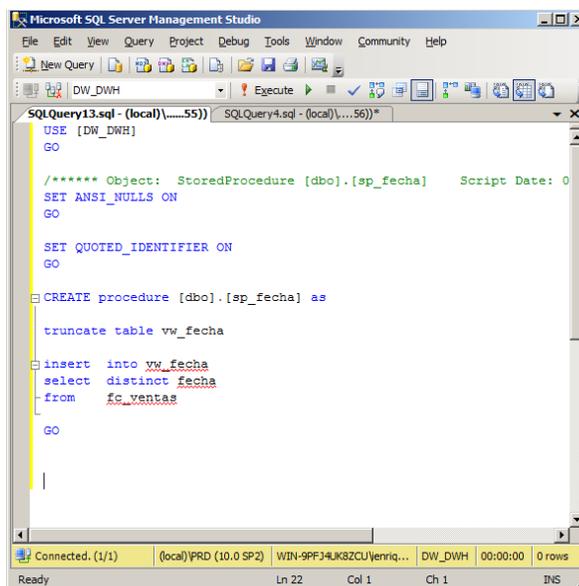
truncate table fc_ventas

insert into fc_ventas
select convert(char, fecha, 112),
id_cliente,
id_vend,
id_usuario_creacion,
subtotal,
porcentaje_descuento,
val_descuento,
iva,
social,
comision
from TANQUE...FACTURA

GO
```

Figura N.-6.28.Creación de procedimientos de carga
Elaborado por: Juan Carlos Pérez

Creación de procedimientos de carga de dimensiones:



```
USE [DW_DWH]
GO

/***** Object: StoredProcedure [dbo].[sp_fecha]   Script Date: 0
SET ANSI_NULLS ON
GO

SET QUOTED_IDENTIFIER ON
GO

CREATE procedure [dbo].[sp_fecha] as

truncate table vw_fecha

insert into vw_fecha
select distinct fecha
from fc_ventas

GO
```

Figura N.-6.29.Creación procedimientos carga de Dimensiones
Elaborado por: Juan Carlos Pérez

*Diagrama final de procesos ETL:*En la siguiente figura se muestran el diseño final con todas las herramientas que se han utilizado como son: Truncate, base de datos origen, destino y sus respectivos procedimientos.

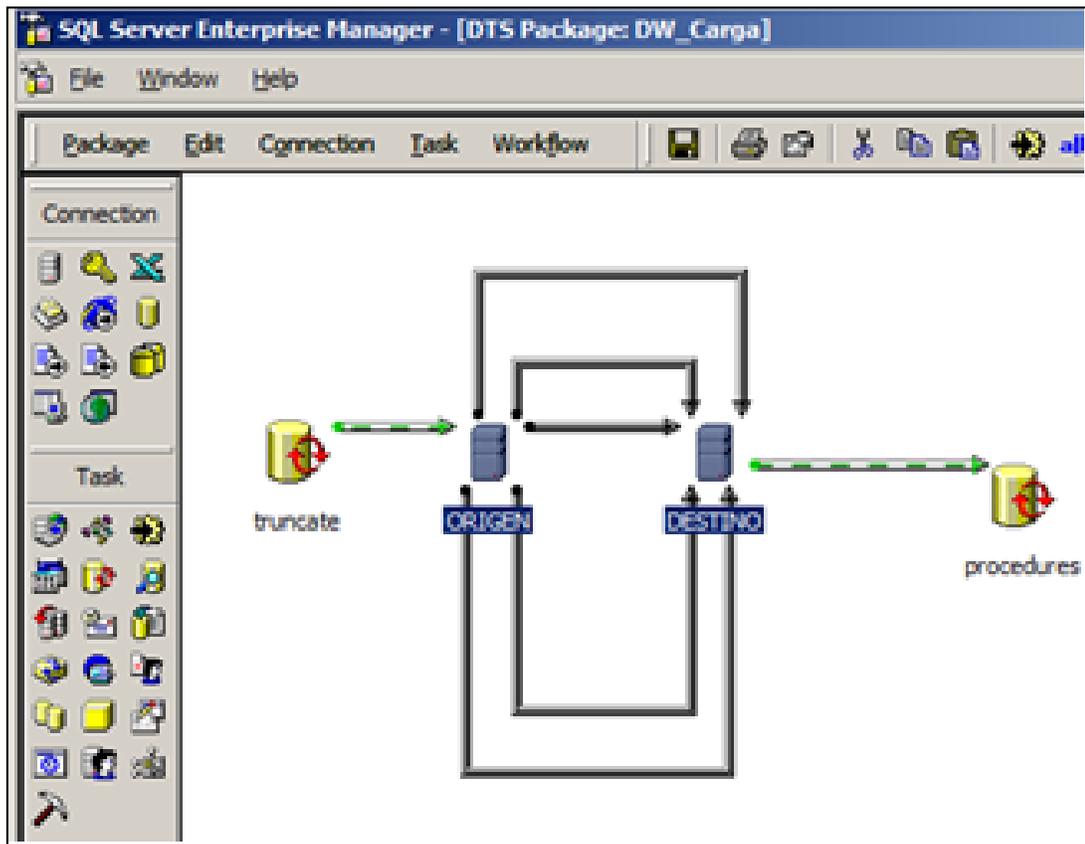
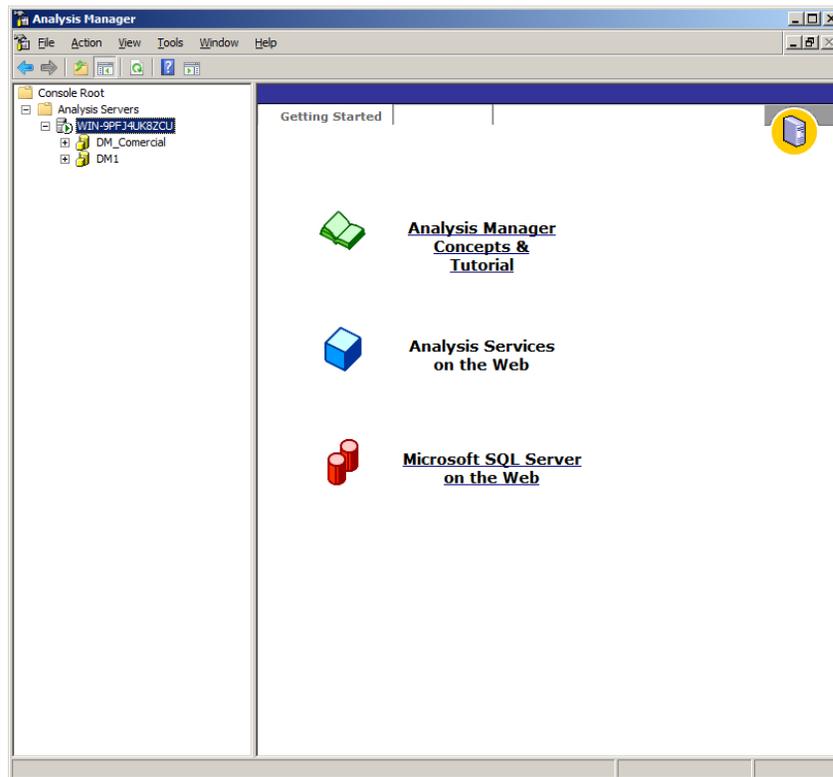


Figura N.-6.30.Diagrama de Procesos ETL
Elaborado por: Juan Carlos Pérez

Construcción

Construcción de Cubos de Información

Desarrollo del DataMart desde AnalysisServices:



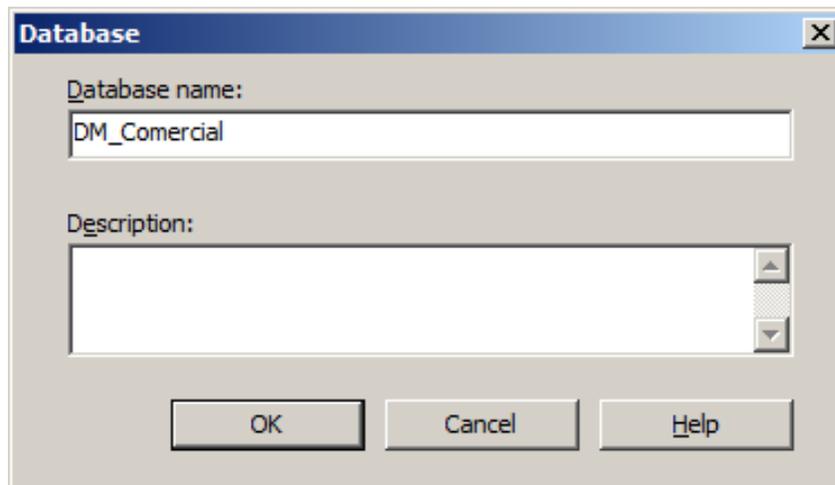


Figura N.-6.31. Desarrollo data Mart desde Analysis Services
Elaborado por: Juan Carlos Pérez

Creación de la conexión al servidor de base de datos:

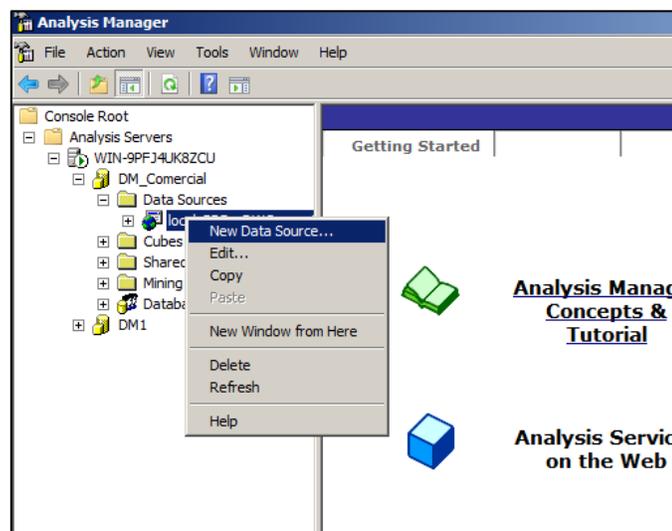


Figura N.-6.32. Conexión al servidor de base de datos
Elaborado por: Juan Carlos Pérez

Configuración enlace de datos:

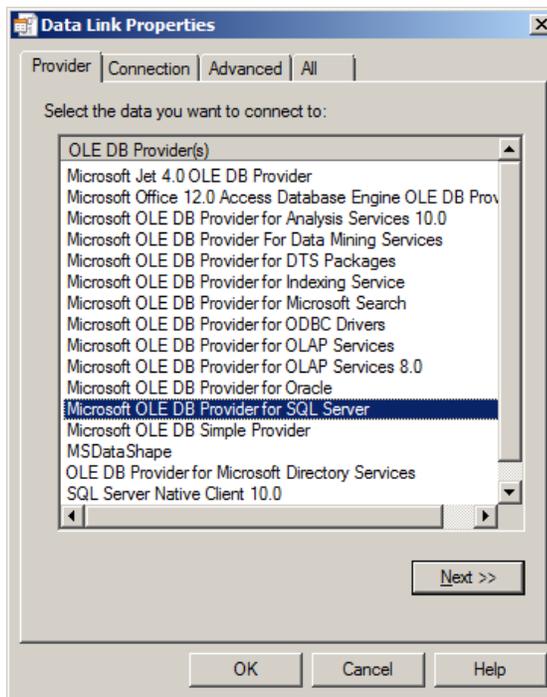


Figura N.-6.33. Configuración Enlace de Datos
Elaborado por: Juan Carlos Pérez

Se especifica la conexión a los datos desde el Servidor SQL, ingresamos el nombre del servidor, especificamos la forma de autenticación y finalmente seleccionamos el DWComercial, que es el nombre de nuestro Datawarehouse.

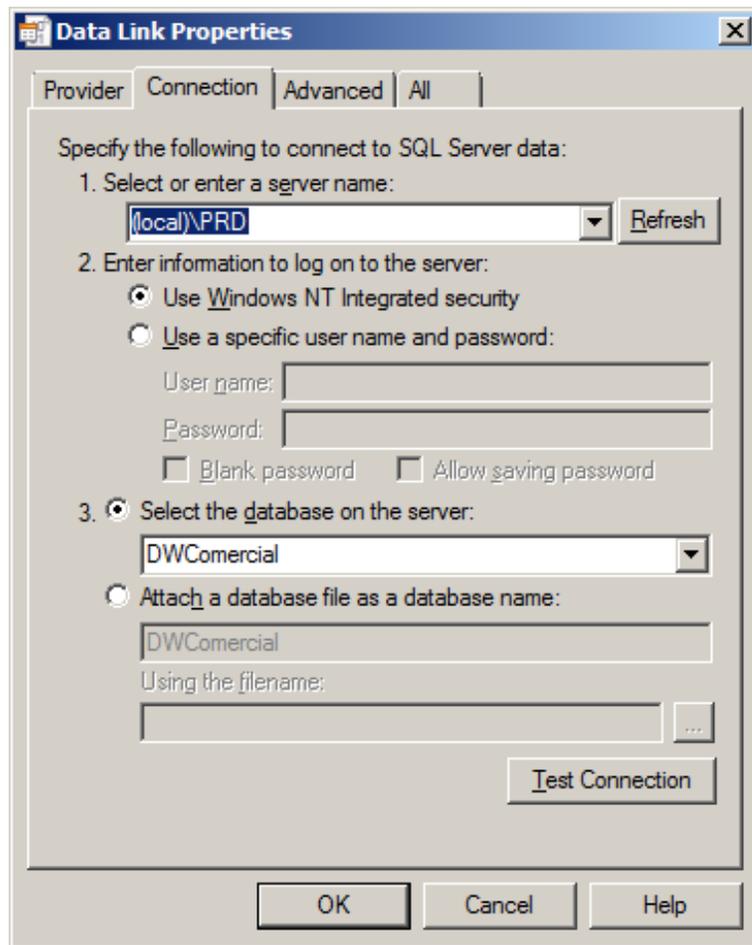


Figura N.-6.34. Prueba de conexión enlace de datos
Elaborado por: Juan Carlos Pérez

Desarrollo de la dimensión de tiempo:

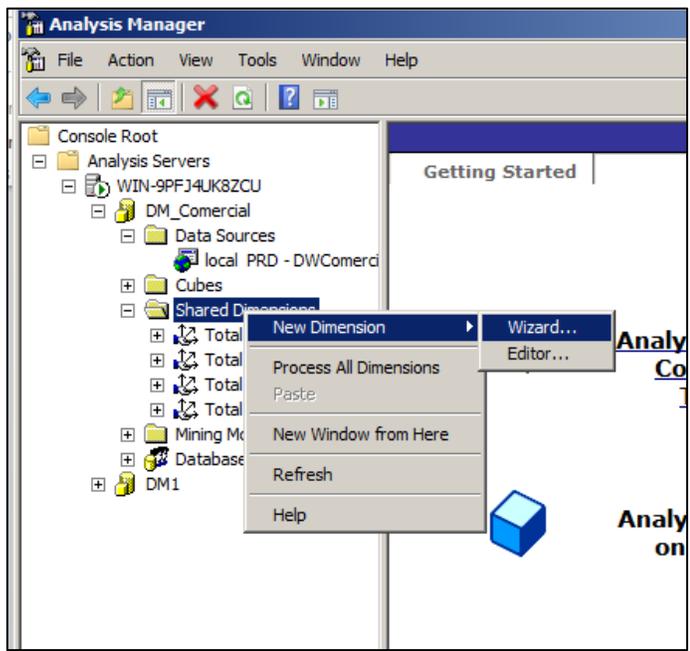


Figura N.-6.35. Desarrollo dimensión de tiempo
Elaborado por: Juan Carlos Pérez

Seleccionamos la forma como deseamos crear la dimensión, en este caso seleccionamos una dimensión jerárquica esquema estrella.

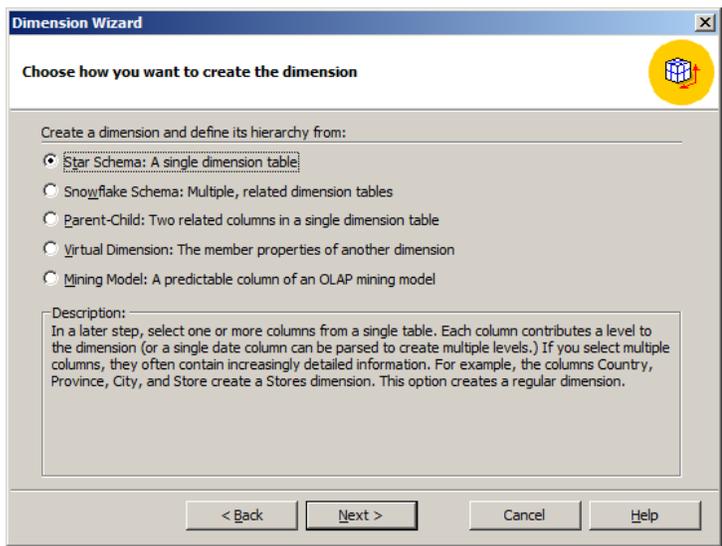


Figura N.-6.36. Creación esquema tipo estrella
Elaborado por: Juan Carlos Pérez

Selección de la tabla de dimensiones:

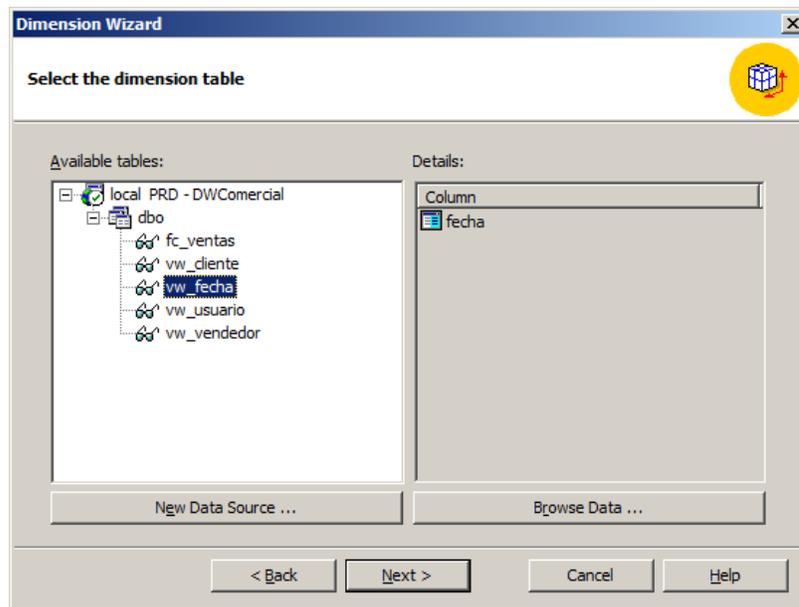


Figura N.-6.37. Selección tabla dimensión
Elaborado por: Juan Carlos Pérez

Se selecciona el tipo de dimensión en este caso el tipo de dimensión es de tiempo.

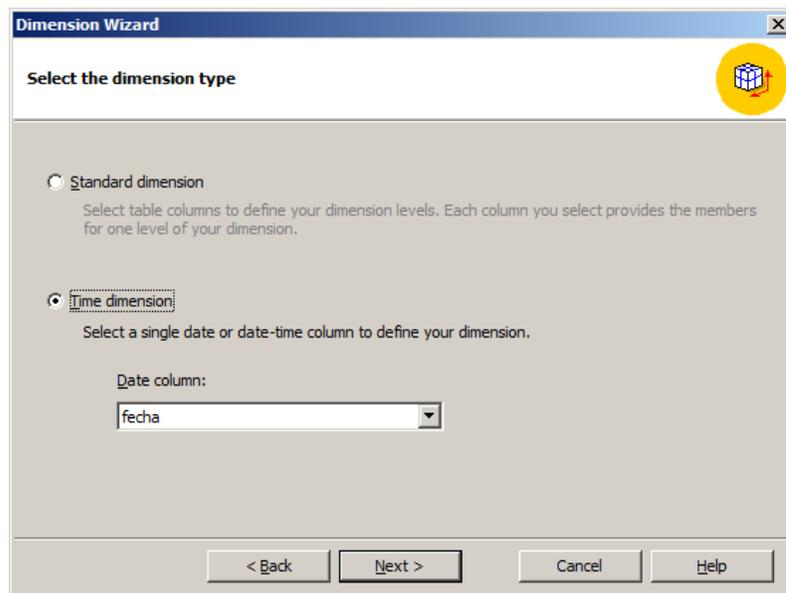


Figura N.-6.38. Selección del tipo de dimensión
Elaborado por: Juan Carlos Pérez

Creación de los niveles de tiempo para la dimensión, se selecciona el formato para los niveles de tiempo.

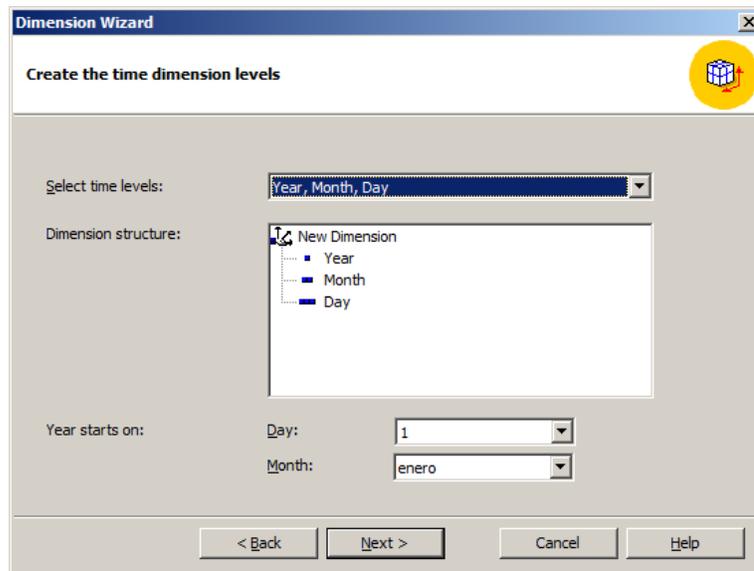


Figura N.-6.39. Creación de los niveles de tiempo para la dimensión
Elaborado por: Juan Carlos Pérez

Selección de opciones avanzadas para la creación de la dimensión, este permite añadir, eliminar y mover los miembros de la dimensión para su reproceso.

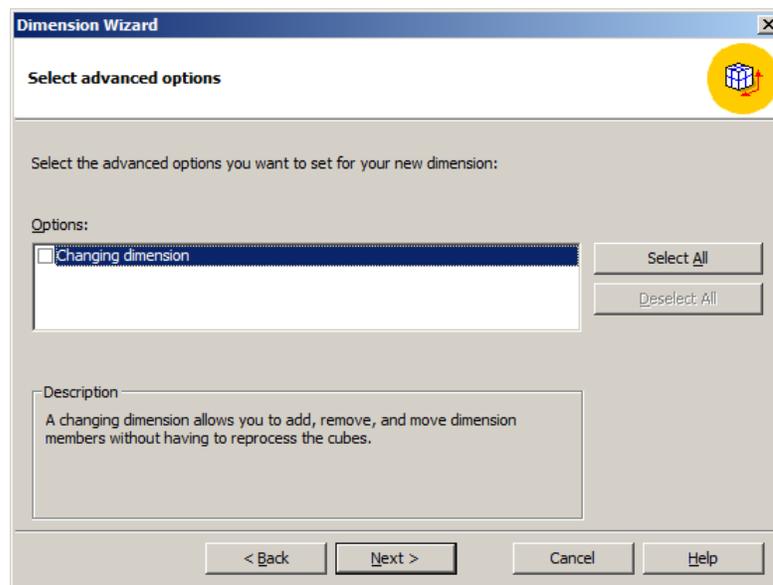


Figura N.-6.40. Opciones avanzadas de la dimensión
Elaborado por: Juan Carlos Pérez

Al finalizar el asistente para la creación de la dimensión, se obtiene como se muestra a continuación.

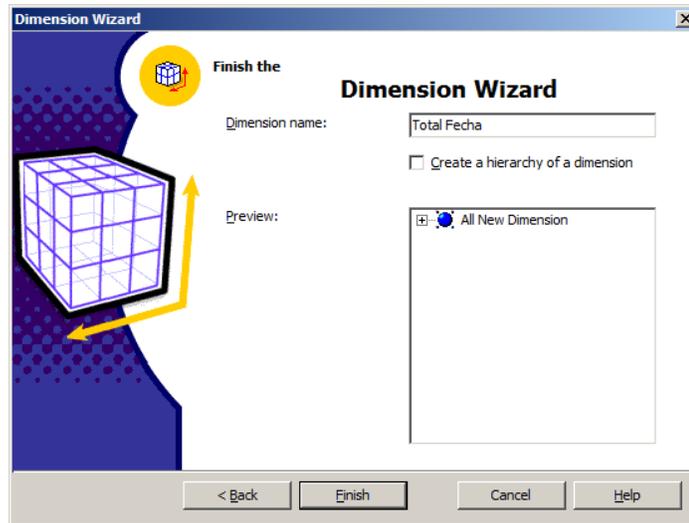


Figura N.-6.41. Fin del asistente para dimensión
Elaborado por: Juan Carlos Pérez

Creación de las dimensiones públicas:

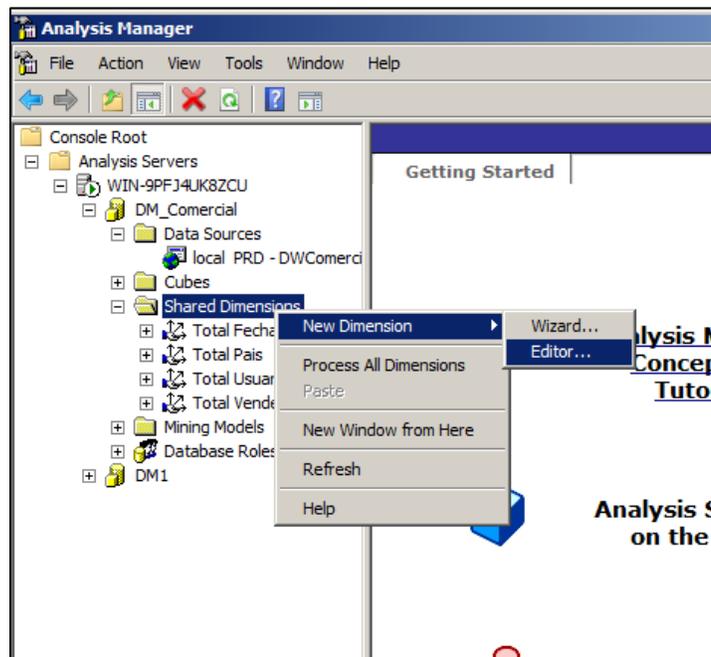


Figura N.-6.42. Creación dimensiones públicas
Elaborado por: Juan Carlos Pérez

De las dimensiones creadas se muestra los detalles de cada uno de estos.

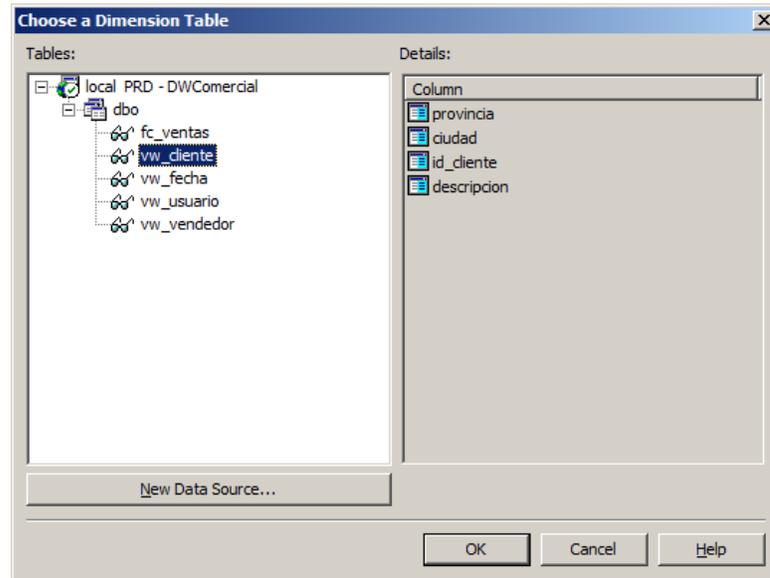


Figura N.-6.43. Dimensiones Creadas
Elaborado por: Juan Carlos Pérez

Para la edición de las dimensiones se cuenta con el editor de dimensiones

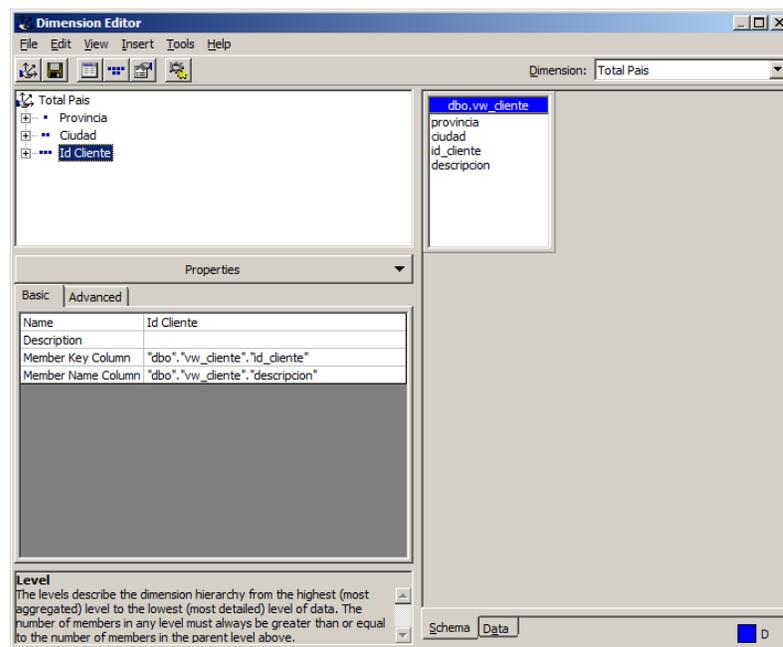


Figura N.-6.44. Editor de Dimensiones
Elaborado por: Juan Carlos Pérez

Revisión preliminar:

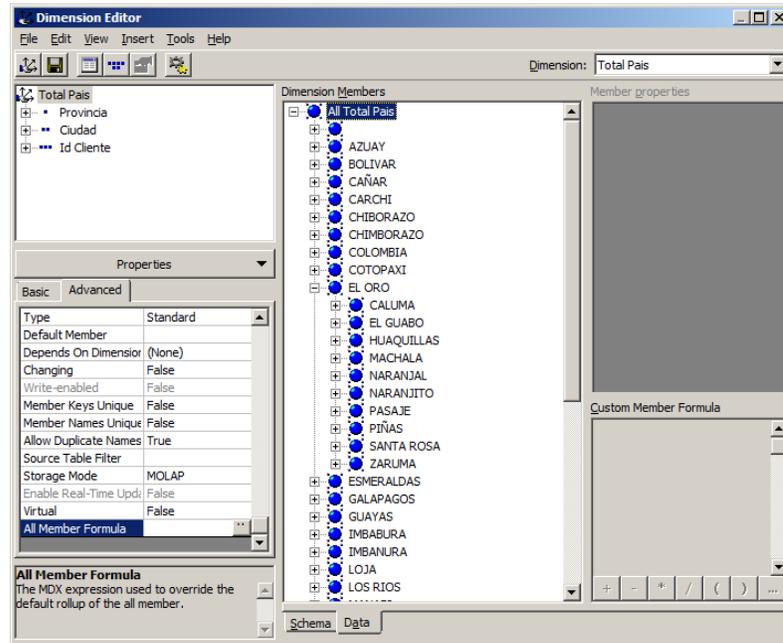


Figura N.-6.45. Revisión preliminar de la dimensión
Elaborado por: Juan Carlos Pérez

Para reconstruir una dimensión, se selecciona el método ya sea incremental o el método de reconstrucción de la estructura de la dimensión.

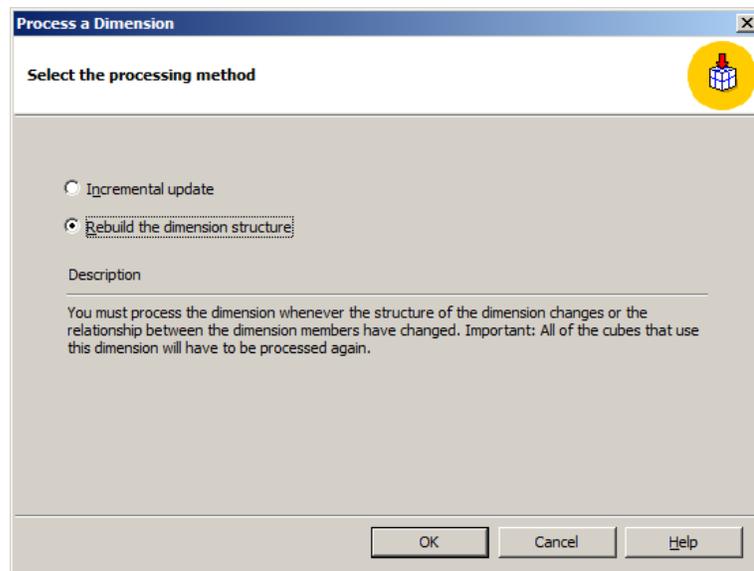


Figura N.-6.46. Reconstrucción estructura de la dimensión
Elaborado por: Juan Carlos Pérez

Creación del cubo de información:

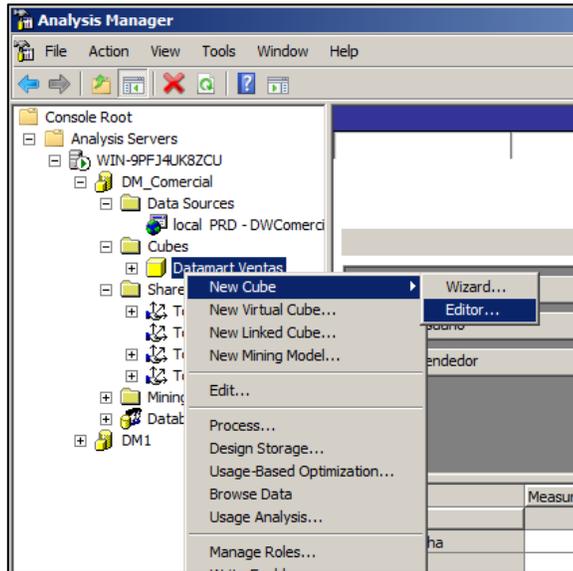


Figura N.-6.47. Creación del cubo de información
Elaborado por: Juan Carlos Pérez

Selección de la tabla de hechos.

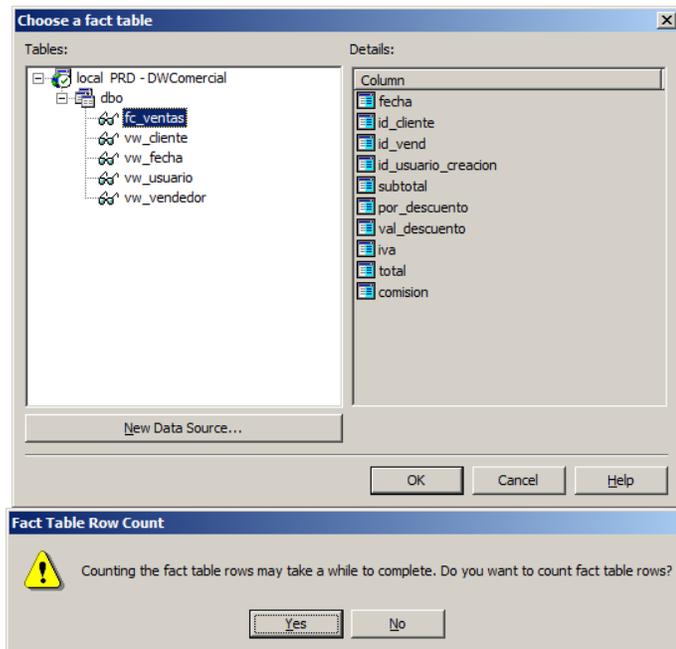


Figura N.-6.48. Selección Tabla de hechos
Elaborado por: Juan Carlos Pérez

Creación de las medidas para el cubo de información:

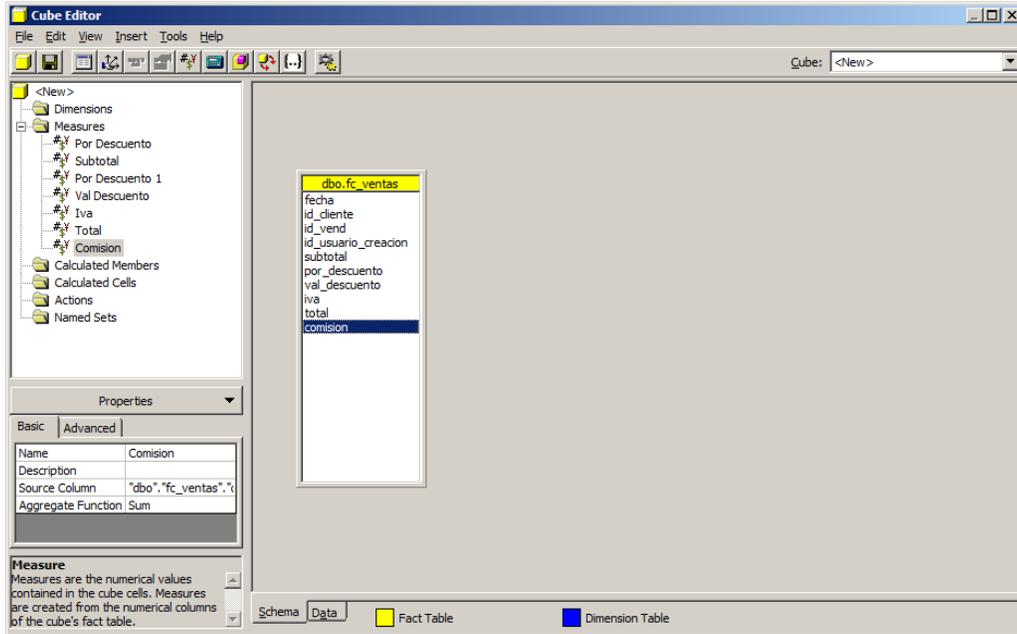


Figura N.-6.49. Creación de las medidas
Elaborado por: Juan Carlos Pérez

Especificar el agregado:

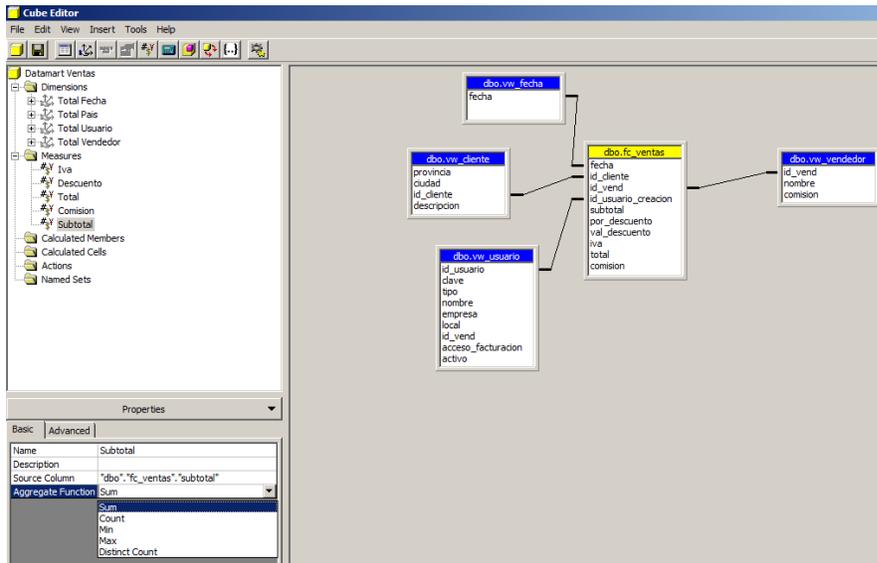


Figura N.-6.50. Especificación de agregados
Elaborado por: Juan Carlos Pérez

Especificar el tipo de dato:

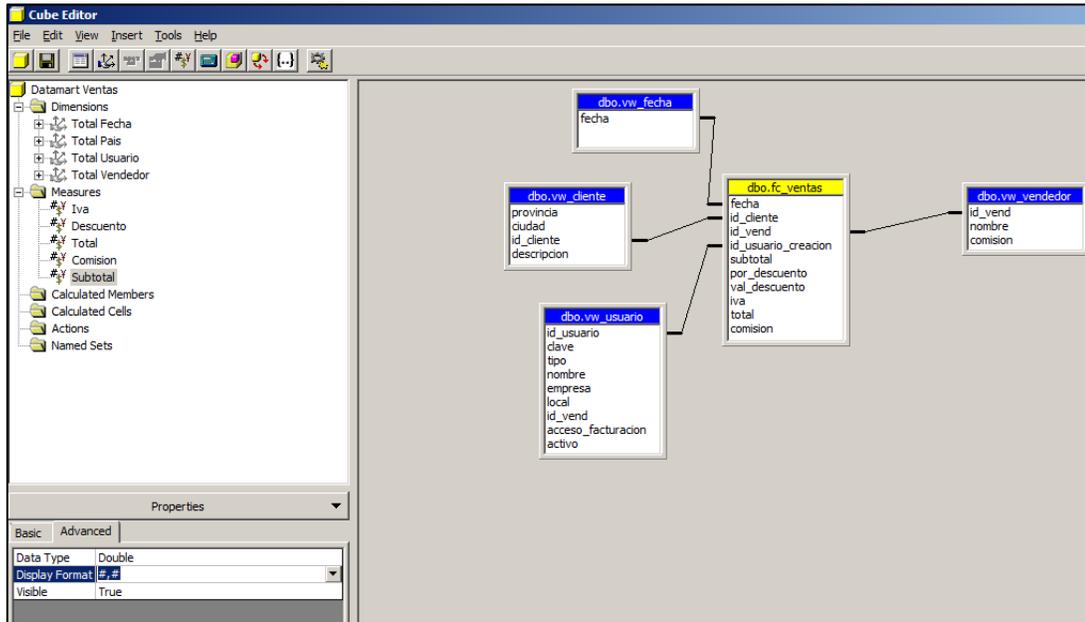


Figura N.-6.51. Especificación tipo de dato
Elaborado por: Juan Carlos Pérez

Agregando las dimensiones publicas:

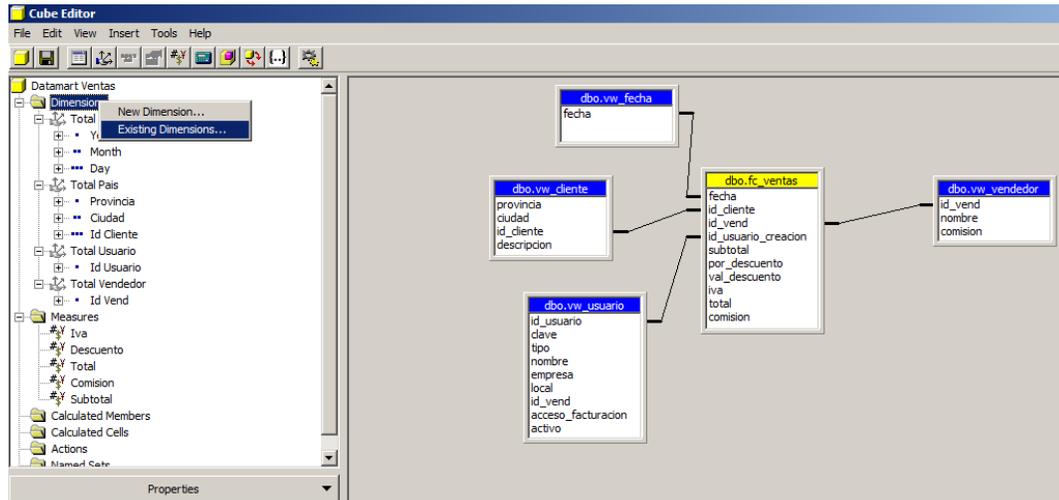


Figura N.-6.52. Agregado de dimensiones públicas
Elaborado por: Juan Carlos Pérez

Selección de las dimensiones para el cubo en el manejador de dimensiones.

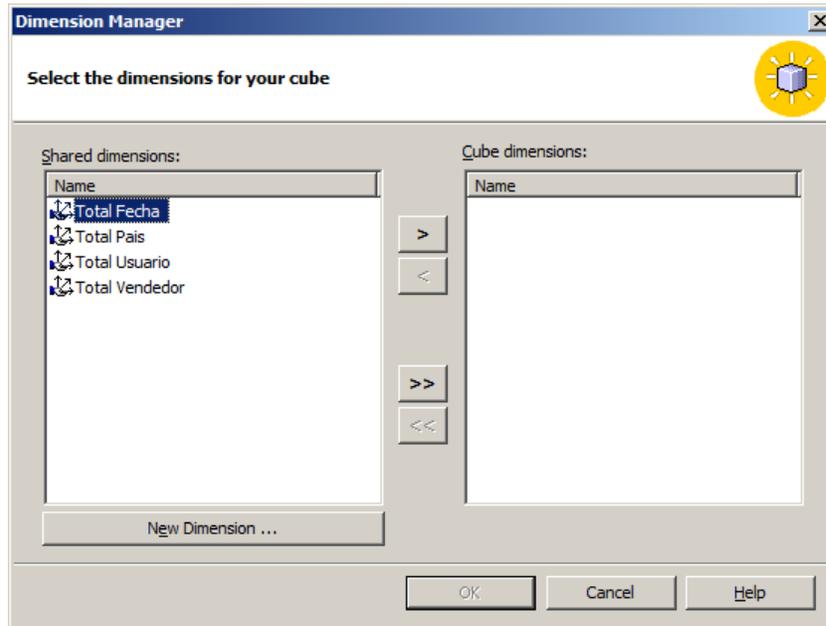


Figura N.-6.53. Selección de dimensiones para el cubo
Elaborado por: Juan Carlos Pérez

Diseño esquema estrella:

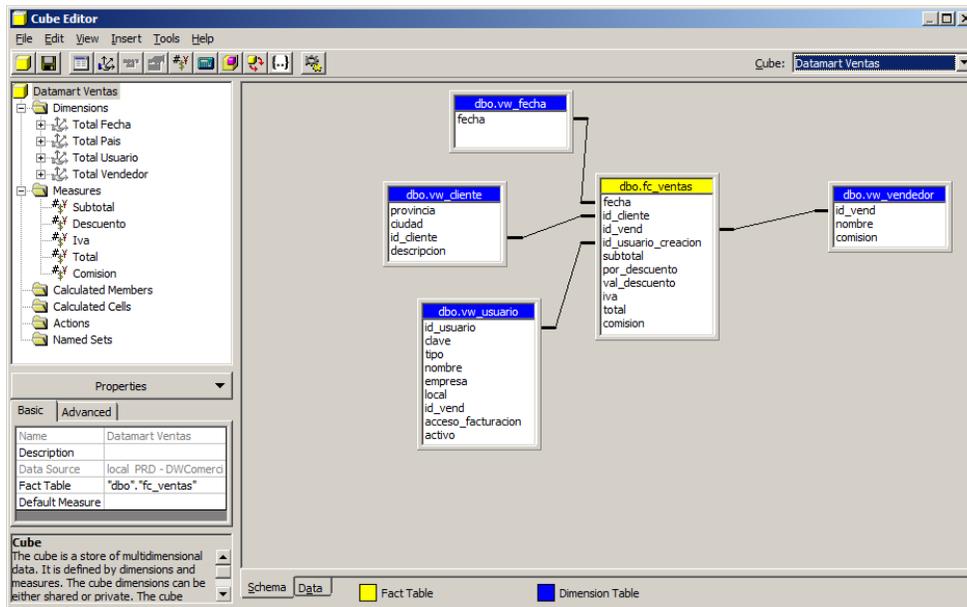


Figura N.-6.54. Diseño esquema estrella
Elaborado por: Juan Carlos Pérez

Cubo Data Mart Ventas, creado

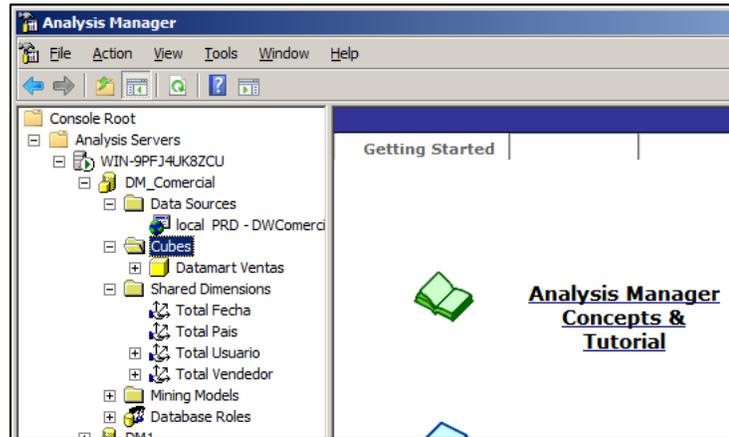


Figura N.-6.55. Cubo data Mart Ventas
Elaborado por: Juan Carlos Pérez

Optimización de consultas

Diseñar agregaciones:

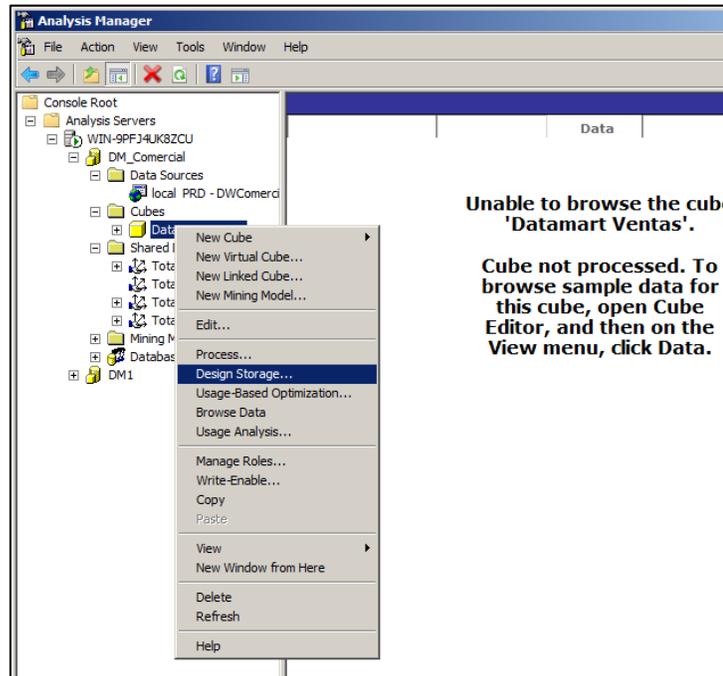


Figura N.-6.56. Diseño de agregaciones
Elaborado por: Juan Carlos Pérez

Se selecciona el tipo de almacenamiento de los datos, MOLAP, ya que esta opción almacena los datos y las agregaciones en estructuras multidimensionales.

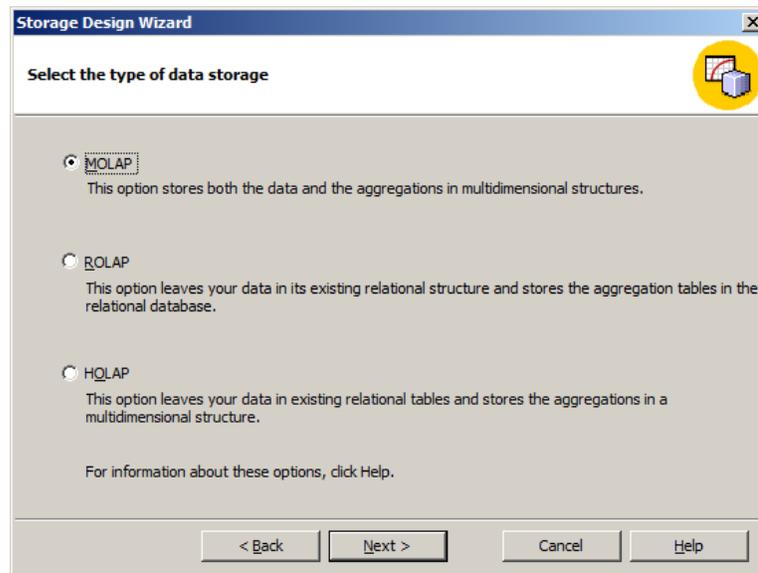


Figura N.-6.57. Selección tipo de almacenamiento de datos
Elaborado por: Juan Carlos Pérez

Se establece las opciones de la agregación, esto nos permite ganar en rendimiento.

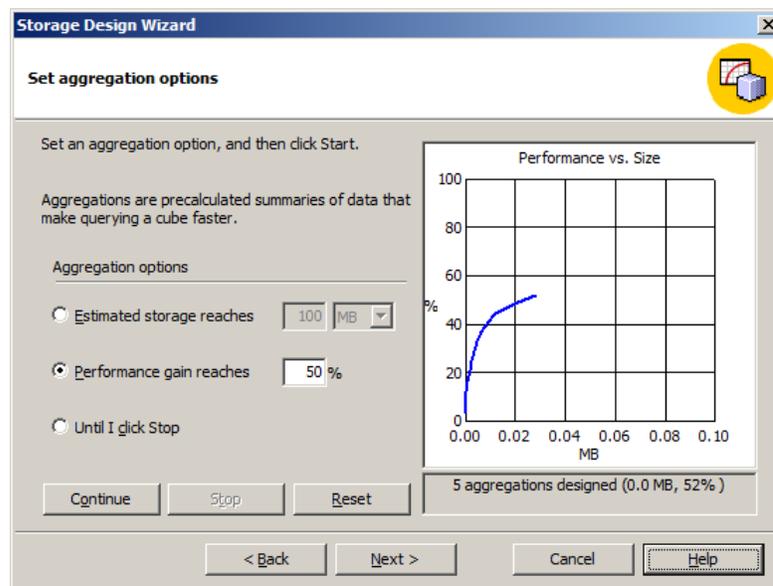


Figura N.-6.58. Configuración opciones de la agregación
Elaborado por: Juan Carlos Pérez

Procesar cubo de información:



Figura N.-6.59. Procesar cubo de información
Elaborado por: Juan Carlos Pérez

Se detalla el procesamiento de la información, y se marca la correcta u errónea finalización de cada uno de estos.

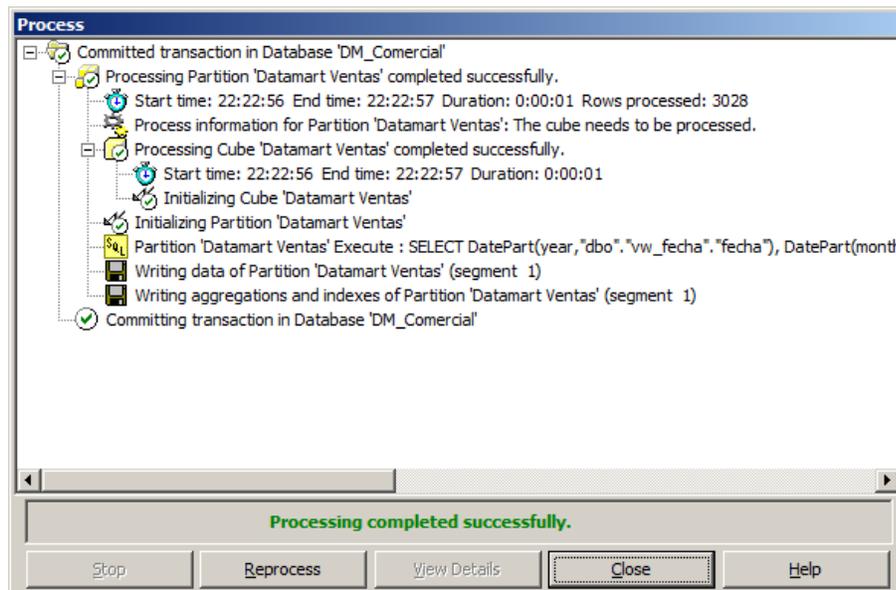


Figura N.-6.60. Resumen de procesamiento del cubo
Elaborado por: Juan Carlos Pérez

Construcción DTS Definitivo

Diagrama DTSDefinitivo: A continuación tenemos el diseño de todo el proceso para obtener los cubos de información requeridos.

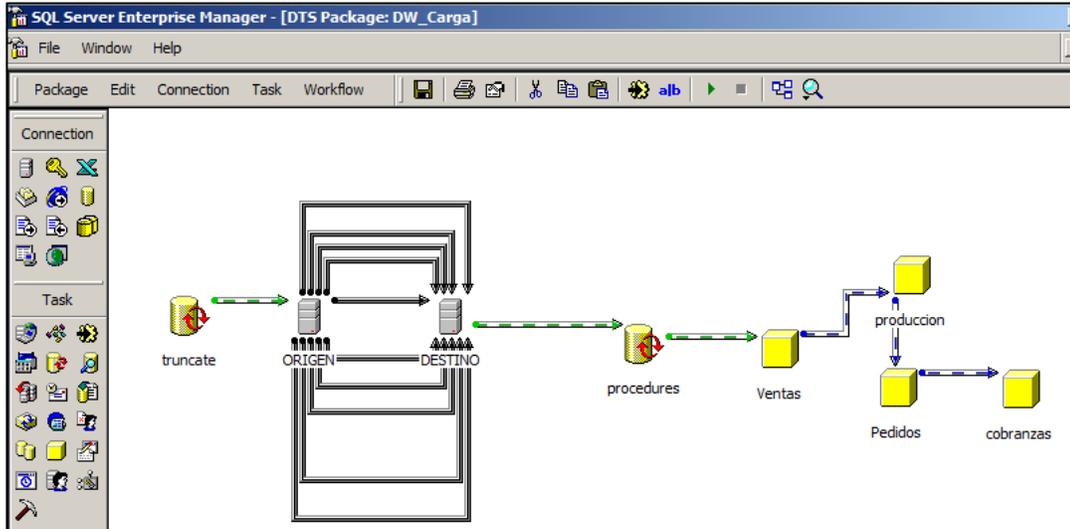


Figura N.-6.61. Diagrama DTS final
Elaborado por: Juan Carlos Pérez

Creación del job para la carga automática de datos:

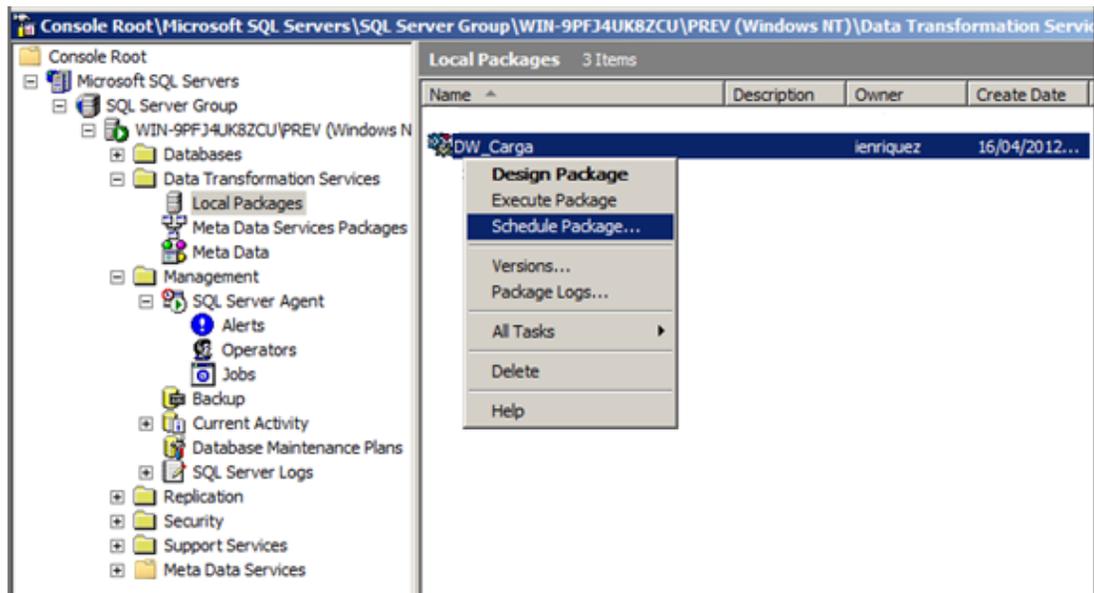


Figura N.-6.62. Creación del job para la carga
Elaborado por: Juan Carlos Pérez

Calendarización del job: Con el objetivo de que los procesos de carga se realicen de forma automática se programan trabajos automáticos calendarizados.

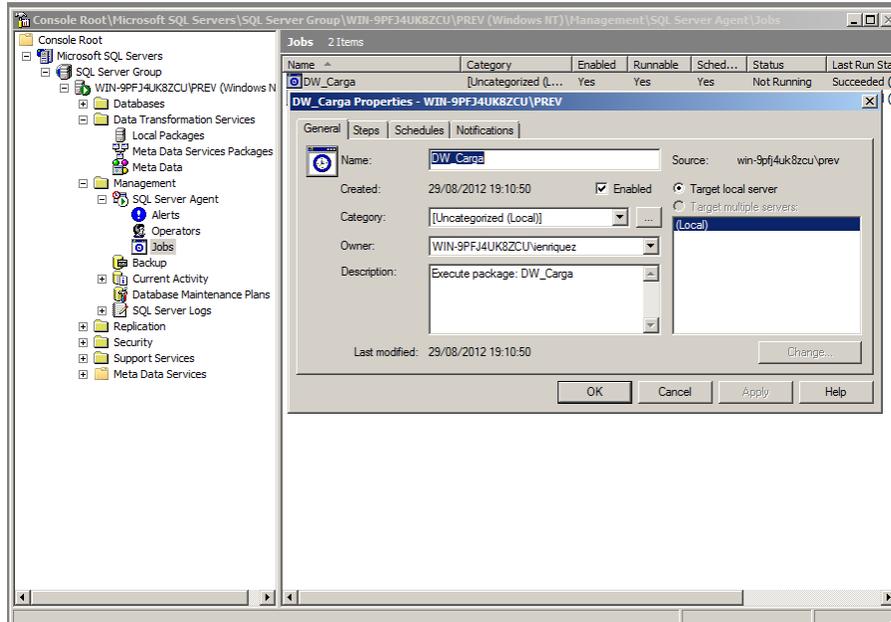


Figura N.-6.63. Calendarización del job

Elaborado por: Juan Carlos Pérez

Diseño del Reporting Services

Reporting Services: Herramienta para mostrar los datos del datawarehouse en línea.

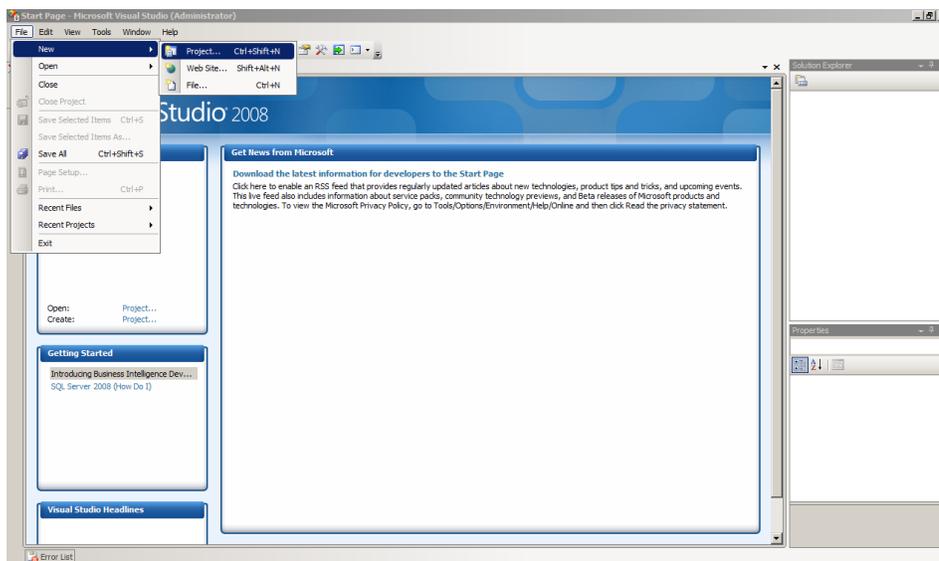


Figura N.-6.64. Reporting Services

Elaborado por: Juan Carlos Pérez

Selección de la fuente de datos desde donde se obtendrán los datos para este reporte

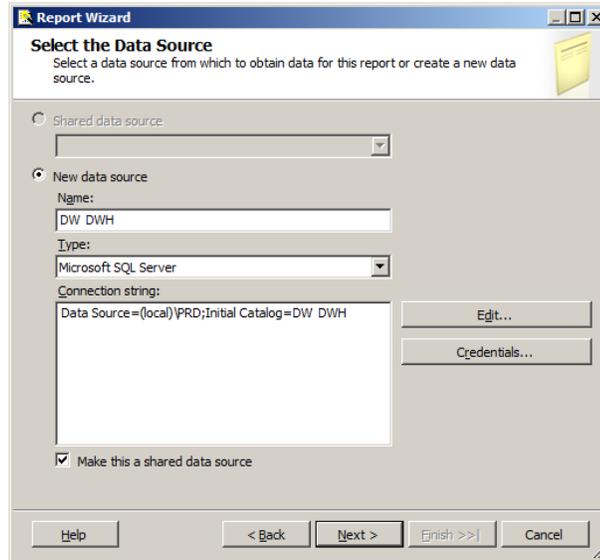


Figura N.-6.65. Selección de la Fuente de datos para el Reporting Services
Elaborado por: Juan Carlos Pérez

Se especifica la consulta a ser ejecutado para obtener los datos para el reporte.

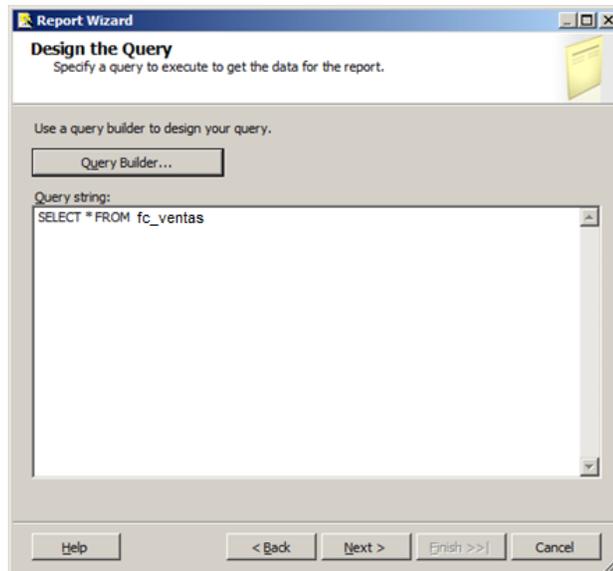


Figura N.-6.66. Query Reporting Services
Elaborado por: Juan Carlos Pérez

Se selecciona el tipo de reporte que se busca crear, en este caso se selecciona tabular.

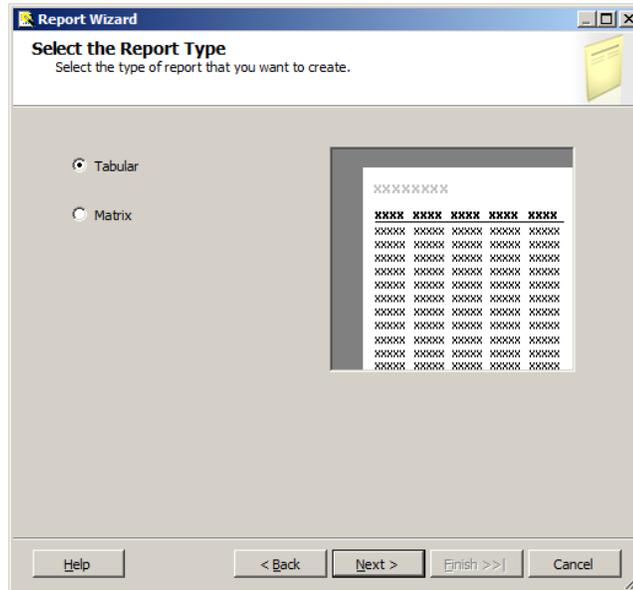


Figura N.-6.67. Tipo de Reporte Reporting Services
Elaborado por: Juan Carlos Pérez

Se selecciona como crear el grupo de datos en la tabla, como parte del diseño del reporte.

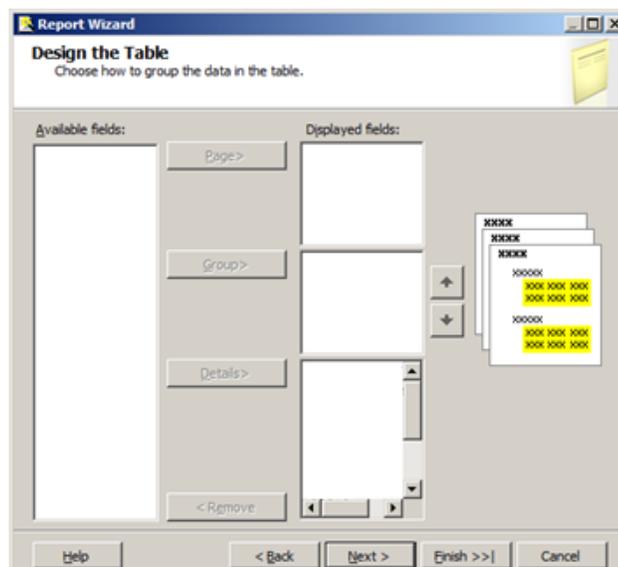


Figura N.-6.68. Diseño de tabla Reporting Services
Elaborado por: Juan Carlos Pérez

Seleccionar el estilo de la tabla, en este caso se selecciona el estilo corporativo.

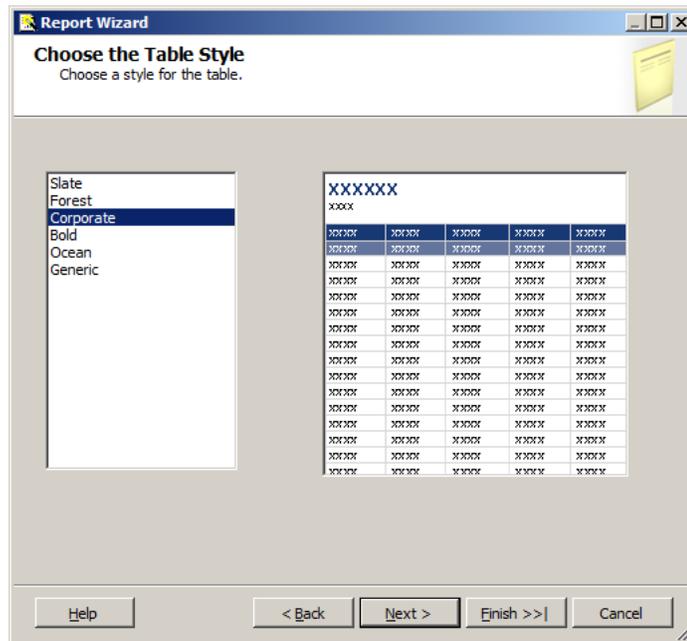


Figura N.-6.69. Selección de estilo Reporting Services
Elaborado por: Juan Carlos Pérez

Seleccionar la ubicación en donde se depositaran los reportes.

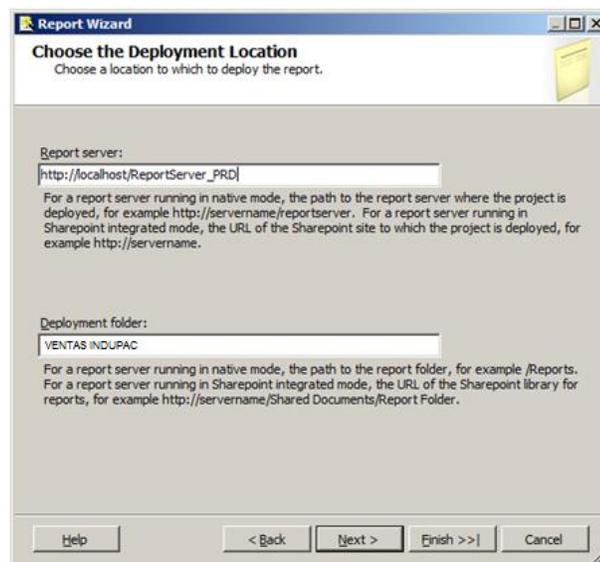


Figura N.-6.70. Ubicación para desplegar el informe Reporting Services
Elaborado por: Juan Carlos Pérez

Implementación

Configuración Consulta de Datos

Configuración de la consulta de los datos.

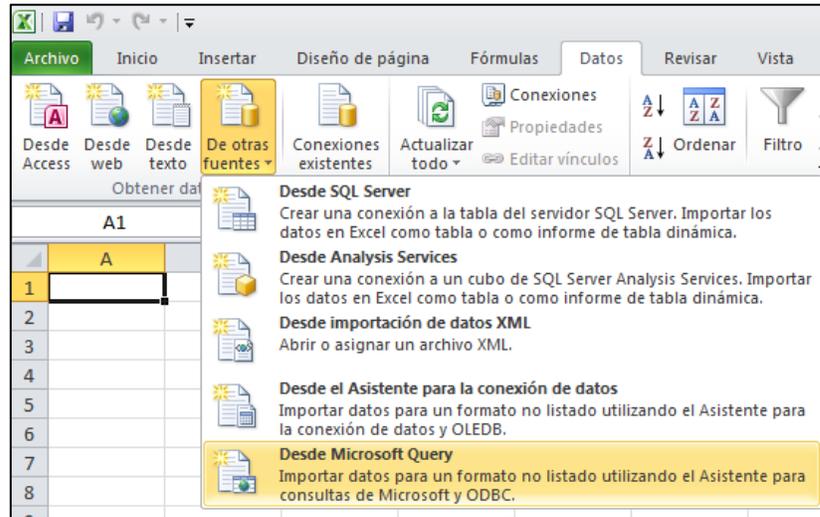


Figura N.-6.71. Configuración consulta de datos
Elaborado por: Juan Carlos Pérez

Seleccionar el origen de los datos, es decir los cubos OLAP

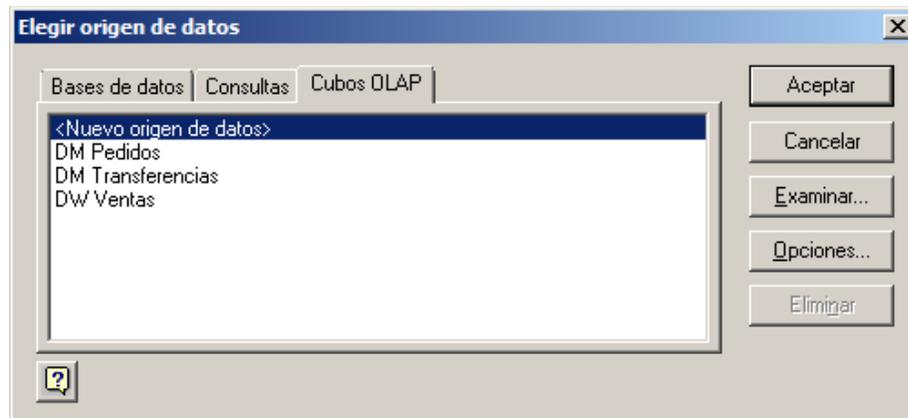


Figura N.-6.72. Elección de origen de datos
Elaborado por: Juan Carlos Pérez

Creación del nuevo origen de datos, le asignamos un nombre, seleccionamos el proveedor para la base de datos que se desea obtener acceso.

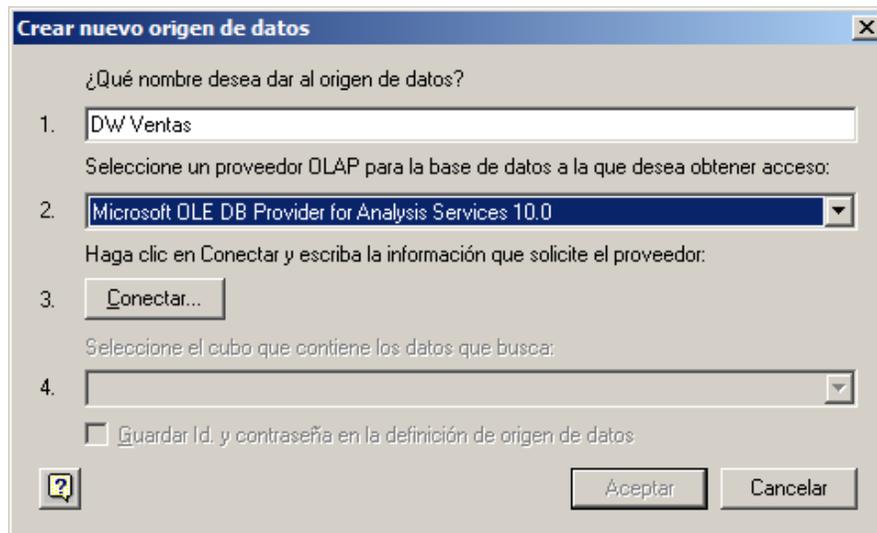


Figura N.-6.73. Creación de un nuevo origen de datos
Elaborado por: Juan Carlos Pérez

Conexión a un origen de datos multidimensional, se selecciona la ubicación del origen de datos multidimensional que se desea usar.

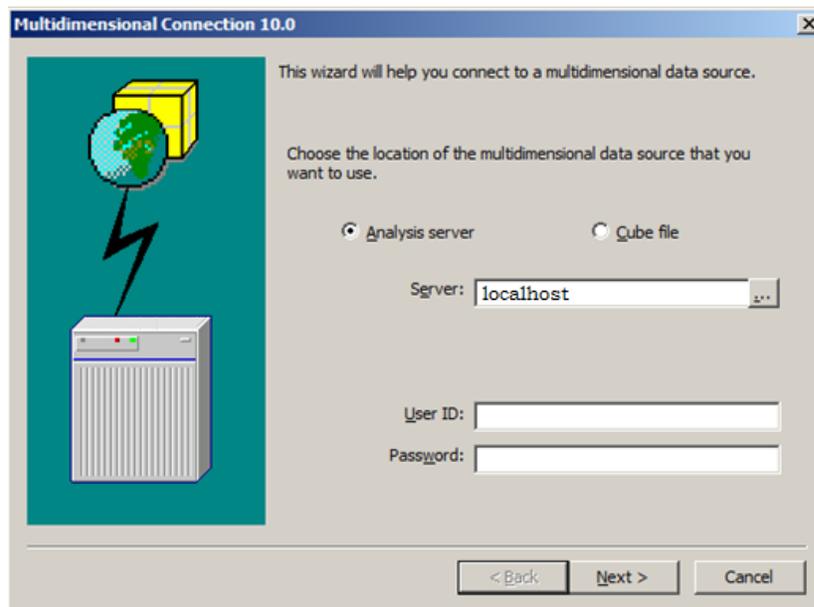


Figura N.-6.74. Conexión almacenamiento de datos multidimensional
Elaborado por: Juan Carlos Pérez

Seleccionar la base de datos Datawarehouse con la que se va a trabajar.

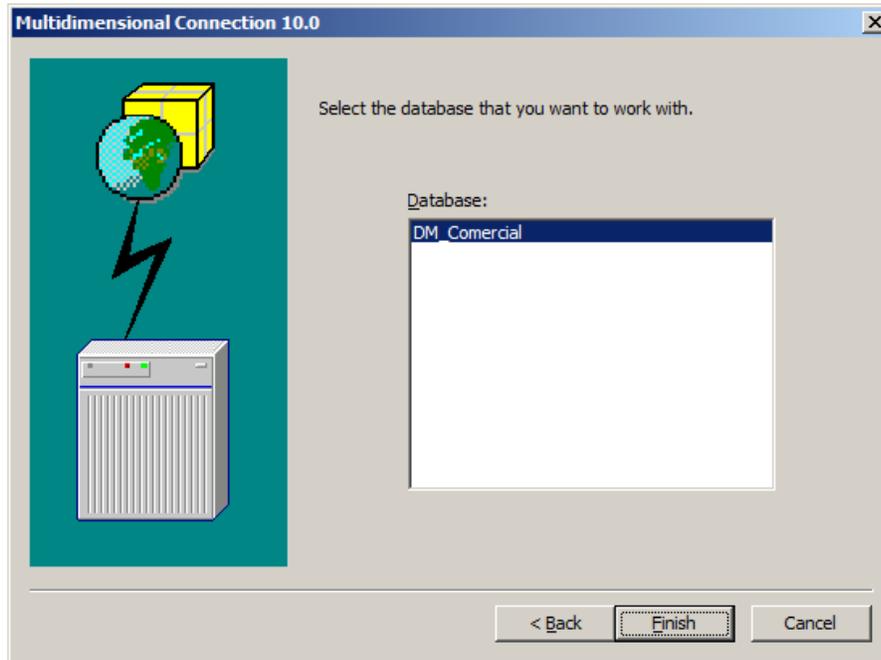


Figura N.-6.75. Conexión multidimensional
Elaborado por: Juan Carlos Pérez

Asignación un nombre al origen de datos, el tipo de proveedor para el acceso, conectar con el Datawarehouse, selección del cubo con el que se va a trabajar.

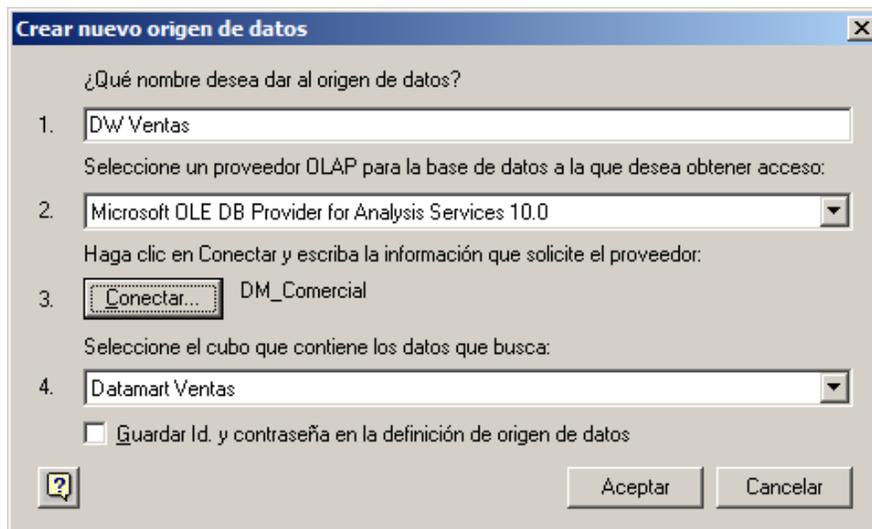


Figura N.-6.76. Origen de datos desde un cubo
Elaborado por: Juan Carlos Pérez

Se elige el cubo de información que contiene el origen de datos

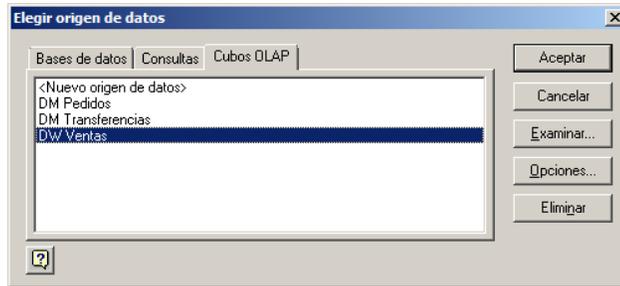


Figura N.-6.77. Elegir origen de datos multidimensional
Elaborado por: Juan Carlos Pérez

Importación de la información del cubo a Microsoft Excel



Figura N.-6.78. Importación de datos a hoja Excel
Elaborado por: Juan Carlos Pérez

Los datos se muestran en una hoja de un libro Excel seleccionado, estos se muestran como una tabla dinámica, en la cual se puede manipular la información en base a las necesidades del usuario, de forma fácil y ágil.

	2010				2011				Total	Total	Total	Total
Etiquetas de fila	Subtotal	Iva	Descuento	Comision	Total	Subtotal	Iva	Descuento	Comision	Total	Total	Total
# AZUAY	889,861	41	95,342	505,682	41	466,914	20	50,026	261,057	20	1,356,775	
# BOLIVAR	63,270	47	6,779	33,710	47	27,773	18	2,976	13,996	18	91,044	
# CAÑAR						33,783	7	3,620	13,766	7	33,783	
# CARCHI	335,983	32	4,431	22,494	32	4,830	6	518	1,938	6	340,813	
# CHIMBORAZO	338,270	192	36,243	187,320	192	183,916	61	19,705	96,356	60	522,186	
# COTOPAXI	325,683	113	34,895	170,829	113	126,549	33	13,559	43,359	33	452,232	
# EL ORO	197,326	73	21,142	116,973	73	262,255	49	28,099	93,856	49	459,582	
# ESMERALDAS	174,540	59	18,701	110,575	59	211,106	42	22,619	83,484	42	385,646	
# GUAYAS	2,935,166	317	314,482	1,976,540	317	3,027,723	348	324,399	1,224,669	348	5,962,889	
# IMBABURA	320,758	94	34,367	167,972	94	100,211	18	10,737	46,011	18	420,969	
# LOJA	280,851	55	30,091	155,696	55	137,346	19	14,716	50,075	19	418,197	
# LOS RIOS	104,855	17	11,234	71,875	17	160,954	19	17,245	64,369	19	265,809	
# MANABI	198,622	55	21,281	119,535	55	275,107	48	29,476	107,582	48	473,729	
# NAPO	23,934	15	2,564	11,314	15						23,934	
# ORELLANA	24,652	5	2,641	11,683	5						24,652	
# PASTAZA	14,169	12	1,518	6,908	12						14,169	
# PICHINCHA	872,140	345	93,444	479,585	345	553,477	143	59,301	260,592	143	1,425,618	
# SANTA ELENA						34,353	2	3,681	12,982	2	34,353	
# SANTO DOMINGO						55,983	11	5,998	15,441	11	55,983	
# TUNGURAHUA	3,500	1	375		1						3,500	
# TUNGURAHUA	5,088	4	545	2,784	4	7,146	3	766	3,990	3	12,234	
# TUNGURAGUA	7,160	16	767	3,966	16	2,541	1	272	1,512	1	9,701	
# TUNGURAHUA	680,129	511	72,871	436,911	510	210,859	155	21,805	108,879	155	890,988	
Total nacional	7,705,067	2,024	803,715	4,504,530	2,022	5,887,878	1,004	620,515	2,503,012	1,003	13,678,704	

Figura N. 6.79. Datos importados en hoja Excel
Elaborado por: Juan Carlos Pérez

Reportes

Información histórica de las cobranzas realizadas por los vendedores

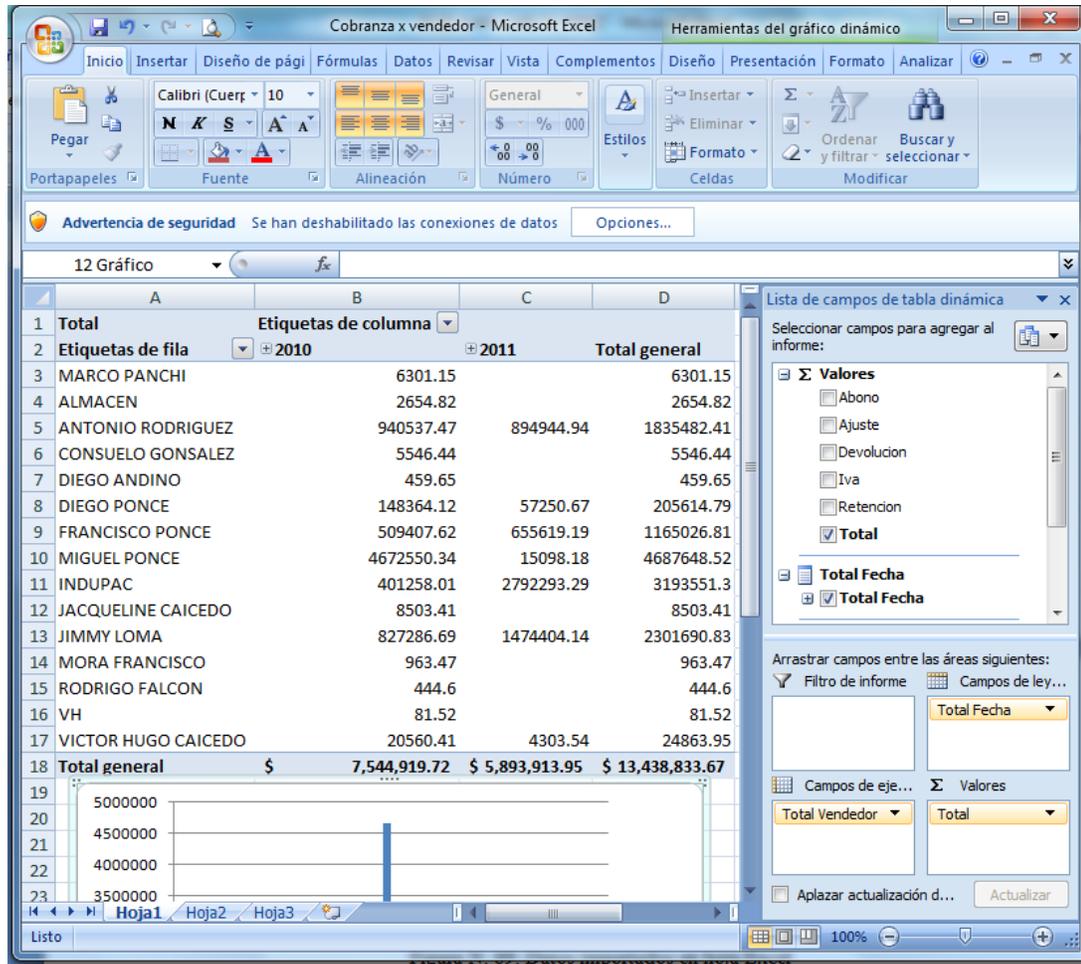


Figura N. 6.80. Información histórica de las cobranzas realizadas por los vendedores
Elaborado por: Juan Carlos Pérez

Información histórica de total recaudado en una determinada región

The screenshot shows a Microsoft Excel spreadsheet titled 'Cobranza x zona'. It features a dynamic table with the following data:

Etiquetas de fila	2010	2011	Total general
Total			
Bolivar	594.74	20594.03	21188.77
CHILLANES	594.74		594.74
GUARANDA		20594.03	20594.03
Chimborazo	2908.04		2908.04
Riobamba	2908.04		2908.04
Costa Centro	299987.14	645401.22	945388.36
CALUMA		10606.25	10606.25
GUAYAQUIL	145597.95	317011.09	462609.04
LA LIBERTAD		34352.63	34352.63
MILAGRO		71527.62	71527.62
QUEVEDO	104294.03	158051.02	262345.05
QUININDE	50095.16	38300.61	88395.77
SANTO DOMINGO		15552	15552
Costa Norte	322775.86	459760.07	782535.93
Costa Sur	2814537.67	2856275.9	5670813.57
Cotopaxi	92299.65	20491.12	112790.77
Oriente Centro	15671.66		15671.66
Oriente Norte	49326.8		49326.8
Sierra Norte	1526278.39	862670.34	2388948.73
Sierra Sur	1498004.62	826299.76	2324304.38
Tungurahua	922535.15	202421.51	1124956.66
Total general	7544919.72	5893913.95	13438833.67

Figura N. 6.81. Información histórica de total recaudado en una determinada región
Elaborado por: Juan Carlos Pérez

Información histórica de pedidos por vendedor

The screenshot shows a Microsoft Excel window with a pivot table titled 'Pedido x vendedor'. The pivot table is structured as follows:

	A	B	D
1	Total Vendedor MORA FRANCISCO		
2			
3	Cantidad	Year	
4	Id Producto	2010	Total general
5	HORARIOS ESCRIBE PEQUEÑOS	115	115
6	CALENDARIOS PARED 2011	160	200
7	MANDILES ESCRIBE	24	24
8	LIBRETAS LOOPY #1 100 HJ. CUADROS	1200	1200
9	LIBRETAS LOOPY #2 100 HJ. 1 LINEA	200	200
10	LIBRETAS LOOPY #2 100 HJ. CUADROS	600	600
11	LIBRETAS LOOPY 100 HJ. TAQUI. CUADROS	200	200
12	CUADERNOS GRAPADOS 30 HJ. DIB#5 C/M	600	600
13	CUADERNOS GRAPADOS 20 HJ. DIB#3X20	900	900
14	CUADERNOS GRAPADOS 30 HJ. MUS. TAM. AC	720	720
15	CUADERNOS GRAPADOS 100 HJ. 1 LINEA	288	288
16	CUADERNOS GRAPADOS 100 HJ. 2 LINEAS	144	144
17	CUADERNOS GRAPADOS 100 HJ. CUADROS	1728	1728
18	CUADERNOS GRAPADOS 100 HJ. 4 LINEAS	1152	1152
19	CUADERNOS GRAPADOS 20 HJ. DIBUJO	600	600
20	CUADERNOS GRAPADOS 20 HJ. 1 LINEA	1200	1200
21	CUADERNOS GRAPADOS 20 HJ. CUADROS	4800	4800
22	CUADERNOS GRAPADOS 30 HJ. 1 LINEA	1200	1200
23	CUADERNOS GRAPADOS 30 HJ. 2 LINEAS	400	400

On the right side of the Excel window, the 'Lista de campos de tabla dinámica' (Dynamic Table Fields List) is visible, showing the following fields:

- Valores** (Values):
 - Cantidad
 - Descuento
 - Despacho
 - Valor Real
 - Valor Total
 - Valor Unitario
- Total Fecha** (Total Date):
 - Total Fecha
- Total Pais** (Total Country):
 - Total Pais

Below the field list, there are sections for 'Arrastrar campos entre las áreas siguientes:' (Drag fields between the following areas:), including 'Filtro de informe' (Report Filter) with 'Total Vendedor' and 'Total Fecha', and 'Rótulos de columna' (Column Labels) with 'Total Producto' and 'Cantidad'. There are also options for 'Rótulos de fila' (Row Labels) and 'Valores' (Values) with 'Total Producto' and 'Cantidad'. An 'Actualizar' (Refresh) button is at the bottom right.

Figura N. 6.82. Información histórica de pedidos por vendedor
Elaborado por: Juan Carlos Pérez

Información histórica de pedidos por zona

The screenshot shows an Excel spreadsheet with a dynamic table. The table has columns for 'Id Producto', 'Year' (with sub-columns for 2010 and 2011), and 'Total'. The data is as follows:

Id Producto	Year		Total
	2010	2011	
HORARIOS ESCRIBE PEQUEÑOS	31786	10551	
CALENDARIOS DE PARED 2010	1848		
CALENDARIOS PARED 2011	360	3044	
MANDILES ESCRIBE	655	4	
CINTA DE EMBALAJE 48X100	180	10	
LIBRETAS LOOPY #1 100 HJ. 1 LINEA	10910	2805	
LIBRETAS LOOPY #1 100 HJ. CUADROS	28032	34009	
LIBRETAS LOOPY #2 100 HJ. 1 LINEA	6115	3004	
LIBRETAS LOOPY #2 100 HJ. CUADROS	26189	18204	
LIBRETAS LOOPY 100 HJ. TAQUI. 1 LINEA	1201	150	
LIBRETAS LOOPY 100 HJ. TAQUI. CUADROS	3807	1158	
CUADERNOS GRAPADOS 30 HJ. DIB#5 S/M	1	450	
CUADERNOS GRAPADOS 30 HJ. DIB#5 C/M	4004	1651	
CUADERNOS GRAPADOS 20 HJ. DIB#3X20	4662	601	
CUADERNOS GRAPADOS 30 HJ. MUS. TAM. AC	3307	3603	
CUADERNOS GRAPADOS 100 HJ. DIBUJO	1012	2980	
CUADERNOS GRAPADOS 100 HJ. 1 LINEA	22947	8507	
CUADERNOS GRAPADOS 100 HJ. 2 LINEAS	14142	10283	
CUADERNOS GRAPADOS 100 HJ. CUADROS	71267	44827	

The 'Total Pais' is set to 'Tungurahua'. The 'Total Fecha' is also visible in the dynamic table options.

Figura N. 6.83. Información histórica de pedidos por zona
Elaborado por: Juan Carlos Pérez

Información histórica de pedidos por producto

Cantidad	Year			
Id Producto	2009	2010	2011	Total general
HORARIOS ESCRIBE PEQUEÑOS		81077	25139	106216
CALENDARIOS DE PARED 2010		1897		1897
CALENDARIOS PARED 2011		4740	5559	10299
MANDILES ESCRIBE		1933	896	2829
CINTA DE EMBALAJE 48X100		180	10	190
ALAMBRE PLASTIFICADO 1.20 BLANCO BRASIL		2998		2998
ESFERO AZUL PUNTA MEDIO BIC			42000	42000
ESFERO NEGRO PUNTA MEDIO BIC			9600	9600
ESFERO ROJO PUNTA MEDIO BIC			18000	18000
ESFERO AZUL PUNTA FINA BIC			98400	98400
ESFERO NEGRO PUNTA FINA BIC			37200	37200
ESFERO ROJO PUNTA FINA BIC			51600	51600
BOND PERF. OFICIO CUADROS 50 HJ. C/P			34110	34110
BOND PERF. OFICIO 1 LINEA 50 HJ. C/P			5152	5152
BOND PERF. OFICIO 2 LINEAS 50 HJ. C/P			33091	33091
BOND PERF. OFICIO 4 LINEAS 50 HJ. C/P			1312	1312
LIBRETAS LOOPY #1 100 HJ. 1 LINEA		142941	58805	201746
LIBRETAS LOOPY #1 100 HJ. CUADROS		334883	175809	510692
LIBRETAS LOOPY #2 100 HJ. 1 LINEA		94530	44304	138834
LIBRETAS LOOPY #2 100 HJ. CUADROS	200	313811	162404	476415
LIBRETAS LOOPY 100 HJ. TAQUI. 1 LINEA		54775	11974	66749
LIBRETAS LOOPY 100 HJ. TAQUI. CUADROS		97475	31174	128649
CUADERNOS GRAPADOS 100 HJ. 1 LINEA TAC		360		360
CUADERNOS GRAPADOS 100 HJ. 2 LINEAS TAC		288	719	1007

Figura N. 6.84. Información histórica de pedidos por producto
Elaborado por: Juan Carlos Pérez

Información histórica de la producción

	2010	2011	Total general
Cantidad			
Etiquetas de fila			
COUCHE 250 GRAMOS 90X130	500	500	
CARTULINA DUPLEX 200 GRAMOS 62X84 CMPC CHILE	99241	43464	142705
CARTULINA DUPLEX 325 GRAMOS 62 X 85 CMPC CHILE	14386	4111	18497
CARTULINA DUPLEX 325 GRAMOS 85 X 62 CMPC CHILE	12005	7193	19198
CARTULINA DUPLEX 345 GRAMOS 90 X 64.5 CMPC CHILE	63722	20220	83942
CARTULINA RIJIDA 390 GRAMOS 84X58 CMPC.CHILE	145008	86658	231666
CARTULINA RIJIDA 390 GRAMOS 42X75 CMPC CHILE	6251	10383	16634
CARTULINA BRISTOL 140 GR. BLANCA 65X90		7500	7500
CARTULINA BRISTOL 70X100 AMARILLO	10500	1250	11750
55 CARTULINA BRISTOL 70X100 BLANCA	48350	19250	67600
CARTULINA BRISTOL 70X100 CELESTE	10000	1250	11250
CARTULINA BRISTOL 70X100 ROSADA	18750	5750	24500
CARTULINA BRISTOL 70X100 VERDE	10000	1250	11250
59 Total general	572953.98	216080	789033.98

Figura N. 6.85. Información histórica de la producción
Elaborado por: Juan Carlos Pérez

Información histórica de ventas por vendedor

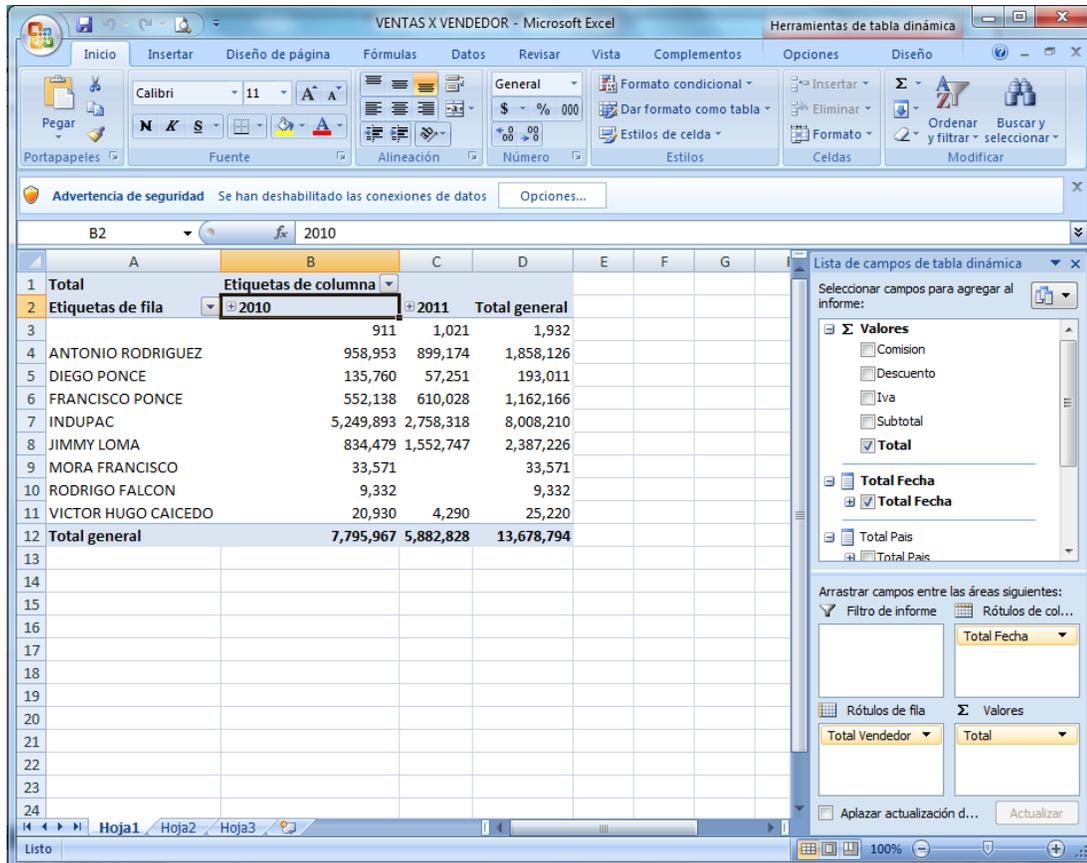


Figura N. 6.86. Información histórica de ventas por vendedor
Elaborado por: Juan Carlos Pérez

Información histórica de ventas por región

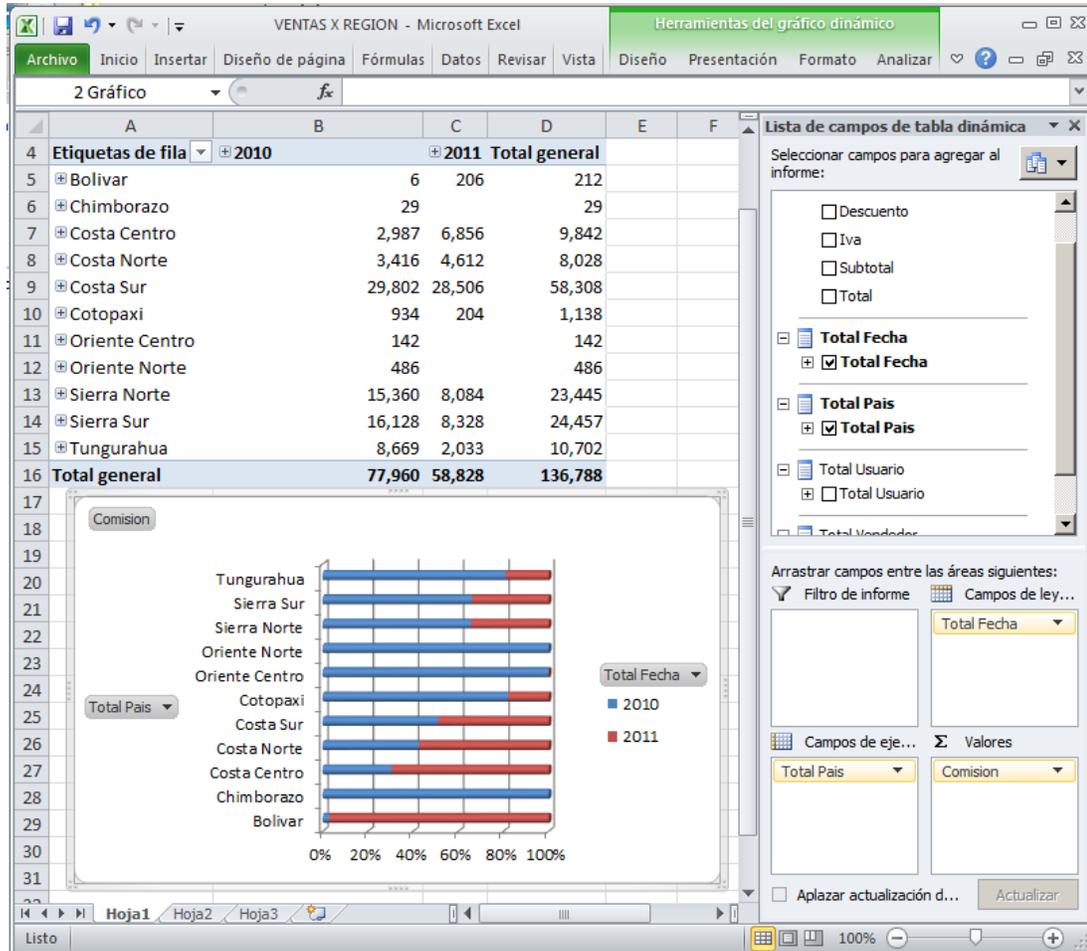


Figura N. 6.87. Información histórica de ventas por región
Elaborado por: Juan Carlos Pérez

Información histórica de las comisiones por vendedor

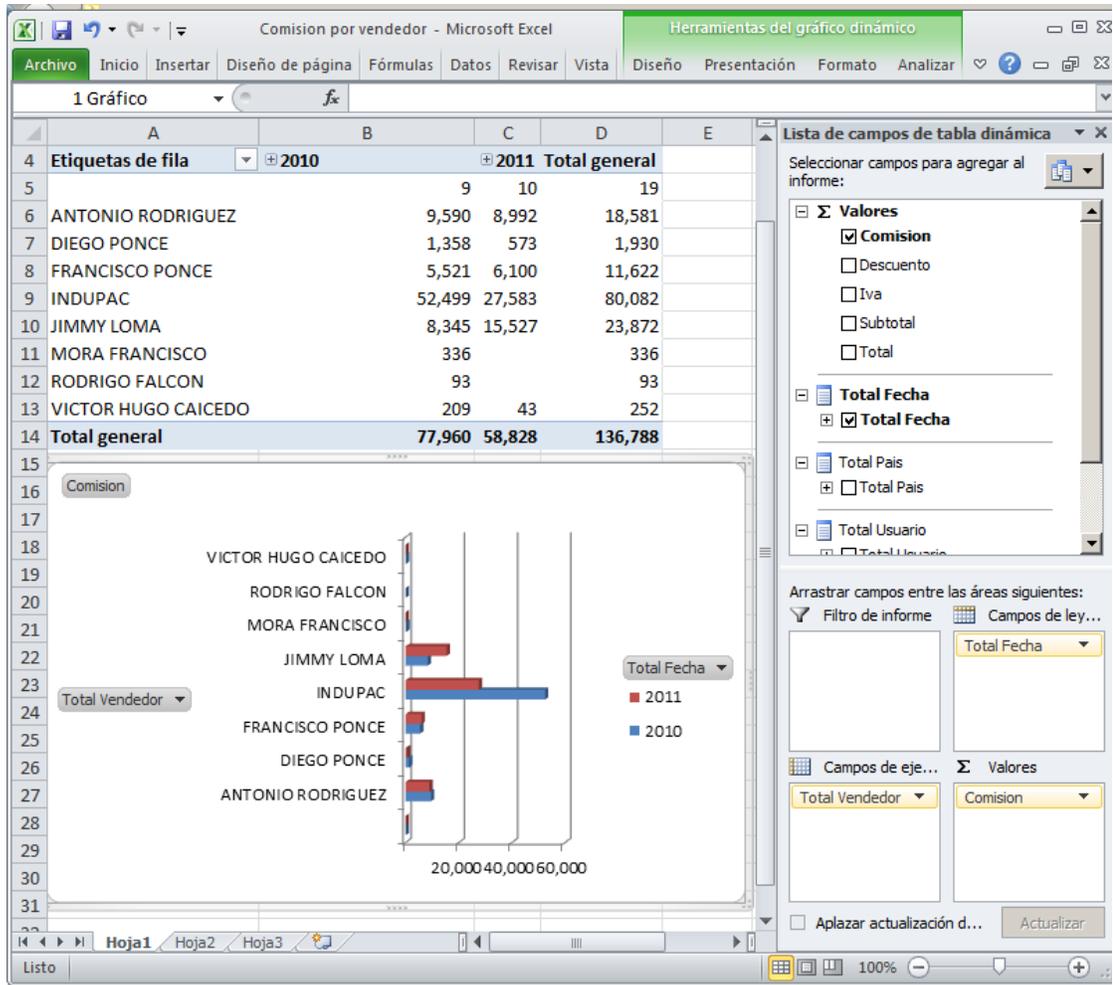


Figura N. 6.88. Información histórica de las comisiones por vendedor
Elaborado por: Juan Carlos Pérez

Reportes En Línea

Mediante el explorador podemos realizar consultas en línea de la información requerida.

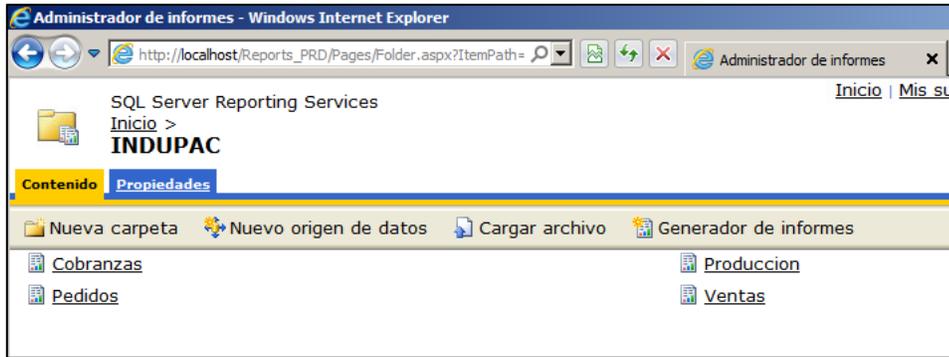


Figura N. 6.89. Consultas en línea
Elaborado por: Juan Carlos Pérez

Vista de los registros detallados de los cubos de información en línea de los pedidos de la empresa.

The screenshot shows a detailed view of the 'Pedidos' report for client '40.GUDE'. The page has a search bar and a 'Ver informe' button. Below the search bar, there is a table with the following columns: fecha, id cliente, id vend, status, comision, val descuentoe f, numeroproveedor, fn, id usuario creacion, id producto, and id. The table contains 14 rows of data, showing various orders from 1/5/2010 to 7/4/2010.

fecha	id cliente	id vend	status	comision	val descuentoe f	numeroproveedor	fn	id usuario creacion	id producto	id
1/5/2010 12:00:00 AM	40.GUDE	INDUPAC	1	1.00	0.0000	4826	F	VEGA	786108421133	UNI0
1/5/2010 12:00:00 AM	40.GUDE	INDUPAC	1	1.00	0.0000	4826	F	VEGA	786108421143	UNI9
1/5/2010 12:00:00 AM	40.GUDE	INDUPAC	1	1.00	0.0000	4826	F	VEGA	786108421331	UNI0
1/5/2010 12:00:00 AM	40.GUDE	INDUPAC	1	1.00	0.0000	4826	F	VEGA	786108421312	UNI9
1/5/2010 12:00:00 AM	40.GUDE	INDUPAC	1	1.00	0.0000	4826	F	VEGA	119210	UNI
1/26/2010 12:00:00 AM	40.GUDE	INDUPAC	1	1.00	0.0000	4845	F	PEREZ	79CBA	PL
1/26/2010 12:00:00 AM	40.GUDE	INDUPAC	1	1.00	0.0000	4845	F	PEREZ	79CBC	PL
1/26/2010 12:00:00 AM	40.GUDE	INDUPAC	1	1.00	0.0000	4845	F	PEREZ	119210	UNI
2/4/2010 12:00:00 AM	40.GUDE	INDUPAC	1	1.00	0.0000	4849	F	PEREZ	786108421133	UNI0
2/4/2010 12:00:00 AM	40.GUDE	INDUPAC	1	1.00	0.0000	4849	F	PEREZ	786108421153	UNI8
7/4/2010	40.GUDE	INDUPAC	1	1.00	0.0000	4849	F	PEREZ	786108421113	UNI

Figura N. 6.90. Consultas En línea Pedidos
Elaborado por: Juan Carlos Pérez

Consulta en línea de las cobranzas por cliente

Administrador de informes - Windows Internet Explorer
 http://localhost/Reports_PRD/Pages/Report.aspx?ItemPath: ...
 SQL Server Reporting Services
 Inicio > INDUPAC >
Cobranzas
 Ver Propiedades Historial Suscripciones
 Cliente: ALMACEN
 1 de 2 ? 100% Buscar

fecha	fec ult	tipo	numero	id cliente
6/2/2010 12:00:00 AM	6/2/2010 12:00:00 AM	R	I000107	ALMACEN
9/22/2009 12:00:00 AM		F	0010010017067	ALMACEN
9/22/2009 12:00:00 AM	4/20/2010 12:00:00 AM	F	0010010017666	ALMACEN
9/22/2009 12:00:00 AM	5/7/2010 12:00:00 AM	F	0010010017667	ALMACEN
9/22/2009 12:00:00 AM	4/20/2010 12:00:00 AM	F	0010010017696	ALMACEN
9/22/2009 12:00:00 AM	5/7/2010 12:00:00 AM	F	0010010017718	ALMACEN
9/22/2009 12:00:00 AM	5/7/2010 12:00:00 AM	F	0010010017719	ALMACEN
10/27/2009 12:00:00 AM	5/7/2010 12:00:00 AM	F	0010010017732	ALMACEN
10/27/2009 12:00:00 AM	5/7/2010 12:00:00 AM	F	0010010017751	ALMACEN
10/27/2009 12:00:00 AM	5/7/2010 12:00:00 AM	F	0010010017776	ALMACEN
12/1/2009 12:00:00 AM	5/7/2010 12:00:00 AM	F	0010010017816	ALMACEN

Figura N. 6.91. Consultas En línea Cliente
 Elaborado por: Juan Carlos Pérez

Consulta en línea de la producción

Administrador de informes - Windows Internet Explorer
 http://localhost/Reports_PRD/Pages/Report.aspx?ItemPath: ...
 SQL Server Reporting Services
 Inicio > INDUPAC >
Producción
 Ver Propiedades Historial Suscripciones
 Id: 002593
 1 de 1 100% Buscar | Siguiete | Seleccionar un formato | Exportar

id trabajo	fecha	descripción	fechafin	numhojas	gramaje	peso	bobinas	ancho
002593	6/6/2011 12:00:00 AM	PASTAS WILL GRAPADOS 43200 PLIEGOS	6/6/2011 12:00:00 AM	0.000000	0.0000	0.000000	0.0000	62.0000

Figura N. 6.92. Consultas En línea Producción
 Elaborado por: Juan Carlos Pérez

Consulta en línea de las ventas por cliente

Administrador de informes - Windows Internet Explorer

http://localhost/Reports_PRD/Pages/Report.aspx?ItemPath=...

Inicio | Mis suscripciones | Configuración del sitio | Ayuda

SQL Server Reporting Services

Inicio > INDUPAC >

Buscar: []

Ver | Propiedades | Historial | Suscripciones

Cliente: 40.GUDE [Ver informe]

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

Ventas

fecha	id cliente	id vend	id usuario creacion	subtotal	por descuento	val descuento	iva	total	comision
1/6/2010 12:00:00 AM	40.GUDE	INDUPAC	VICTOR	2973.4400	0.0000	1338.0400	196.2500	1831.6500	18.3165
4/14/2011 12:00:00 AM	40.GUDE	AR	VICTOR	12656.1600	0.0000	3923.4100	1047.9300	9780.6800	97.8068
4/26/2011 12:00:00 AM	40.GUDE	AR	VICTOR	876.7400	0.0000	271.7900	72.5900	677.5400	6.7754
9/27/2010 12:00:00 AM	40.GUDE	INDUPAC	VICTOR	7257.9600	0.0000	2758.0100	539.9900	5039.9400	50.3994
10/5/2010 12:00:00 AM	40.GUDE	INDUPAC	VICTOR	952.0800	0.0000	361.8000	70.8300	661.1100	6.6111
10/13/2010 12:00:00 AM	40.GUDE	INDUPAC	VICTOR	1741.4700	0.0000	661.7600	129.5700	1209.2800	12.0928
11/1/2010 12:00:00 AM	40.GUDE	INDUPAC	VICTOR	13688.4000	0.0000	5201.5900	1018.4200	9505.2300	95.0523
11/24/2010 12:00:00 AM	40.GUDE	INDUPAC	VICTOR	993.3400	0.0000	307.9400	82.2500	767.6500	7.6765
12/7/2010 12:00:00 AM	40.GUDE	INDUPAC	VICTOR	269.0100	0.0000	83.3900	22.2700	207.8900	2.0789
1/5/2011 12:00:00 AM	40.GUDE	INDUPAC	VICTOR	950.2800	0.0000	294.5800	78.6800	734.3800	7.3438
1/11/2011 12:00:00 AM	40.GUDE	INDUPAC	VICTOR	443.3400	0.0000	137.4400	36.7100	342.6100	3.4261
1/11/2011 12:00:00 AM	40.GUDE	INDUPAC	VICTOR	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Start | Administrador de info... | Microsoft SQL Server Ma... | SQL Server Enterprise M... | 1:32

Figura N. 6.93. Consultas En línea Ventas
Elaborado por: Juan Carlos Pérez

Plan de Capacitación a usuarios.

Un recurso muy importante definitivamente lo conforma el talento humano que acciona cada actividad del giro del negocio de una empresa, por lo que es necesario plantear un plan de capacitación para garantizar el correcto uso y explotación del datawarehouse

El siguiente plan de capacitación está desarrollado a través de un conjunto de metas, alcances, políticas, actividades y competencias.

Meta

Capacitar a los usuarios sobre la manera adecuada de obtener información del datawarehouse a través de reportes construidos en Microsoft Excel 2010.

Alcance

El plan de capacitación es aplicado a los siguientes grupos conformados por la función que desempeñan:

GRUPO N° 1: PRODUCCION	
Participante	Función
Carlos Luzuriaga	Jefe de producción
Gioconda Saltos	Coordinador
Wilson Beltrán	Coordinador
Pilar Ortega	Coordinador
Milton Saltos	Coordinador

GRUPO N° 2: VENTAS	
Participante	Función
Francisco Ponce	Jefe de ventas
Marco Villacís	Vendedor Senior
Eduardo Ponce	Vendedor Junior
Luis Ponce	Vendedor Junior

Políticas

Con el objetivo de modernizar la Planificación de la Producción y Ventas de nuestra empresa, fomentando la innovación, el desarrollo, la motivación de su capital humano, sus principios y valores, se determina la presente política de capacitación para respaldar el buen funcionamiento del sistema Datawarehouse:

- La elaboración del Plan de Capacitación es responsabilidad de la Coordinación de Sistemas, quien será el encargado de impartir el conocimiento, conjuntamente con el Jefe de RRHH, quien coordinará el horario y fecha de la capacitación, así como solicitar los recursos económicos requeridos.
- Las jornadas de capacitación pueden ser desarrolladas y organizadas directamente por Industrial Papelera Caicedo Miño S.A., según las disposiciones generales vigentes.
- Las jornadas de capacitación pueden ser de dos tipos: In-house (dentro de Industrial Papelera Caicedo Miño S.A.) o, Out-house cuyo servicio externo puede ser debidamente contratado.

- El presupuesto del recurso económico que demande el Plan de Capacitación debe ser parte de la tarea de su formulación. El estimado de la inversión en capacitación debe ser incluida en el presupuesto general de funcionamiento de la empresa, según las normas y procedimientos establecidos para tal fin.
- Los criterios que deben considerar los Jefes al momento de postular a los colaboradores a las distintas jornadas de capacitación, son: necesidad del conocimiento con relación a su puesto de trabajo, desempeño en su puesto, nivel de cooperación con su unidad de trabajo, potencial para multiplicar los conocimientos a adquirir.
- Todos los participantes de la capacitación deben completar el formulario de Informe de Participación de capacitación. El informe puede ser presentado en forma individual o por grupo, según como haya sido la participación en la capacitación. El informe debe ser remitido al Departamento de Recursos Humanos en los siguientes 5 días hábiles a la conclusión de la capacitación.

Actividades

Para alcanzar la meta propuesta, se aplicará una metodología exposición – diálogo. Los materiales a utilizar son carpeta con la exposición impresa para que puedan tomar notas u aportaciones personales los involucrados lo cual se adjunta en el Anexo 4.

Temario:

Herramienta Microsoft Excel 2010.

1. Definiciones generales
 - 1.1 Inteligencia de negocios, Datawarehouse
2. Configuración de la consulta
 - 2.1 Crear, editar conexión origen de datos
 - 2.2 Selección del Servidor y datawarehouse a trabajar
 - 2.3 Selección del cubo de información
 - 2.4 Importación de datos a Microsoft Excel
3. Obtención y manipulación de una tabla dinámica a través de cubos

OLAP

- 3.1 Edición de una tabla dinámica
- 3.2 Tipos de campos
- 3.3 Cálculos
- 3.4 Subtotales

Cronograma

Se define el cronograma de capacitación organizados por grupos:

Nombre de tarea	Duración	Comienzo	Fin	09 sep '12							16 sep '12						
				D	L	M	X	J	V	S	D	L	M	X	J	V	S
Capacitacion Producción	5 días	lun 10/09/12	vie 14/09/12														
Capacitación Ventas	5 días	lun 17/09/12	vie 21/09/12														

Competencias

Después del análisis correspondiente las personas que participen de la capacitación habrán adquirido las siguientes competencias:

- Uso solvente de los recursos de información
- Capacidad para encontrar, relacionar y estructurar información proveniente del datawarehouse y de establecer dicha información en reportes.
- Capacidad de tomar decisiones basadas en criterios objetivos
- Capacidad de planificación y organización del trabajo personal
- Capacidad para el razonamiento crítico, lógico y matemático.
- Demostrar conocimiento en el manejo de la herramienta que permite generar reportes.
- Identificar tecnologías actuales para satisfacer las necesidades de los usuarios.
- Analizar, identificar y definir los requisitos que debe cumplir un sistema informático para resolver problemas o conseguir objetivos de organizaciones y personas.

CONCLUSIONES

Como principal conclusión la herramienta datawarehouse incide significativamente en la planificación de la producción y ventas de la empresa, ya que el manejo de la información confiable ayuda a tomar las mejores decisiones a la gerencia y a los departamentos involucrados.

En conclusión el establecimiento de una herramienta datawarehouse se puede tornar difícil si no se establece una metodología de configuración de la herramienta, ya que se debe tomar en cuenta el estado de la información de la base de datos transaccional.

Se concluye que disponer de 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 los mismos.

RECOMENDACIONES

Se recomienda al desarrollador del que durante el desarrollo del datawarehouse se tome especial énfasis en la fase de análisis, con el objetivo de evitar el desperdicio de tiempo y recursos en la estructuración de los data Marts.

Es importante recomendar al desarrollador verificar el tipo de dato de los campos del datawarehouse con el tipo de datos de las tablas de la base de datos transaccional ya que estos deben del mismo tipo para evitar problemas de compatibilidad durante los procesos ETL.

Una vez implementado el datawarehouse, se recomienda a los usuarios del datawarehouse realizar un estándar gráfico y numérico, para la presentación de los reportes con el mismo formato, de forma que los directivos puedan visualizar la información de mejor manera, todo esto se logra aprovechando las herramientas ofimáticas disponibles.

Se sugiere aplicar la política de capacitación de la empresa para el personal, dando prioridad en el tema del uso de la herramienta Datawarehouse de forma que puedan adquirir experticia en la manipulación de la herramienta Excel, en el uso de tablas dinámicas, las conexiones ODBC con los Data Marts, de manera que puedan hacer uso fácil de la información contenida en el datawarehouse.

BIBLIOGRAFÍA

Libros:

Microsoft, (2000). Microsoft SQL Server 7.0 Data Warehousing Training Kit, Microsoft Press.

WREMBEL, R. (2007). Concepts, Architectures and Solutions. Canada

Kimball, R., & Margy Ross, W. T. (2008). *The Data Warehouse Lifecycle Toolkit (2nd ed.)*. Wiley.

Microsoft, (2000). SQL Server 2000 Books En línea, Microsoft Press.

Valencia, J.y Guevara, L., (2007). Data Warehousepara el Análisis Académico de la Escuela Politécnica Nacional”, Editorial EPN, Quito.

Inmon, B. (1992). *Building the Data Warehouse. 1st Edition.* . Wiley and Sons.

Abad, E., (2004). Sistema de Gestión de Inventarios y Ventas Utilizando Tecnología de Datawarehouse, Editorial de la EPN, Quito.

CAICEDO, H. (2005). Política de Capacitación Industrial Papelera Caicedo Miño, Ambato.

Enlaces Internet:

Contraloría General del Estado, N. V. (Abril de 2010). *Contraloría General del Estado, Normativa Vigente*. Recuperado el 27 de Octubre de 2012, de Contraloría General del Estado, Normativa Vigente:

http://www.contraloria.gob.ec/documentos/normatividad/constitucion_de_bolsillo.pdf

Galarraga, X. (Junio de 2000). *Datawarehouse y sus principales características*. Recuperado el 20 de Agosto de 2012, de

<http://egkafati.bligoo.com/content/view/302166/Datawarehouse-y-sus-principales-caracteristicas.html#.UIx1g2dI2vs>

García, I. E. (16 de 02 de 2010). *Organización y planificación de la producción*. Recuperado el 15 de 06 de 2012, de Producción, procesos y operaciones: <http://www.gestiopolis.com/administracion-estrategia/organizacion-planificacion-produccion.htm>

Lara Treviño, J. I. (2012). *Sistemas de soporte a la decisión: ¿qué son y por qué son necesarios?* Recuperado el 08 de Abril de 2012, de <http://www.gestiopolis.com/canales2/gerencia/1/ssdjose.htm>

Mailvaganam, H. (2007). *Introduction to OLAP* .Recuperado el 2012 de Julio de 23, de http://www.dwreview.com/OLAP/Introduction_OLAP.html

Robles, D. (09 de Octubre de 2010). *Conceptos Basicos de Business Intelligence y Datawarehouse*. Recuperado el 08 de Enero de 2011, de Conceptos Basicos de Business Intelligence y Datawarehouse:

<http://www.danterobles.com.mx/?p=310>

Real Tech, (2001), DATAWAREHOUSING, Extraído el 30 diciembre 2011 <http://www.sqlmax.com/dataw1.asp>

Valle, J. (2005), TECNOLOGIA DE BASES DE DATOS, Extraído el 15 febrero 2012

<http://www.monografias.com/trabajos24/bodega-de-datos/bodega-de-datos.shtml>

Castro, E. (2008) MODELADO DE LA DATA WAREHOUSE, Extraído el 18 de abril de 2012
<http://www.slideshare.net/ecastrom/modelado-de-data-warehouse>

MSDN Microsoft, (2012) SQL Server 2008, REPORTING SERVICES, Extraído el 05 de mayo de 2012
<http://msdn.microsoft.com/es-es/sqlserver/cc511478.aspx>

González, J. (2008), SQL Server 2008: CREANDO INFORMES CON SSRS 2008, Extraído el 20 de mayo de 2012
<http://geeks.ms/blogs/ciin/archive/2008/04/28/sql-server-2008-creando-informes-con-ssrs-2008-i.aspx>

Accelebrate, (2011), SQL SERVER 2008 REPORTING SERVICES TUTORIAL, Extraído el 25 de mayo de 2012
http://www.accelebrate.com/sql_training/ssrs_2008_tutorial.htm

Escobar, R. (2011), TUTORIAL PASAO A PASO CREAR REPORTE EN REPORTING SERVICES, Extraído el 01 de junio de 2012
<http://interdata.cl/?p=1005>

(2011, 07). Metodología Del Datawarehouse. *BuenasTareas.com*. Recuperado 07, 2011, de <http://www.buenastareas.com/ensayos/Metodologia-Del-Datawarehouse/2530059.html>

SAS Institute, I. (2001). En *Rapid Warehouse Methodology*.EEUU: SAS Institute White Paper.

ANEXOS

ANEXO 1

Ambato, 01 de Mayo de 2011

Sr.

Victor Hugo Caicedo Miño.

Gerente General de Industrial Papelera Caicedo Miño S.A.

Presente.-

De mis consideraciones:

Reciba un cordial saludo y por su digno intermedio a toda la Junta Directiva de la prestigiosa Industria a al cual Usted acertadamente dirige.

En atención al Art. 5 del reglamento Interno de Control y Administración de Bienes de Industrial Papelera Caicedo Miño S.A. que reza: “Los materiales y suministros que requieran diferentes unidades que conforman la estructura orgánica, deben ser solicitados a la Junta de Accionistas, utilizando para el efecto un oficio, el mismo que llevara la firma del Jefe de la unidad solicitante”, me permito solicitar la autorización respectiva para el uso de los equipos informáticos necesarios para la realización de mi tesis de grado con el tema: “LA TECNOLOGÍA DATAWAREHOUSE Y SU INCIDENCIA EN LA PLANIFICACIÓN DE LA PRODUCCIÓN Y VENTAS DE LA INDUSTRIA PAPELERA CAICEDO MIÑO CIA. LTDA. DE LA CIUDAD DE AMBATO.”.

Por la favorable atención a la presente, agradezco y suscribo.

Atte.,

Ing. Juan Carlos Pérez

CI: 1803465903

ANEXO 2



ASAMBLEA CONSTITUYENTE

El Estado garantizará los recursos y la infraestructura necesaria para estas actividades. Los recursos se sujetarán al control estatal, rendición de cuentas y deberán distribuirse de forma equitativa.

Art. 382.- Se reconoce la autonomía de las organizaciones deportivas y de la administración de los escenarios deportivos y demás instalaciones destinadas a la práctica del deporte, de acuerdo con la ley.

Art. 383.- Se garantiza el derecho de las personas y las colectividades al tiempo libre, la ampliación de las condiciones físicas, sociales y ambientales para su disfrute, y la promoción de actividades para el esparcimiento, descanso y desarrollo de la personalidad.

Sección séptima Comunicación social

Art. 384.- El sistema de comunicación social asegurará el ejercicio de los derechos de la comunicación, la información y la libertad de expresión, y fortalecerá la participación ciudadana.

El sistema se conformará por las instituciones y actores de carácter público, las políticas y la normativa; y los actores privados, ciudadanos y comunitarios que se integren voluntariamente a él. El Estado formulará la política pública de comunicación, con respeto irrestricto de la libertad de expresión y de los derechos de la comunicación consagrados en la Constitución y los instrumentos internacionales de derechos humanos. La ley definirá su organización, funcionamiento y las formas de participación ciudadana.

Sección octava Ciencia, tecnología, innovación y saberes ancestrales

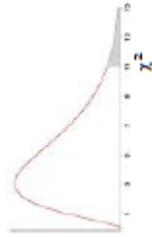
Art. 385.- El sistema nacional de ciencia, tecnología, innovación y saberes ancestrales, en el marco del respeto al ambiente, la naturaleza, la vida, las culturas y la soberanía, tendrá como finalidad:

1. Generar, adaptar y difundir conocimientos científicos y tecnológicos.
2. Recuperar, fortalecer y potenciar los saberes ancestrales.
3. 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.

ANEXO 3

Tabla D.7: VALORES CRÍTICOS DE LA DISTRIBUCIÓN JI CUADRADA

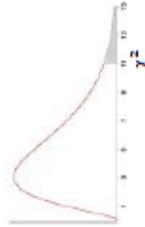
edra: Probabilidad y Estadística
 uldad Regional Mendoza



g.d.i	0,001	0,005	0,01	0,02	0,025	0,03	0,04	0,05	0,10	0,15	0,20	0,25	0,30	0,35	0,40	g.d.i
1	10,828	7,879	6,635	5,412	5,024	4,709	4,218	3,841	2,706	2,072	1,642	1,323	1,074	0,873	0,708	1
2	13,816	10,597	9,210	7,824	7,378	7,013	6,438	5,991	4,605	3,794	2,773	2,403	2,100	1,833	1,608	2
3	16,266	12,838	11,345	9,837	9,348	8,947	8,311	7,815	6,251	5,317	4,642	4,108	3,665	3,283	2,946	3
4	18,467	14,860	13,277	11,668	11,143	10,712	10,026	9,488	7,779	6,745	5,989	5,385	4,878	4,438	4,045	4
5	20,515	16,750	15,086	13,388	12,833	12,375	11,644	11,070	9,236	8,115	7,289	6,626	6,064	5,573	5,132	5
6	22,458	18,548	16,812	15,033	14,449	13,968	13,198	12,592	10,645	9,446	8,558	7,841	7,231	6,695	6,211	6
7	24,322	20,278	18,475	16,622	16,013	15,509	14,703	14,067	12,017	10,748	9,803	9,037	8,383	7,806	7,283	7
8	26,124	21,955	20,090	18,168	17,535	17,010	16,171	15,507	13,362	12,027	11,030	10,219	9,524	8,909	8,351	8
9	27,877	23,589	21,666	19,679	19,023	18,480	17,608	16,919	14,684	13,288	12,242	11,389	10,656	10,006	9,414	9
10	29,588	25,188	23,209	21,161	20,483	19,922	19,021	18,307	15,987	14,534	13,442	12,549	11,781	11,097	10,473	10
11	31,264	26,757	24,725	22,618	21,920	21,342	20,412	19,675	17,275	15,767	14,631	13,701	12,899	12,184	11,530	11
12	32,909	28,300	26,217	24,054	23,337	22,742	21,785	21,026	18,549	16,989	15,812	14,845	14,011	13,266	12,584	12
13	34,528	29,819	27,688	25,472	24,736	24,125	23,142	22,362	19,812	18,202	16,985	15,984	15,119	14,345	13,636	13
14	36,123	31,319	29,141	26,873	26,119	25,493	24,485	23,685	21,064	19,406	18,151	17,117	16,222	15,421	14,685	14
15	37,697	32,801	30,578	28,259	27,488	26,848	25,816	24,996	22,307	20,603	19,311	18,245	17,322	16,494	15,733	15
16	39,252	34,267	32,000	29,633	28,845	28,191	27,136	26,296	23,542	21,793	20,465	19,369	18,418	17,565	16,780	16
17	40,790	35,718	33,409	30,995	30,191	29,523	28,445	27,587	24,769	22,977	21,615	20,489	19,511	18,633	17,824	17
18	42,312	37,156	34,805	32,346	31,526	30,845	29,745	28,869	25,989	24,155	22,760	21,605	20,601	19,699	18,868	18
19	43,820	38,582	36,191	33,687	32,852	32,158	31,037	30,144	27,204	25,329	23,900	22,718	21,689	20,764	19,910	19
20	45,315	39,997	37,566	35,020	34,170	33,462	32,321	31,410	28,412	26,498	25,038	23,828	22,775	21,826	20,951	20
21	46,797	41,401	38,932	36,343	35,479	34,759	33,597	32,671	29,615	27,662	26,171	24,935	23,858	22,888	21,991	21
22	48,268	42,796	40,289	37,659	36,791	36,049	34,867	33,924	30,813	28,822	27,301	26,039	24,939	23,947	23,031	22
23	49,728	44,181	41,638	38,968	38,076	37,332	36,131	35,172	32,007	29,979	28,429	27,141	26,018	25,006	24,069	23
24	51,179	45,559	42,980	40,270	39,364	38,609	37,389	36,415	33,196	31,132	29,553	28,241	27,096	26,063	25,106	24
25	52,620	46,928	44,314	41,566	40,646	39,880	38,642	37,652	34,382	32,282	30,675	29,339	28,172	27,118	26,143	25
26	54,052	48,290	45,642	42,856	41,923	41,146	39,889	38,895	35,563	33,429	31,795	30,435	29,246	28,173	27,179	26
27	55,476	49,645	46,963	44,140	43,195	42,407	41,132	40,113	36,741	34,574	32,912	31,528	30,319	29,227	28,214	27
28	56,892	50,993	48,278	45,419	44,461	43,662	42,370	41,337	37,916	35,715	34,027	32,620	31,391	30,279	29,249	28
29	58,301	52,336	49,588	46,693	45,722	44,913	43,604	42,557	39,087	36,854	35,139	33,711	32,461	31,331	30,283	29
30	59,703	53,672	50,892	47,962	46,979	46,160	44,834	43,773	40,256	37,990	36,250	34,800	33,530	32,382	31,316	30
31	61,098	55,003	52,191	49,226	48,232	47,402	46,059	44,985	41,422	39,124	37,359	35,887	34,598	33,431	32,349	31
32	62,487	56,328	53,486	50,487	49,480	48,641	47,282	46,194	42,585	40,256	38,466	36,973	35,665	34,480	33,381	32
33	63,870	57,648	54,776	51,743	50,725	49,876	48,500	47,400	43,745	41,386	39,572	38,058	36,731	35,529	34,413	33
34	65,247	58,964	56,061	52,995	51,966	51,107	49,716	48,602	44,903	42,514	40,676	39,141	37,795	36,576	35,444	34
35	66,619	60,275	57,342	54,244	53,203	52,335	50,928	49,802	46,059	43,640	41,778	40,223	38,859	37,623	36,475	35
40	73,402	66,766	63,691	60,436	59,342	58,428	56,946	55,758	51,805	49,244	47,269	45,616	44,165	42,848	41,622	40
60	99,607	91,952	89,379	84,580	83,298	82,225	80,482	79,082	74,397	71,341	68,972	66,981	65,227	63,628	62,135	60
80	124,839	116,321	114,329	108,069	106,629	105,422	103,459	101,879	96,578	93,106	90,405	88,130	86,120	84,284	82,566	80
90	137,208	128,299	124,116	119,648	118,136	116,869	114,806	113,145	107,565	103,904	101,054	98,650	96,524	94,581	92,761	90
100	149,449	140,169	135,807	131,142	129,561	128,237	126,079	124,342	118,498	114,659	111,667	109,141	106,906	104,862	102,946	100
120	173,617	163,648	158,950	153,918	152,211	150,780	148,447	146,567	140,233	136,062	132,806	130,055	127,616	125,383	123,289	120
140	197,451	186,847	181,840	176,471	174,648	173,118	170,624	168,613	161,827	157,352	153,854	150,894	148,269	145,863	143,604	140

Distribución ji cuadrada - Pág. 1

Tabla D.7: VALORES CRÍTICOS DE LA DISTRIBUCIÓN JI CUADRADA



g.d.l	0,45	0,50	0,55	0,60	0,65	0,70	0,75	0,80	0,85	0,90	0,95	0,975	0,98	0,99	0,995	g.d.l
1	0,571	0,455	0,357	0,275	0,206	0,148	0,102	0,064	0,036	0,016	0,004	0,001	0,001	0,000	0,000	1
2	1,597	1,386	1,196	1,022	0,862	0,713	0,575	0,446	0,325	0,211	0,103	0,051	0,040	0,020	0,010	2
3	2,643	2,366	2,109	1,869	1,642	1,424	1,213	1,005	0,798	0,584	0,352	0,216	0,185	0,115	0,072	3
4	3,687	3,357	3,047	2,753	2,470	2,195	1,923	1,649	1,366	1,064	0,711	0,484	0,429	0,297	0,207	4
5	4,728	4,351	3,996	3,655	3,325	3,000	2,675	2,343	1,994	1,610	1,145	0,831	0,752	0,554	0,412	5
6	5,765	5,348	4,952	4,570	4,197	3,828	3,455	3,070	2,661	2,204	1,635	1,237	1,134	0,872	0,676	6
7	6,800	6,346	5,913	5,493	5,082	4,671	4,255	3,822	3,358	2,833	2,167	1,690	1,564	1,239	0,989	7
8	7,833	7,344	6,877	6,423	5,975	5,527	5,071	4,594	4,078	3,490	2,733	2,180	2,032	1,646	1,344	8
9	8,863	8,343	7,843	7,357	6,876	6,393	5,899	5,380	4,817	4,168	3,325	2,700	2,532	2,088	1,735	9
10	9,892	9,342	8,812	8,295	7,783	7,267	6,737	6,179	5,570	4,865	3,940	3,247	3,059	2,558	2,156	10
11	10,920	10,341	9,783	9,237	8,695	8,148	7,584	6,989	6,336	5,578	4,575	3,816	3,609	3,053	2,603	11
12	11,946	11,340	10,755	10,182	9,612	9,034	8,438	7,807	7,114	6,304	5,226	4,404	4,178	3,571	3,074	12
13	12,972	12,340	11,729	11,129	10,532	9,926	9,299	8,634	7,901	7,042	5,892	5,009	4,765	4,107	3,565	13
14	13,996	13,339	12,703	12,078	11,455	10,821	10,165	9,467	8,696	7,790	6,571	5,629	5,368	4,660	4,075	14
15	15,020	14,339	13,679	13,030	12,381	11,721	11,037	10,307	9,499	8,547	7,261	6,259	5,985	5,229	4,601	15
16	16,042	15,338	14,655	13,983	13,310	12,624	11,912	11,152	10,309	9,312	7,962	6,908	6,614	5,812	5,142	16
17	17,065	16,338	15,633	14,937	14,241	13,531	12,792	12,002	11,125	10,085	8,672	7,564	7,255	6,408	5,697	17
18	18,086	17,338	16,611	15,893	15,174	14,440	13,675	12,857	11,946	10,865	9,390	8,231	7,906	7,015	6,265	18
19	19,107	18,338	17,589	16,850	16,109	15,352	14,562	13,716	12,773	11,651	10,117	8,907	8,567	7,633	6,844	19
20	20,127	19,337	18,569	17,809	17,046	16,266	15,452	14,578	13,604	12,443	10,851	9,591	9,237	8,260	7,434	20
21	21,147	20,337	19,548	18,768	17,984	17,182	16,344	15,445	14,439	13,240	11,591	10,283	9,915	8,897	8,034	21
22	22,166	21,337	20,529	19,729	18,924	18,101	17,240	16,314	15,279	14,041	12,338	10,982	10,600	9,542	8,643	22
23	23,185	22,337	21,510	20,690	19,866	19,021	18,137	17,187	16,122	14,848	13,091	11,689	11,293	10,196	9,260	23
24	24,204	23,337	22,491	21,652	20,808	19,943	19,037	18,062	16,969	15,659	13,848	12,401	11,982	10,856	9,886	24
25	25,222	24,337	23,472	22,616	21,752	20,867	19,939	18,940	17,818	16,473	14,611	13,120	12,687	11,524	10,520	25
26	26,240	25,336	24,454	23,579	22,697	21,792	20,843	19,820	18,671	17,292	15,379	13,844	13,409	12,198	11,160	26
27	27,257	26,336	25,437	24,544	23,644	22,719	21,749	20,703	19,527	18,114	16,151	14,573	14,125	12,879	11,808	27
28	28,274	27,336	26,419	25,509	24,591	23,647	22,657	21,588	20,386	18,939	16,928	15,308	14,847	13,565	12,461	28
29	29,291	28,336	27,402	26,475	25,539	24,577	23,567	22,475	21,247	19,768	17,708	16,047	15,574	14,256	13,121	29
30	30,307	29,336	28,386	27,442	26,488	25,508	24,478	23,364	22,110	20,599	18,493	16,791	16,306	14,953	13,787	30
31	31,323	30,336	29,369	28,409	27,438	26,440	25,390	24,255	22,976	21,434	19,281	17,539	17,042	15,655	14,458	31
32	32,339	31,336	30,353	29,376	28,389	27,373	26,304	25,148	23,844	22,271	20,072	18,291	17,783	16,362	15,134	32
33	33,355	32,336	31,337	30,344	29,340	28,307	27,219	26,042	24,714	23,110	20,867	19,047	18,527	17,074	15,815	33
34	34,371	33,336	32,322	31,313	30,293	29,242	28,136	26,938	25,586	23,952	21,664	19,806	19,275	17,789	16,501	34
35	35,386	34,336	33,306	32,282	31,246	30,178	29,054	27,836	26,460	24,797	22,465	20,569	20,027	18,509	17,192	35
40	40,459	39,335	38,233	37,134	36,021	34,872	33,650	32,345	30,956	29,051	26,509	24,433	23,838	22,164	20,707	40
60	60,713	59,335	57,978	56,620	55,239	53,809	52,284	50,641	48,759	46,459	43,188	40,482	39,699	37,485	35,534	60
80	80,927	79,334	77,688	76,188	74,583	72,915	71,145	69,207	66,994	64,278	60,391	57,153	56,213	53,540	51,172	80
90	90,1023	89,334	87,666	85,993	84,285	82,511	80,625	78,558	76,195	73,291	69,126	65,647	64,635	61,754	59,196	90
100	101,115	99,334	97,574	95,808	94,005	92,159	90,133	87,945	85,441	82,358	77,929	74,222	73,142	70,065	67,328	100
120	121,285	119,334	117,404	115,465	113,483	111,419	109,220	106,806	104,037	100,624	95,705	91,573	90,367	86,923	83,828	120
140	141,441	139,334	137,248	135,149	133,003	130,766	128,380	125,758	122,748	119,029	113,659	109,157	107,815	104,034	100,655	140