

UNIVERSIDAD NACIONAL DEL ALTIPLANO
FACULTAD DE INGENIERÍA ESTADÍSTICA E INFORMÁTICA
ESCUELA PROFESIONAL DE INGENIERÍA ESTADÍSTICA E INFORMÁTICA



DATAMART DE LA GERENCIA DE SUPERMERCADO MIA
MARKET DE JULIACA - 2018

TESIS

PRESENTADA POR:

Bach. ALAIN YUCRA HALANOCA

PARA OPTAR EL TÍTULO PROFESIONAL DE:

INGENIERO ESTADÍSTICO E INFORMÁTICO

PUNO – PERÚ

2018

**UNIVERSIDAD NACIONAL DEL ALTIPLANO
FACULTAD DE INGENIERÍA ESTADÍSTICA E INFORMÁTICA
ESCUELA PROFESIONAL DE INGENIERÍA ESTADÍSTICA E INFORMÁTICA**

DATAMART DE LA GERENCIA DE SUPERMERCADO MIA MARKET DE

JULIACA - 2018

TESIS PRESENTADA POR:

Bach. ALAIN YUCRA HALANOCA

PARA OPTAR EL TITULO PROFESIONAL DE:

INGENIERO ESTADÍSTICO E INFORMÁTICO



APROBADO POR EL JURADO REVISOR CONFORMADO POR:

PRESIDENTE:

Dr. PAREDES QUISPE, JUAN REYNALDO

PRIMER MIEMBRO:

D.Sc. APAZA TARQUI, ALEJANDRO

SEGUNDO MIEMBRO:

M.Sc. CHOQUEJAHUA ACERO, REMO

DIRECTOR / ASESOR:

Dr. PEREZ QUISPE, SAMUEL DONATO

Área

: Informática

Tema

: Base De Datos y Sistemas De Información

Fecha de sustentación 27 de Diciembre del 2018

DEDICATORIA

A mis Padres Leoncio Persi y Nora quienes
son el soporte y mi guía para seguir
adelante, a pesar de los obstáculos en mi
vida siempre están presente conmigo
apoyándome.

A mis hermanos Gabriel y Betsabeth que
siempre están ahí, en cada momento de mi
vida aconsejando sabiamente.

Alain

AGRADECIMIENTOS

Primero quiero agradecer a Dios por llenar mi vida de dichas y bendiciones, por estar conmigo en cada paso que doy, por fortalecer mi corazón e iluminar mi mente, por haber puesto en mi camino a aquellas personas que han sido mi soporte y compañía durante todo el período de estudio.

A la Universidad Nacional del Altiplano Puno, a la Facultad de Ingeniería Estadística e Informática. En especial a los docentes, que con su amabilidad y apoyo contribuyen con mi formación profesional impartiendo sus sabios conocimientos.

A mi director de investigación Dr. PEREZ QUISPE, SAMUEL DONATO, docente de la Escuela Profesional de Ing. Estadística e Informática por colaborador durante todo este proceso, quien con su dirección, conocimiento y enseñanza permitiendo el desarrollo de este trabajo de investigación. Gracias por guiarme.

Mi reconocimiento a mi Presidente. Dr. PAREDES QUISPE, JUAN REYNALDO, a mi primer miembro de Jurado D.Sc. APAZA TARQUI, ALEJANDRO, a mi Segundo miembro de Jurado, M.Sc. CHOQUEJAHUA ACERO, REMO, por sus acertadas observaciones y las sugerencias para el desarrollo y culminación del estudio de Investigación.

Alain

ÍNDICE GENERAL

RESUMEN	14
----------------------	-----------

CAPÍTULO I

INTRODUCCIÓN	16
---------------------------	-----------

1.1 DESCRIPCIÓN DEL PROBLEMA	17
1.2 FORMULACIÓN DEL PROBLEMA	24
1.3 JUSTIFICACIÓN DEL PROYECTO	24
1.4 HIPÓTESIS DEL TRABAJO	25
1.5 OBJETIVO GENERAL	25
1.6 OBJETIVOS ESPECÍFICOS:	25
1.7 LIMITACIONES	25

CAPITULO II

REVISIÓN DE LITERATURA	27
-------------------------------------	-----------

2.1 ANTECEDENTES DEL PROYECTO	27
2.2 MARCO TEÓRICO	33
2.2.1. Inteligencia de Negocios (BI)	33
2.2.2. Metodología De Ralph Kimball.....	35
2.2.3. Justificación De Utilización De La Metodología Ralph Kimball.....	46
2.2.4. Proceso ETL.	48
2.2.5. Integration Services SQL Server (SSIS).....	49
2.2.6. Analysis Services SQL Server (SSAS).....	50
2.3 DEFINICIÓN DE TÉRMINOS	51
2.3.1. Sistemas Transaccionales (OLTP).....	51
2.3.2. OLAP- OnLine Analytical Processing	52
2.3.3. OLTP vs OLAP	54

2.3.4.	Sistema de soporte de decisiones (DSS).....	58
2.3.5.	Data Warehouse.....	58
2.3.6.	Datamart.....	59
2.3.7.	Dimensiones.....	60
2.3.8.	Métricas	61
2.3.9.	Indicadores.....	62
2.3.10.	Fact Table	62
2.3.11.	Esquema Estrella.....	64
2.3.12.	Esquema Copo De Nieve.....	65
2.3.13.	Enfoque “Bottom-up”	66
2.3.14.	Enfoque “Top-down”.....	67
2.3.15.	Cubos OLAP.....	67
2.4	OPERACIONALIZACIÓN DE VARIABLES	67

CAPITULO III

MATERIALES Y MÉTODOS	69
3.1 LOCALIZACIÓN DEL PROYECTO.....	69
3.2 MATERIAL DE ANÁLISIS	69
3.2.1. Entrevistas.....	70
3.2.2. Guía de observaciones	70
3.3 MÉTODO DE RECOLECCIONES DATOS.....	70
3.3.1. Entrevistas.....	70
3.3.2. Guía de observaciones	70
3.4 POBLACIÓN.....	71
3.5 DISEÑO DE LA MUESTRA	71
3.6 MÉTODO.....	71
3.7 MÉTODO DE ANÁLISIS DE DATOS.....	71
3.7.1 La entrevista.....	71
3.7.2 Encuestas	72
3.8 METODO DE TRATAMIENTO DE DATOS.....	72

3.9	BASE DE DATOS TRANSACCIONAL EN SQL SERVER.....	73
3.10	DICCIONARIO DE LA BASE DE DATOS GENERAL.....	74
3.11	BASE DE DATOS TRANSACCIONAL.....	74
3.11.1	Descripción de la tabla DATE	74
3.11.2	Descripción de la tabla PRODUCTOS	75
3.11.3	Descripción de la tabla categorías	76
3.11.4	Descripción de la tabla FactInternetSales.....	76
3.11.5	Descripción de la tabla PROVEEDORES.....	77
3.11.6	Descripción de la tabla EMPLEADOS.....	78
3.11.7	Descripción de la tabla CURRENCY	78
3.11.8	Descripción de la tabla REGION.....	79
3.11.9	Descripción de la tabla SalesTerritory.....	79
3.11.10	Descripción de la tabla GEOGRAPHICS.....	79
3.11.11	Descripción de la tabla CLIENTES.....	80
3.12	MODELADO DIMENSIONAL.....	80
3.12.1.	Crear Un Proyecto De Analysis Services	81
3.12.2.	Definir Un Origen De Datos	81
3.12.3	Definir Una Vista De Origen De Datos	84
3.12.4.	Definir e Implementar Un Cubo	86
3.12.5	Definir Un Cubo	88
3.12.6.	Agregar Atributos a Dimensiones.....	92
3.12.7.	Agregar atributos a la dimensión Productos	95
3.12.8.	Implementar un proyecto de Analysis Services.....	99
3.12.9	Examinar el cubo	102

CAPITULO IV

RESULTADOS Y DISCUSIÓN	106
4.1 IDENTIFICACIÓN DE REQUERIMIENTOS	106
4.2 REQUERIMIENTOS FUNCIONALES	106
4.3 REQUERIMIENTOS NO FUNCIONALES	107

4.4 ANÁLISIS DE REQUERIMIENTOS.....	108
4.5 DISEÑO Y DESARROLLO DE LA PRESENTACIÓN DE DATOS	110
4.6 CONSTRUCCIÓN DE PROCESOS ETL.....	111
4.7 DISEÑO DE LA ARQUITECTURA TÉCNICA.....	112
4.8 IMPLEMENTACIÓN DE NEGOCIOS	112
4.9 PRUEBA DE VALIDEZ Y CONSISTENCIA DE DATOS.....	113
4.10 CONTRASTACIÓN DE HIPÓTESIS	114
4.11 SITUACIÓN ACTUAL	115
4.12 RESULTADOS DE LA HIPÓTESIS ESTADÍSTICA.....	119
CONCLUSIONES	122
RECOMENDACIONES	124
REFERENCIAS BIBLIOGRÁFICAS.....	125
ANEXOS.....	129
ANEXO A: Guía de Entrevista N° 01	129
ANEXO B: Guía de Entrevista N° 02.....	130
ANEXO C: Funciones de la Gerencia de Mia Market SRL	132
ANEXO D: Entrevista de evaluación de resultados	135
ANEXO E: Guía de Observación	136

ÍNDICE DE FIGURAS

Figura 1: Gerencia Mía Market	18
Figura 2: Organigrama de soporte y desarrollo supermercado Mia Market	19
Figura 3: Solicitud de reportes al SISDE	20
Figura 4: Entrevista con Ing. Informático	21
Figura 5: Componentes de Inteligencia de negocios.....	34
Figura 6: Fases de la metodología de Ralph Kimball	36
Figura 7: Enfoque del proceso ETL	48
Figura 8: Esquema Estrella.....	64
Figura 9: Esquema copo de nieve.....	65
Figura 10: Ubicación del establecimiento	69
Figura 11: Base de datos transaccional en SQL Server.....	73
Figura 12: Crear nuevo proyecto en Analysis Service	81
Figura 13: Definir un nuevo origen de datos en Analysis Service	82
Figura 14: Definir la conexión con la base de datos	82
Figura 15: Asistente de origen de datos	83
Figura 16: Finalización del Asistente de origen de datos.....	83
Figura 17: Definir Vista de origen de datos	84
Figura 18: Asistente para vistas de origen de datos	84
Figura 19: Asistente de dimensiones.....	86
Figura 20: Selección de dimensiones de la tabla DATE	88
Figura 21: Explorador de soluciones para nuevo cubo	89
Figura 22: Asistente para definir cubos.....	90

Figura 23: Asistente para definir cubos seleccionando nuevas dimensiones	90
Figura 24: Finalizar el asistente renombrando el cubo.....	91
Figura 25: Esquema con implementación completa.....	92
Figura 26: Implementación de la dimensión clientes	94
Figura 27: Implementación de la dimensión Productos	96
Figura 28: Definir medidas del cubo	97
Figura 29: Uso de dimensiones	98
Figura 30: Particiones.....	98
Figura 31: Proceso y ejecución del cubo 1	101
Figura 32: Proceso y ejecución del cubo 2.....	101
Figura 33: Proceso y ejecución del cubo 3.....	102
Figura 34: Examinador del cubo implementado	103
Figura 35: Cubo Implementado.....	104
Figura 36: Realizar Conexión.....	104
Figura 37: Expandir cubo implementado	105
Figura 38: Proceso de diseño de 4 etapas.....	108
Figura 39: Arquitectura de la plataforma de negocios	110
Figura 40: Conectando servidor con SQL Server 2008 R2.....	113
Figura 41: Grafica de T-Student.....	120

ÍNDICE DE TABLAS

Tabla 1: Lista de tiempos de espera de los reportes por módulo del SISDE.....	22
Tabla 2: Justificación de la metodología Ralph Kimball	47
Tabla 3: Diferencias entre las tecnologías OLTP Y OLAP	54
Tabla 4: Operacionalización de variables.....	67
Tabla 5: Descripción de la base de datos transaccional SQL Server	74
Tabla 6: Descripción de la tabla DATE	74
Tabla 7: Descripción de la tabla PRODUCTOS	75
Tabla 8: Descripción de la tabla categorías.....	76
Tabla 9: Descripción de la tabla FACTINTERNETSALES	76
Tabla 10: Descripción de la tabla PROVEEDORES	77
Tabla 11: Descripción de la tabla EMPLEADOS	78
Tabla 12: Descripción de la tabla CURRENCY	78
Tabla 13: Descripción de la tabla REGION.....	79
Tabla 14: Descripción de la tabla SALESTERRITORY	79
Tabla 15: Descripción de la tabla GEOGRAPHICS	79
Tabla 16: Descripción de la tabla CLIENTES	80
Tabla 17: Requerimientos Funcionales	106
Tabla 18: Leyenda de prioridad.....	107
Tabla 19: Leyenda de dificultad.....	107
Tabla 20: Pre-Test del sistema actual para tomar decisiones del supermercado Mia Market	116

Tabla 21: Post-Test del sistema actual para tomar decisiones del supermercado Mia Market	117
Tabla 22: Abreviaturas de calificación.....	118
Tabla 23: Definición de variables.....	118
Tabla 24: Contrastación entre el Pre-test y Post-test.....	118

ÍNDICE DE ACRÓNIMOS

(Inteligencia de Negocios)	BI
(Sistema de información ejecutiva)	EIS
(Sistema de soporte de decisiones)	DSS
(Tecnologías de Información)	TI
(Datawarehouse)	DW
(Procesamiento de transacciones en línea)	OLTP
(Extracción, Transformación y carga)	ETL
(Procesamiento analítico en línea)	OLAP
(Procesamiento analítico multidimensional en línea)	MOLAP
(Sistema de soporte de decisiones)	DSS
(Servicios de análisis SQL server)	SSAS
(Integración de servicios SQL server)	SSIS
(Lenguaje de consulta estructurada)	SQL
(Sistema de soporte y desarrollo)	SISDE

RESUMEN

La presente investigación se desarrolla para el área de gerencia del supermercado Mía Market que requiere acceder a información confiable de manera rápida. El problema carece en solicitar los reportes de las actividades diarias a la oficina de informática, la exportación de la base de datos en un archivo Excel y Pdf; proceso que manualmente genera diversos inconvenientes, como la dependencia del Departamento de Informática, y la información obtenida suele tener grandes posibilidades de error, tiempo en procesos mecánicos con reportes dispersos en hojas impresas, dificultando la consulta de las últimas actualizaciones diarias de los reportes, no generando información de inmediata, a partir de esto se propone como objetivo desarrollar un Datamart para la gerencia Mía Market que requiere tomar decisiones con el fin solucionar este problema, contándose con la tecnología y el personal para la construcción de la herramienta utilizando la metodología de Ralph Kimball, basada en inteligencia de negocios para su futura aplicación con Analysis Services, Cubos OLAP y aplicaciones de software recopilando los datos en SQL Server 2008 Express R2, Visual Studio 2008, que viene a ser un conjunto de programas que gestionan enormes cantidades de información generando inteligencia empresarial para evitar grandes costos en software con licencias. Finalmente se tiene como conclusión, que el sistema desarrollado automatizará el procedimiento en un entorno amigable, permitirá a los directivos de gerencia acceder a toda la información de mejor calidad, más confiable, en el menor tiempo y en un repositorio que facilite acceder a los resultados, lográndose validar y cumplir con el objetivo principal de la investigación.

Palabras Claves: Datamart, Mía Market, Gerencia, Juliaca

ABSTRACT

This research is developed for the management area of the Mía Market supermarket that requires access to reliable information quickly. The problem is not to request the reports of the daily activities to the computer office, the export of the database in an Excel file and Pdf; process that manually generates various problems, such as the dependency of the IT Department, and the information obtained usually has great potential for error, time in mechanical processes with scattered reports in printed sheets, making it difficult to consult the latest daily updates of the reports, generating immediate information, from this is proposed to develop a Datamart for management Mia Market that requires making decisions in order to solve this problem, with technology and staff to build the tool using Ralph's methodology Kimball, based on business intelligence for its future application with Analysis Services, OLAP Cubes and software applications collecting data in SQL Server 2008 Express R2, Visual Studio 2008, which is a set of programs that manage huge amounts of information generating business intelligence to avoid big costs in software with licenses. Finally, it is concluded that the developed system will automate the procedure in a friendly environment, will allow management managers to access all the information of better quality, more reliable, in the shortest time and in a repository that facilitates access to the results, being able to validate and comply with the main objective of the investigation.

Keywords: Datamart, Mia Market, Management, Juliaca

CAPÍTULO I

INTRODUCCIÓN

Durante la última década los supermercados a nivel nacional cuentan con sistemas muy eficientes para mantener sus actividades en un nivel superior controlando sus ventas, empleados, proveedores, clientes, productos y más, ya que el apoyo de la tecnología ha ido mejorando en favor de las necesidades que cubran los requerimientos actuales y competitivos; a nivel nacional Plaza Vea, Oeschle, Tottus y otras entidades de supermercado ya cuentan con dichos sistemas de Datamart, Datawarehouse que ayuda enormemente a las gerencias de las empresas tomar decisiones en el menor tiempo posible, y gracias a estas herramientas de desarrollo se requiere una inmediata implementación de Datamarts y Datawarehouse a nivel regional y local, contando con esas herramientas las pequeñas y micro empresas tendrán. Sin embargo llegar a tomar una decisión no es fácil, para ello se debe de realizar un estudio previo de todas las ventajas y desventajas que esta implica, la investigación plantea desarrollar un Datamart para el área de gerencia que optimice los reportes para almacenar enormes cantidades de datos de proveedores, empleados, productos y clientes que puedan organizar su información y procesarla con las herramientas adecuadas para la toma de decisiones en cuanto al control eficiente de los productos, generación de reportes de ventas y registro de clientes para seguimiento post-venta, los cuales sin un control adecuado generan insatisfacción y pérdida de clientes. Es por esto, que se ve la necesidad de realizar un proyecto que permita brindar una herramienta automatizando el procedimiento necesario, para tomar esta información estructurada como base para que los usuarios de la gerencia puedan analizar y explotar esta información, a continuación, el proyecto de investigación está dividido en capítulos los cuales se detallan en forma breve.

En el capítulo I, aborda la introducción, haciendo una definición y formulación del problema a solucionar que se estará justificando la investigación dando a conocer el objetivo principal y los objetivos específicos de la investigación.

En el capítulo II, se presenta las referencias teóricas y antecedentes que sustentan el trabajo de investigación para cada uno de los objetivos propuestos de una forma crítica, explicando su importancia y relevancia con el trabajo realizado.

En el capítulo III, contiene con detalle el material experimental utilizado en la investigación. Así mismo, los métodos materiales y técnicas empleadas para cada uno de los objetivos señalando la metodología experimental empleada y el análisis utilizado en la interpretación de los datos de la investigación.

En el capítulo IV, se comparan o contrastan los resultados obtenidos con aquellos previamente señalados en la sección de Revisión de Literatura, destacando la nueva información lograda a partir del trabajo de investigación desarrollada.

En el capítulo V, se precisan las conclusiones demostrando el objetivo general y los objetivos específicos de la investigación, contrastando la hipótesis general y específica.

En el capítulo VI, presenta las recomendaciones para el futuro de la investigación.

En el capítulo VII, se da a conocer las referencias bibliográficas, textos, tesis y direcciones web que contribuyeron en el desarrollo de la investigación.

1.1 DESCRIPCIÓN DEL PROBLEMA

Mía Market, es un supermercado situada en la ciudad de Juliaca cuyo lema es: “Compra más, pagando menos”. Para el cumplimiento se estructura de la siguiente manera.

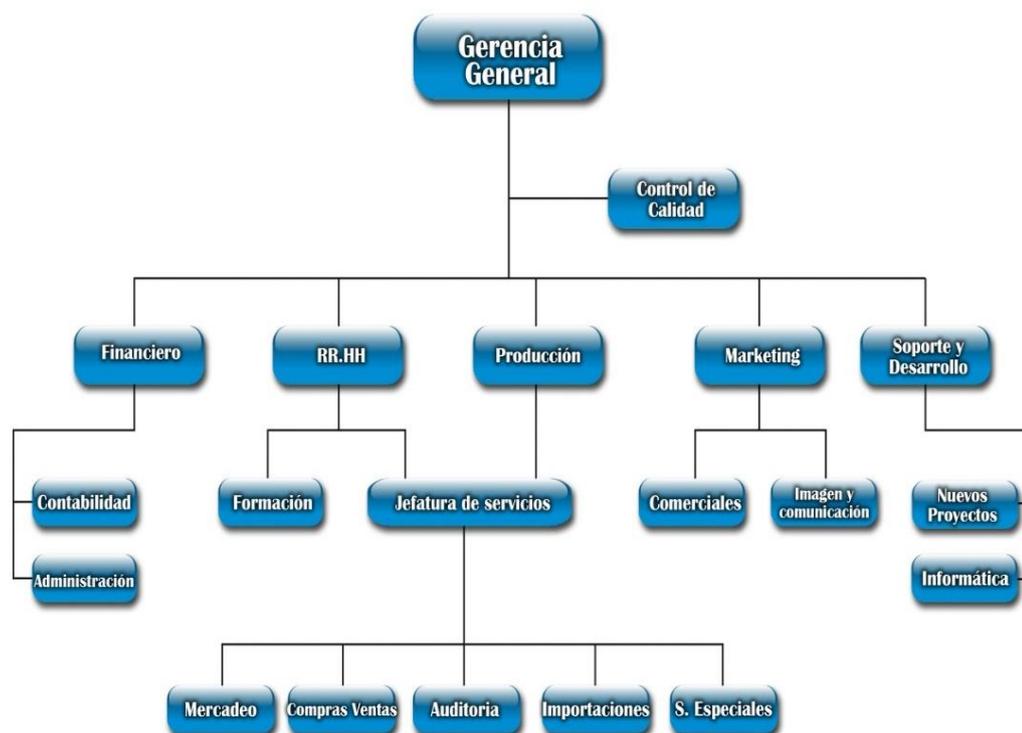


Figura 1: Gerencia Mía Market

La Gerencia de Mía Market se encarga de organizar, estructurar las áreas de Finanzas, Recursos humanos, Producción, Marketing y Soporte y desarrollo; las políticas y procesos para la asignación presupuestal por unidad orgánica, incluyendo las responsabilidades de gasto por partida presupuestal y destino del gasto; el sistema de información de gestión administrativo y financiera y la implementación de la organización para soportar el desarrollo y mantenimiento de los sistemas requeridos por el plan de transformación. De igual forma, conduce la al Control mediante el soporte y desarrollo mediante la oficina de informática y sistemas que está a cargo de un equipo de Ingenieros Informáticos y sistemas; conformada por las unidades orgánicas siguientes:



Figura 2: Organigrama de soporte y desarrollo supermercado Mía Market

Dentro de soporte y desarrollo podemos encontrar el Departamento de informática y nuevos proyectos.

El Departamento de Informática es el área dependiente de la Gerencia, encargado de efectuar la programación, ejecución y evaluación de las actividades relacionadas con los procesos de control informático, contabilidad y tesorería, así como elaborar e interpretar los estados financieros y presupuestales de la empresa Mía Market.

Dentro de las funciones del Departamento de informática, las que están más ligadas a la Alta Dirección son: “Organizar, programar y ejecutar los procesos que permitan presentar a la alta dirección la información económica y financiera oportuna para la toma de decisiones, formular el presupuesto institucional en coordinación con las unidades orgánicas, emitir los informes de disponibilidad presupuestaria para la adquisición o contratación de bienes, servicios u obras que requiera la entidad, formular la consolidación mensual de gastos y efectuar la consolidación con los saldos de balance y presentar a la alta dirección los estados financieros y presupuestales del pliego”.

El departamento de informática, realizan reportes mediante su sistema transaccional actual: Sistema de soporte y desarrollo (SISDE), donde cada usuario del área obtiene sus reportes de manera propia. Cada gerente encuentra los reportes en un menú en el sistema actual.

Está información económica y financiera es solicitada por la Alta Dirección mensualmente para la adecuada toma de decisiones.

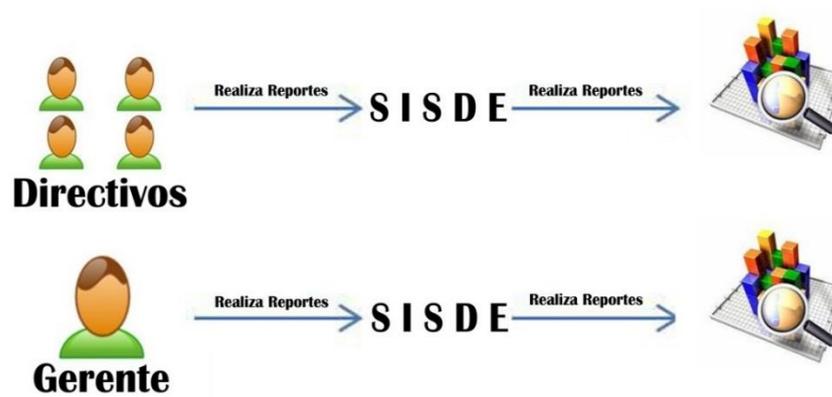


Figura 3: Solicitud de reportes al SISDE

Por otro lado, existe otra forma de realizar reportes, que consiste en presentar una solicitud al Departamento de Tecnologías de la Información. Para la atención de estas solicitudes, el personal accede al servidor de base de datos de SISDE (Sistema de Soporte y Desarrollo) y realiza la exportación de los datos de las tablas involucradas en el requerimiento. Esta área tiene dos opciones de reporte: integrarlo al Sistema de Gestión Administrativo o enviarlo a la Alta Dirección en el momento que lo pide, mediante un archivo de tipo excel o pdf.



Figura 4: Entrevista con Ing. Informático

Al efectuar cualquiera de los dos procedimientos para la generación de informes de entrega a la Alta Dirección o Departamentos que lo soliciten, nos encontramos con un factor muy importante: el tiempo. Estos reportes, debido a la gran cantidad de procesamiento de datos, sus respuestas tienen un tiempo de espera de 8, 10 hasta 30 minutos. Motivo por el cual, la Alta Dirección o Departamentos que lo soliciten, tienen que esperar hasta que el reporte necesario esté listo para su interpretación y la posterior toma de decisiones.

La Alta Dirección no está dispuesta a esperar tanto tiempo para tener el resultado de los reportes, y es una queja común de cada mes; entonces, necesitan información inmediata, precisa, rápida para tomar decisiones. Los tiempos de espera, en cada módulo son:

Tabla 1: Lista de tiempos de espera de los reportes por módulo del SISDE

Módulo	Tiempo de espera SISDE (mensual)
Contabilidad	20 min
Financiero	30 min
Informática	10 min
Tesorería	29 min
Caja Chica	15 min

Entrevista con el encargado de contabilidad (Fuente: Elaboración propia)

Es por eso, que se recurre a una solución de inteligencia de negocios para optimizar los tiempos de espera de reportes, obteniéndolos de manera inmediata y confiable para la toma de decisiones.

Mia Market nace en 2002 con la denominación de minimarket en la ciudad de Juliaca y quince años después nos volvimos parte del Grupo de supermercado siendo la primera en iniciar netamente desde abajo con nuestros propios recursos de la zona, grupo que nos dio el respaldo económico para seguir creciendo y buscando mejorar la vida de más familias juliaqueñas.

En el 2012 comenzamos nuestra expansión de sucursales, siendo los primeros en abrir un nuevo local con todas las tecnologías de punta; logrando presencia de más clientes compitiendo con enormes cadenas de supermercados como plaza vea, Real plaza, Oechle y demás que se instalan en nuestra zona.

En el 2012 conseguimos la certificación internacional HACCP para los alimentos frescos. El 2015 renovamos nuestra imagen buscando refrescarla y adecuarla frente a un Perú cada vez más moderno y en crecimiento. Buscamos ser la primera opción de compra para todos los juliaqueños y damos empleo a más de 300 colaboradores en quienes inculcamos nuestros valores:

Pensando en ti y en tu familia, en el año 2002, lanzamos los primeros productos con nuestro nombre y personalidad considerando la canasta básica de nuestros clientes: aceite, arroz y azúcar.

Hasta la fecha seguimos innovando día a día para ofrecerte más productos. Hoy nos sentimos orgullosos de ofrecer más de 800 productos en las siguientes categorías:

- *Alimentos de mascotas*
- *Abarrotes dulces y salados*
- *Bebidas y helados*
- *Cuidado personal y limpieza*
- *Frutas y verduras*
- *Comidas al paso*
- *Panadería y pastelería*

Nuestros productos Mía Market son elaborados por los principales productores de alimentos en el Perú y el mundo, pasan por los más altos estándares de calidad y se caracterizan por tener un inigualable sabor e increíbles precios. ¡De esta forma tú y tu familia pueden tenerlos siempre en sus hogares!

Misión

Ser una empresa Juliaqueña líder en la comercialización de productos de consumo, teniendo como objetivo central satisfacer las necesidades de nuestros clientes, con por tal motivo brindamos productos de calidad a precios bajos y nos esforzamos por ofrecer un excelente servicio, con una atención personalizada y de esta manera crear una agradable experiencia de compra para que nuestros leales clientes regresen. Asimismo, nos preocupamos por el bienestar y desarrollo de nuestros trabajadores.

Visión

Ser el supermercado más importante del sur reconociendo y agradeciendo a todos nuestros clientes y proveedores para poder seguir creciendo y aperturando mas sucursales; y así aumentar nuestras ventas a nuestros clientes.

1.2 FORMULACIÓN DEL PROBLEMA

¿Cuál será la eficiencia del Datamart desarrollado para la gerencia de supermercado Mia Market de la ciudad de Juliaca?

1.3 JUSTIFICACIÓN DEL PROYECTO

El presente proyecto nace por la necesidad de generar información consistente en tiempo real de los datos que se almacenan en nuestro Datawarehouse y Datamart. También requiere reportes que se tiene directamente de la base de datos sin la necesidad de hacer consultas manuales desde SQL server que no todos los directivos de la empresa tienen conocimiento del manejo de la herramienta que no ofrecen cierta conformidad por parte de las estrategias para promocionar los productos que se oferta, para esta investigación se tomó como referencia los supermercados más grandes a nivel nacional e internacional que ya poseen su propio sistema en cargado de optimizar sus negocios comerciales, creándose con anterioridad la base de datos que servirá para este proyecto de investigación y que será de utilidad para las futuras investigaciones, haciendo una cadena de distintos proyectos a futuro, siendo el mismo equipo de desarrolladores en diseño, programación y usuarios para que los problemas sean más sencillos de identificar teniendo de base un Datamart y haciendo mucho más uso de este sistema se podrá realizar todo tipo de análisis y reportes.

1.4 HIPÓTESIS DEL TRABAJO

El desarrollo de un Datamart para la gerencia del supermercado Mía Market tendrá la aceptación correspondiente para seguir expandiendo el mercado.

1.5 OBJETIVO GENERAL

Desarrollar un Datamart para la gerencia de supermercado Mía Market de la ciudad de Juliaca, el cual se verificará con los siguientes:

1.6 OBJETIVOS ESPECÍFICOS:

- Identificar las fuentes de datos requeridas del supermercado y procesarlas en un Datamart del supermercado Mía Market de la ciudad de Juliaca.
- Analizar el modelamiento multidimensional del proyecto para las distintas áreas del supermercado Mía Market de la ciudad de Juliaca.
- Diseñar el Datamart en base a la metodología Ralph Kimball que cumpla con los requerimientos necesarios para la gerencia del supermercado Mía Market de la ciudad de Juliaca.
- Implementar la herramienta para desplegar el cubo OLAP del supermercado Mía Market de la ciudad de Juliaca.

1.7 LIMITACIONES

El proyecto solo será utilizado para el área de gerencia por los directivos autorizados y jefes de cada área por motivos de que la información que se obtendrá sirve para las proyecciones a futuro de la empresa sin intervención de los demás Departamentos.

Será implementado solo en la ciudad de Juliaca, limitando a las demás sucursales de contar con la solución de inteligencia de negocios.

Debido a la pronta migración de supermercado a Mall Mia Market, no fue autorizado realizar un Datamart con inteligencia de negocios, la cual se realizará en futuros proyectos de investigación.

CAPITULO II

REVISIÓN DE LITERATURA

2.1 ANTECEDENTES DEL PROYECTO

A NIVEL INTERNACIONAL

LOPEZ DILMAR, (2007). “Análisis, Diseño e Implementación de un Data Mart para la dirección financiera y recursos humanos de la Escuela Politécnica del Ejército para una toma de decisión Efectiva”. Tesis de Bachiller publicada. Departamento de ciencias de la computación de la Escuela Politécnica del Ejército, Sagolqui, Ecuador. Cuyo objetivo general es brindar una herramienta de apoyo en la tecnología Business Intelligence de Oracle que sea capaz de solventar la información solicitada en la Dirección Financiera y Recursos Humanos de la Escuela Politécnica del Ejército, dando la posibilidad de formar cubos dimensionales de decisión. Para el trabajo de la tesis, se utilizó la metodología Rapid Warehousing, este método es iterativo, y está basado en el desarrollo incremental del proyecto de Data Warehouse, que se basa en la definición de un DataMart como la combinación de estructuras de datos especializados para resolver requerimientos de análisis de información en distintas áreas de una institución. La tesis se enfoca en el control de recursos financieros y humanos y guarda relación con el tema de investigación, de ahí que, ambos se enfocan en la ejecución de una Data Mart, para el Departamento de Finanzas y así conseguir una efectiva toma de decisiones.

ILLBAY MAYDANA, (2009). “Propuesta Metodológica para Aplicar Business Intelligence Caso Práctico Cohervi S.A.” Tesis de Bachiller. Facultad de ciencias de la computación de la carrera profesional de Ingeniería Informática, Santiago, Chile. El

cual su compromiso principal es la satisfacción de las necesidades y expectativas del cliente, suministrándole productos de bajos precios. Llega a la siguiente conclusión: “La metodología planteada para implementar una solución Business Intelligence en la empresa COHERVI S.A., obteniéndose una mejora en la elaboración de informes de análisis y satisfacción de las áreas departamentales, considerando tres variables: el tiempo de elaboración de informes, complejidad de elaboración de los informes y satisfacción del usuario estratégico. Esto a su vez, permite ejecutar una planificación más adecuada con respecto a: la de cantidad de compra de productos proyecciones de ventas, analizar el movimiento de los distintos clientes y buscar nuevos mercados”

GUTIERREZ COLMENARES, (2016). “Construcción de un sistema de apoyo a la toma de decisiones para el área gerencial del Hospital de Clínicas” Tesis de PostGrado. Facultad de ciencias de la computación de la Escuela profesional de Ingeniería Informática y ciencias de la computación, Asunción, Paraguay. En su trabajo de investigación, analiza distintos sistemas de manejo de información, resaltando las ventajas de los más avanzados frente a los primitivos, que ofrecen mínima interacción entre el usuario y la máquina, son de respuesta lenta y no permiten respuestas a cuestiones complejas de interés para el usuario. Sistemas más modernos ofrecen mejoras, en cuanto a herramientas de análisis, mayor capacidad de manejo de información, interacciones complejas, y uso de sistemas operacionales presentes en la empresa. Llega a la siguiente conclusión: “El proceso de implementación del sistema automatizado permitió a los usuarios conocer de manera detallada las actividades y tiempos necesarios para la completa puesta en marcha de tal sistema, identificándose

los cuellos de botella, que necesitan ser atendidos previamente para una exitosa implementación.”

A NIVEL NACIONAL

NUÑEZ BEJARANO, (2010). "Análisis, Diseño e Implementación de una Solución de Inteligencia de Negocio para el área de Finanzas de la Municipalidad de Lima". Tesis de Bachiller publicada. Pontificia Universidad Católica del Perú, Lima, Perú. Esta investigación tiene como objetivo analizar, diseñar e implementar una solución de inteligencia de negocios para el Área de Finanzas, de manera que, los usuarios cuenten con información sólida, confiable y de alta disponibilidad. La realidad problemática de esta tesis consiste en que cada uno de los usuarios al realizar los procedimientos de reportes de forma manual, invierte mucho tiempo y los resultados obtenidos no siempre son exactos. Esto debido a que, al manipular una gran cantidad de datos, están propensos al error. La metodología que usó el proyecto fue una adaptación de la guía del PMBOK (Project Management Body of Knowledge—guía de los fundamentos de la Dirección de Proyectos), desarrollada por el PMI (Project Management Institute - Instituto de Administración de Proyectos). La guía del PMBOK es un estándar reconocido internacionalmente que agrupa un conjunto de buenas prácticas generalmente aceptadas en lo que se refiere a la gestión de proyectos. Está estructurado en cinco grupos de procesos básicos y nueve áreas de conocimiento comunes a casi todos los tipos de proyectos. La relación que guarda con esta tesis es el tema de investigación, porque ambos se enfocan en el desarrollo de un DataMart para una entidad pública, específicamente para el Área de Finanzas. Otro aspecto

relacional, se encuentra en la realidad problemática, por lo mismo, que la información requerida no se proporciona en el momento oportuno para la toma de decisiones.

FERNANDEZ BERMUDEZ, (2010). "Análisis, Diseño e Implementación de un Datamart de clientes para el área de Marketing de una Entidad Aseguradora. Tesis de Bachiller publicada. Pontificia Universidad Católica del Perú, Lima, Perú". La empresa en donde se implementó la presente solución maneja en su sistema transaccional información de clientes, productos, siniestros, pólizas, certificados y planes. El Área de Marketing necesita explotar la misma para poder analizar la información y tomar mejores decisiones. La tesis consiste en atender las necesidades de explotación de información del Área de Marketing mediante reportes e indicadores de una entidad aseguradora. La metodología de la elaboración del producto es Business Intelligence Roadmap, que consiste en una guía de buenas prácticas que sigue 16 pasos donde se recogen las actividades a realizar en un proyecto de inteligencia de negocios.

VILLANUEVA FUENTES, (2008). "Análisis, Diseño e Implementación de un Datawarehouse de soporte de decisiones para un Hospital del Sistema de Salud Público. Tesis de Bachiller publicada. Pontificia Universidad Católica del Perú, Lima, Perú". La realidad problemática de esta tesis precisa que al generar un reporte en el sistema transaccional de salud pública demora días e incluso semanas, ya que se tiene que cumplir obligatoriamente una serie de pasos administrativos para elaborarlo. En cambio, con herramientas tecnológicas, estos reportes son generados, en forma personalizada, en minutos o en segundos. Este proyecto de tesis, tiene como objetivo implementar una Data Warehouse que ayude al proceso de la toma de decisiones para un hospital del sistema de salud pública.

A NIVEL REGIONAL

QUISPE PACO, (2012). “Sistema de inteligencia de negocios para el soporte de toma de decisiones en el área de nutrición y aprendizaje infantil del programa nacional Wawa Wasi sede Puno”. Tesis de bachiller publicada de la Universidad Nacional del Altiplano, Facultad de Ingeniería Mecánica Eléctrica, Electrónica y Sistemas, Escuela Profesional de Ingeniería de Sistemas. Esta investigación tiene como objetivo general determinar que el sistema de inteligencia de negocios dará el soporte a la toma de decisiones en el área de nutrición y aprendizaje infantil del Programa Nacional Wawa Wasi Sede Puno. El cual se verificará con los siguientes objetivos específicos: Obtener los requerimientos del área de Nutrición y Aprendizaje Infantil para diseñar la base de datos multidimensional, Implementar el sistema de inteligencia de negocios para el soporte de toma de decisiones y validar la aceptación del Sistema de Inteligencia de Negocios para el Soporte de Toma de Decisiones. Este investigación concluye con: PRIMERA: Con el sistema de inteligencia de negocios para el soporte de toma de decisiones en el área de nutrición y aprendizaje infantil del Programa Nacional Wawa Wasi Sede Puno, se logró determinar cuál es la cantidad de niñas y niños en situación de desnutrición aguda (peso/talla) y desnutrición crónica (talla/edad), los cuales son factores determinantes en el desarrollo del niño de la primera infancia, para tomar las decisiones adecuadas recurriendo a las líneas de acción para su correcto desarrollo, de acuerdo a los lineamientos que posee en el Programa Nacional Wawa Wasi, también se pudo consolidar información acerca de las zonas en las cuales las niñas y niños son vulnerables de caer en riesgo de desnutrición. SEGUNDA: Con el sistema de inteligencia de negocios para el soporte de toma de decisiones en el área de nutrición y

aprendizaje infantil del Programa Nacional Wawa Wasi Sede Puno, se logró detectar a los niños en los cuales la estimulación de Aprendizaje Infantil Temprano no se está logrando, o no se está concretando de acuerdo a los lineamientos del Programa Nacional Wawa Wasi, para lo cual es personal toma las decisiones adecuadas para revertir este inconveniente que es otro punto fundamental en el desarrollo del niño.

GUILLÉN QUILCA, (2017). “Sistema de soporte de decisiones con tecnología data Warehouse para la gestión de la información de la empresa Mallku Import SAC - Juliaca 2016”, Tesis de bachiller publicada de la Universidad Nacional del Altiplano, Facultad de Ingeniería Mecánica Eléctrica, Electrónica y Sistemas, Escuela Profesional de Ingeniería de Sistemas, La presente investigación tiene por finalidad el desarrollo de un Sistema de Soporte de Decisiones con Tecnología Data Warehouse que optimice la Gestión de la Información de la Empresa Mallku Import SAC, la empresa en mención tiene su centro de operaciones en la ciudad de Juliaca, región de Puno y se dedica al rubro de la importación de máquinas remanufacturadas como son fotocopiadoras, impresoras, repuestos, insumos y accesorios; donde sus operaciones diarias son la venta de estos productos. Encontrando como uno de los principales problemas la mala gestión de la información, el desaprovechamiento de información almacenada a diario, no generando información a partir de esto. Por lo cual se plantea un sistema de soporte de decisiones basado en inteligencia de negocios que tendrá como fin solucionar este problema. Para el desarrollo del software se utilizó la metodología de Moss Larissa que es basada en Inteligencia de Negocios de Roadmap y el enfoque de Ralph Kimball para la construcción del Data Warehouse y como herramientas de software se utilizó la suite de Inteligencia de Negocios “Pentaho” que

viene a ser un conjunto de programas libres para generar inteligencia empresarial escrita en Java con un ambiente de implementación basado en Java, para evitar costos altísimos en las licencias de software. Finalmente se tiene como principal conclusión, que el sistema desarrollado ha optimizado la gestión de la información en un más del 83%, logrando validar y cumplir con el objetivo principal de la investigación.

2.2 MARCO TEÓRICO

2.2.1. Inteligencia de Negocios (BI)

En el “Boletín de Asesoría Gerencial” proporcionado por Espiñera, Sheldon y asociados, indican que las definiciones de inteligencia de negocios incluyen una amplia categoría de metodologías, aplicaciones y tecnologías que permiten reunir, acceder, transformar y analizar los datos, transacciones e información no estructurada (interna y externa), con el propósito de ayudar a los usuarios de una compañía a tomar mejores decisiones de negocio.

Por otra parte, en el libro Business Intelligence: Josep Lluís Cano, menciona que el objetivo básico de inteligencia de negocios es apoyar de forma sostenible y continuada a las organizaciones para mejorar su competitividad, facilitando la información necesaria para la toma de decisiones.

Asimismo, describe tres tipos de beneficios que se pueden obtener a través del uso de inteligencia de negocios, los cuales son:

- Beneficios tangibles: por ejemplo, reducción de costos, generación de ingresos, reducción de tiempos para las distintas actividades del negocio.

- Beneficios intangibles: el hecho de que tengamos disponible la información para la toma de decisiones hará que más usuarios la utilicen para tomar decisiones y mejorar nuestra posición competitiva.
- Beneficios estratégicos: la formulación de estrategias, especificarán a qué clientes, mercados o con qué productos dirigirnos. A su vez, el mismo autor, nos indica los componentes de inteligencia de negocios:

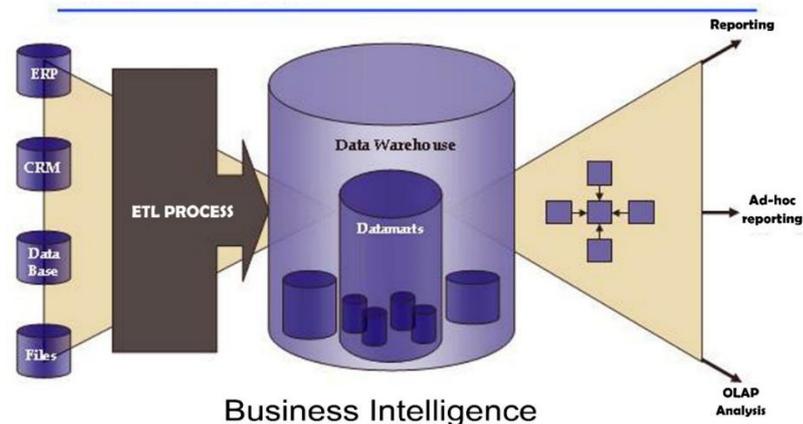


Figura 5: Componentes de Inteligencia de negocios

Fuentes de información, de las cuales partiremos para alimentar de información el Data Warehouse.

Proceso ETL de extracción, transformación y carga de los datos en la Data Warehouse. Antes de almacenar los datos en una Data Warehouse, éstos deben ser transformados, limpiados, filtrados y redefinidos. Normalmente, la información que tenemos en los sistemas transaccionales no está preparada para la toma de decisiones.

La propia Data Warehouse: se busca almacenar los datos de una forma que maximice su flexibilidad, facilidad de acceso y administración.

El motor OLAP, nos provee capacidad de cálculo, consultas, funciones de planeamiento, pronóstico y análisis de escenarios en grandes volúmenes de datos. En la actualidad, existen otras alternativas tecnológicas al OLAP.

Las herramientas de visualización, nos permiten el análisis y la navegación a través de los mismos.

2.2.2. Metodología De Ralph Kimball

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

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

- Centrarse en el negocio
- Construir una infraestructura de información adecuada
- Realizar entregas en incrementos significativos (este principio consiste en crear el almacén de datos (DW) en incrementos entregables en plazos de 6 a 12 meses, en este punto, la metodología se parece a las metodologías ágiles de construcción de software)

- Ofrecer la solución completa (En este se punto proporcionan todos los elementos necesarios para entregar valor a los usuarios de negocios, para esto ya se debe tener un almacén de datos bien diseñado, se deberán entregar herramientas de consulta ad hoc, aplicaciones para informes y análisis avanzado, capacitación, soporte, sitio web y documentación).

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

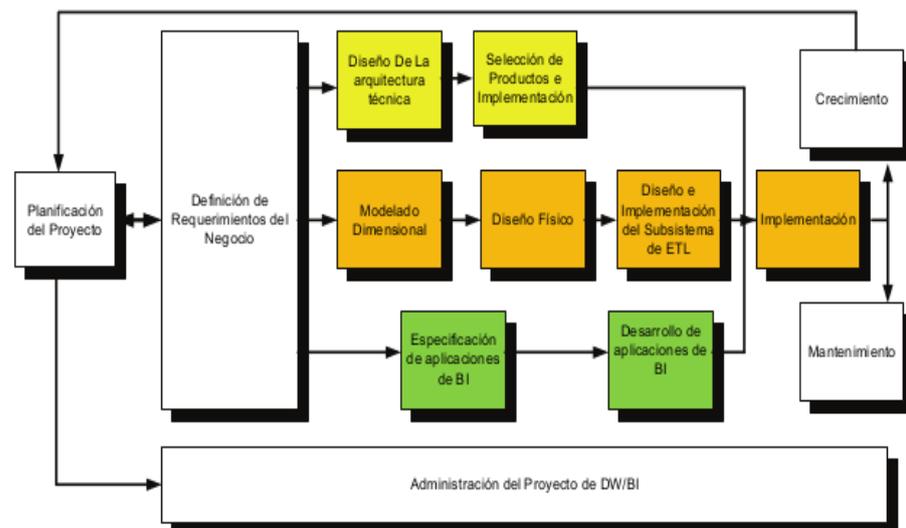


Figura 6: Fases de la metodología de Ralph Kimball

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

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

Definir el alcance (entender los requerimientos del negocio).

Identificar las tareas

Programar las tareas

Planificar el uso de los recursos.

Asignar la carga de trabajo a los recursos

Elaboración de un documento final que representa un plan del proyecto.

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

Monitoreo del estado de los procesos y actividades.

Rastreo de problemas

Desarrollo de un plan de comunicación comprensiva que dirija la empresa y las áreas de TI

Definición De Requerimientos Del Negocio

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

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

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

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

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

Modelado Dimensional.

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

- A.** Elegir el proceso de negocio: que consiste en, elegir el área a modelizar.

Esta es una decisión de la dirección, y depende fundamentalmente del

análisis de requerimientos y de los temas analíticos anotados en la etapa anterior.

- B.** Establecer el nivel de granularidad: La granularidad significa especificar el nivel de detalle. La elección de la granularidad depende de los requerimientos del negocio y lo que es posible a partir de los datos actuales. La sugerencia general es comenzar a diseñar el DW al mayor nivel de detalle posible, ya que se podrían realizar agrupamientos posteriores, al nivel deseado.
- C.** Elegir las dimensiones: Las dimensiones surgen naturalmente de las discusiones del equipo, y facilitadas por la elección del nivel de granularidad y de la matriz de procesos/dimensiones. Las tablas de dimensiones tienen un conjunto de atributos (generalmente textuales) que brindan una perspectiva o forma de análisis sobre una medida en una tabla hechos. Una forma de identificar las tablas de dimensiones es que sus atributos son posibles candidatos para ser encabezado en los informes, tablas pivot, cubos, o cualquier forma de visualización, unidimensional o multidimensional.
- D.** Identificar medidas y las tablas de hechos: Este paso, consiste en identificar las medidas que surgen de los procesos de negocios. Una medida es un atributo (campo) de una tabla que se desea analizar, sumando o agrupando sus datos y usando los criterios de corte conocidos como dimensiones. Las medidas habitualmente se vinculan con el nivel de granularidad del punto 2, y se encuentran en tablas que

denominamos tablas de hechos (fact en inglés). Cada tabla de hechos tiene como atributos una o más medidas de un proceso organizacional, de acuerdo a los requerimientos. Un registro contiene una medida expresada en números, como ser cantidad, tiempo, dinero, etc., sobre la cual se desea realizar una operación de agregación (promedio, conteo, suma, etc.) en función de una o más dimensiones. La granularidad, en este punto, es el nivel de detalle que posee cada registro de una tabla de hechos.

Diseño Físico

En esta tarea, se contestan las siguientes preguntas:

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

- ¿Debe usarse la partición en las tablas relacionales?

Diseño e Implementación Del Subsistema De Extracción, Transformación Y Carga (ETL)

Esta etapa es típicamente la más subestimada de las tareas en un proyecto de Data Warehouse. Las principales sub-etapas de esta zona del ciclo de vida son: la extracción, la transformación y la carga

Se define como proceso de extracción a aquellos requeridos para obtener los datos que permitirán efectuar la carga del modelo físico acordado. También se definen como proceso de transformación el convertir o recodificar los datos fuente, a fin poder efectuar la carga efectiva del modelo físico. Por otra parte, los procesos de carga de datos son los requeridos para poblar la Data Warehouse.

Todas estas tareas son altamente críticas, pues tienen que ver con la materia prima de la Data Warehouse: los datos. La desconfianza y pérdida de credibilidad de la Data Warehouse serán resultados inmediatas e inevitables, si el usuario chocó con información inconsistente. Es por ello, que la calidad de los datos es un factor determinante en el éxito de un proyecto de Data Warehousing. Es en esta etapa donde deben sanearse todos los inconvenientes relacionados con la calidad de los datos fuente.

Como advierte Kimball, el proceso de Data Staging es el iceberg de un proyecto de Data Warehousing. Son muchos los desafíos que deben enfrentarse para lograr datos de alta calidad de los sistemas fuentes. En

general, es una de las etapas más subestimadas que siempre termina tomando más tiempo del previsto.

Mantenimiento y Crecimiento Del Data Warehouse

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

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

Selección del producto e implementación

Utilizando como marco el diseño de arquitectura técnica es necesario evaluar y seleccionar componentes específicos de la arquitectura como ser la plataforma de hardware, el motor de base de datos, la herramienta de ETL o el desarrollo pertinente, herramientas de acceso, etc.

Una vez evaluados y seleccionados los componentes determinados se procede con la instalación y prueba de los mismos en un ambiente integrado de Data Warehousing.

Especificación De Aplicaciones De BI

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

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

transaccionales basándose en los conocimientos obtenidos del uso de la aplicación de BI. Algunas aplicaciones analíticas comunes incluyen:

- a. Análisis de la eficacia de las promociones
- b. Análisis de rutas de acceso en un sitio Web
- c. Análisis de afinidad de programas
- d. Planificación del espacio en espacios comerciales
- e. Detección de fraudes
- f. Administración y manejo de categorías de productos

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

Diseño De La Arquitectura Técnica

El área de arquitectura técnica cubre los procesos y herramientas que se aplican a los datos. En el área técnica existen dos conjuntos que tienen distintos requerimientos, brindan sus propios servicios y componentes de almacenaje de datos, por lo que se consideran cada uno aparte: El back room (habitación trasera) y el front room (habitación frontal). El back room es el responsable de la obtención y preparación de los datos, por lo que también se conoce como adquisición de datos y el front room es responsable de entregar los datos a la comunidad de usuario y también se le conoce como acceso de datos.

Desarrollo de aplicaciones de BI

Siguiendo a la especificación de las aplicaciones para usuarios finales, el desarrollo de las aplicaciones de los usuarios finales involucra configuraciones de la meta data y construcción de reportes específicos. Una vez que se ha cumplido con todos los pasos de la especificación y se tiene la posibilidad de trabajar con algunos datos de prueba, comienza el desarrollo de la aplicación.

A. Implementación

La implementación representa la convergencia de la tecnología, los datos y las aplicaciones de usuarios finales accesible desde el escritorio del usuario del negocio. Existen factores extras que aseguran el correcto funcionamiento de todas estas piezas, entre ellos, se encuentran la capacitación, el soporte técnico, la comunicación, las estrategias de feedback. Todas estas tareas deben ser tenidas en cuenta antes de que cualquier usuario pueda tener acceso al Data Warehouse.

B. Mantenimiento y Crecimiento

El Data Warehousing es un proceso (etapas bien definidas, con comienzo y fin, pero de naturaleza espiral), pues acompaña a la evolución de la organización durante toda su historia. Se necesita continuar con los relevamientos de forma constante para poder seguir la evolución de las metas por conseguir. Según, afirma Kimball, “Si se ha utilizado el ciclo de vida dimensional del negocio, la Data Warehouse está preparado para evolucionar y crecer”. Al contrario de los sistemas tradicionales, los cambios en el desarrollo deben ser

vistos como signos de éxito y no de falla. Es importante establecer las prioridades para poder manejar los nuevos requerimientos de los usuarios y de esa forma poder evolucionar y crecer.

C. Administración Del Proyecto

El gerenciamiento del proyecto asegura que las actividades del ciclo de vida dimensional del negocio se lleven en sincronizadas y en la mejor forma. Entre las actividades principales se encuentra el monitoreo del estado del proyecto, la comunicación entre los requerimientos del negocio y las restricciones de información para poder manejar correctamente las expectativas en ambos sentidos.

2.2.3. Justificación De Utilización De La Metodología Ralph Kimball

A continuación, se presenta un cuadro resumido a cerca de la metodología con objetivos, diseño del datamart, enfoque, tiempo de implementación del Datamart, costos y su respectivo modelado de datos

Tabla 2: Justificación de la metodología Ralph Kimball
RALPH KIMBALL

Objetivo	Todas las empresas necesitan almacenar, analizar e interpretar los datos que van generando y acumulando, para luego tomar decisiones críticas que les permitan maximizar la prosperidad. Para ello, se necesita un sistema que les ayude a entender los datos y logren cumplir sus objetivos, de esta forma nace la idea de “implementar una Data Warehouse”.
Diseño del Datamart	Utiliza el enfoque “Bottom – Up”
Enfoque	Tiene un enfoque por procesos que son manejados por las diferentes áreas del proceso. Trata de responder necesidades específicas según el tema.
Tiempo de implementación del Datamart	Debido a que en primer lugar debemos implementar los Datamarts, el tiempo de implementación es rápido. Sin embargo, se tiene que tener cuidado ya que si se trabaja de forma independiente cada Datamart el entorno del DWH se desintegraría rápidamente.
Costos	Implementar cada Datamart permite que la solución no presente un alto costo.
Modelos de datos	Kimball plantea usar el modelamiento dimensional: esquema estrella. Identificación de dimensiones y hechos.

Fuente: Metodología Ralph Kimball

Para la construcción del Datamart del área de gerencia del supermercado Mía Market, usaremos la metodología propuesta por Ralph Kimball, dado que el ámbito y los recursos destinados al proyecto son de un área en específico, proporcionando un enfoque de menor a mayor, muy versátil y dejando la puerta abierta a una ampliación posterior dentro del ámbito de la Contraloría General de la República. Kimball asegura la usabilidad de los usuarios con un sistema fácil de entender y con un rápido desarrollo.

2.2.4. Proceso ETL.

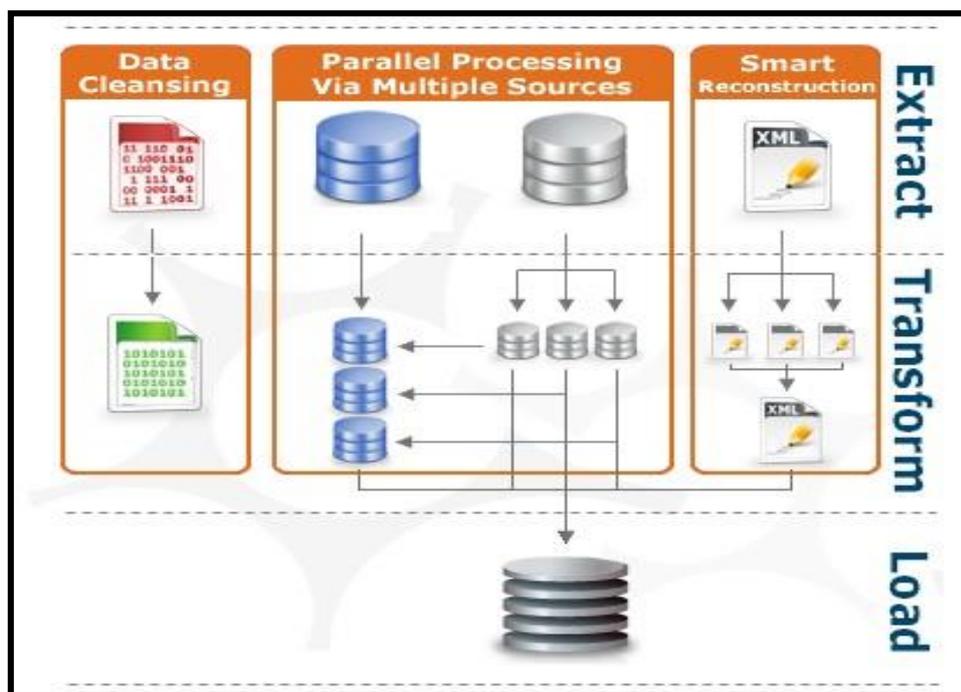


Figura 7: Enfoque del proceso ETL

Los procesos ETL son procesos que leen los registros de las fuentes de datos, aplican las transformaciones necesarias para prepararlos y los cargan en el destino de datos.

El proceso ETL se divide en 5 subprocesos:

- **Extracción:** recupera los datos físicamente de las distintas fuentes de información. En este momento disponemos de los datos en bruto.
- **Limpieza:** recupera los datos en bruto y comprueba su calidad, elimina los duplicados y, cuando es posible, corrige los valores erróneos y completa los valores vacíos, es decir se transforman los datos -siempre que sea posible para reducir los errores de carga. En este momento disponemos de datos limpios y de alta calidad.

- **Transformación:** recupera los datos limpios, de alta calidad, los estructura y resume en los distintos modelos de análisis. El resultado de este proceso es la obtención de datos limpios, consistentes, resumidos y útiles.
- **Integración:** valida que los datos que cargamos en la Data Warehouse sean consistentes con las definiciones y formatos del Data Warehouse; los integra en los distintos modelos de las distintas áreas de negocio que hemos definido en el mismo. Estos procesos pueden ser complejos.
- **Actualización:** es el que nos permite añadir los nuevos datos a los Datamarts

2.2.5. Integration Services SQL Server (SSIS)

Salvador Ramos, en el documento “Integration Services, pieza fundamental en los proyectos de BI” define a Microsoft Integration Services como una plataforma para la creación de soluciones empresariales de transformaciones de datos e integración de datos. Integration Services sirve para resolver complejos problemas empresariales mediante la copia o descarga de archivos, el envío de mensajes de correo electrónico como respuesta a eventos, la actualización de almacenamientos de datos, la limpieza y minería de datos, y la administración de objetos y datos de SQL Server. Los paquetes pueden funcionar por separado o conjuntamente con otros paquetes para hacer frente a las complejas necesidades de la empresa. Integration Services puede extraer y transformar datos de muchos orígenes distintos,

como archivos de datos XML, archivos planos y orígenes de datos relacionales, y, posteriormente, cargarlos en uno o varios destinos.

Integration Services contiene un variado conjunto de tareas y transformaciones integradas, herramientas para la creación de paquetes y el servicio Integration Services para ejecutar y administrar los paquetes. Las herramientas gráficas de Integration Services se pueden usar para crear soluciones sin escribir una sola línea de código. También, se puede programar el amplio modelo de objetos de Integration Services para crear paquetes

Mediante programación y codificar tareas personalizadas y otros objetos de paquete. Integration Services no es solo una herramienta para ETL, sino que tiene una serie de tareas orientadas a la administración que serán utilizadas por los BA's. Integration Services es una herramienta muy completa y con muchas funcionalidades, e incluso con ciertos riesgos si no hacemos un uso apropiado de ella.

2.2.6. Analysis Services SQL Server (SSAS)

La empresa consultec en el documento Microsoft SQL Server 2008 Express R2 funcionalidades de business intelligence define que Microsoft SQL Server Analysis Services es una herramienta que permite a las organizaciones mejorar y optimizar la toma de decisiones empresariales dentro de la organización.

Incorpora funcionalidades OLAP (Online Analytical Processing) y de data Mining para las aplicaciones de inteligencia empresarial. Analysis Services soporta la creación y administración OLAP, que son unas bases de datos

analíticas, permitiéndole diseñar, crear y gestionar estructuras multidimensionales que contienen datos agregados a partir de orígenes diversos, como bases de datos relacionales.

2.3 DEFINICIÓN DE TÉRMINOS

2.3.1. Sistemas Transaccionales (OLTP)

Los OLTP (Online Transaction Processing) son sistemas transaccionales que están altamente afinados para realizar su trabajo rápidamente, usualmente en tiempo real, y a menudo con el uso de y otros servidores grandes. Capturan las transacciones de un negocio y las persisten en estructuras relacionales llamadas base de datos. Las características principales de los sistemas OLTP son:

Realizan transacciones en tiempo real del proceso de un negocio, con lo cual los datos almacenados cambian continuamente. Los sistemas OLTP en sus transacciones conducen procesos esenciales del negocio.

Los sistemas OLTP son los responsables del mantenimiento de los datos, ya sea agregando datos, realizando actualizaciones o bien eliminándolos.

Las estructuras de datos deben estar optimizadas para validar la entrada de los mismos, y rechazarlos si no cumplen con determinadas reglas de negocio.

Para la toma de decisiones, proporciona capacidades limitadas ya que no es su objetivo, por lo tanto, no es prioridad en su diseño. Si se quisiera obtener determinada información histórica relativa al negocio consultando un sistema OLTP, se produciría un impacto negativo en el funcionamiento del sistema.

2.3.2. OLAP- OnLine Analytical Processing

La tecnología OLAP es una forma específica para representar datos financieros, operacionales, comerciales y estadísticos orientados a los ejecutivos, especialistas y analistas. Está diseñada para ayudar a la toma de decisiones y una mejor comprensión de la información. La idea central es poder contestar las preguntas de los usuarios, de una forma fácil, poderosa e intuitiva. Un sistema OLAP permite a los usuarios entrar en detalles y generalizar, filtrar, ordenar, rankear y reagrupar datos, calculándose totales intermediarios y finales en forma instantánea.

La tecnología OLAP permite un uso más eficaz de los almacenes de datos para el análisis en línea, lo que proporciona respuestas rápidas a consultas analíticas complejas e iterativas. Los modelos de datos multidimensionales de OLAP y las técnicas de agregados de datos organizan y resumen grandes cantidades de datos para que puedan ser evaluados con rapidez mediante el análisis en línea y las herramientas gráficas. Los sistemas OLAP proporcionan la velocidad y la flexibilidad necesarias para dar apoyo al analista en tiempo real. Cabe indicar que la tecnología OLAP tiene como base el proceso de transacciones en línea (OLTP).

Las siguientes son características que la tecnología OLAP posee:

Las bases de datos de OLAP tienen un esquema que está optimizado para que las preguntas realizadas por los usuarios sean respondidas rápidamente.

Las preguntas que se le hacen a un OLAP, deben permitir un uso interactivo con los usuarios.

Los cubos de OLAP almacenan varios niveles de datos conformados por estructuras altamente optimizadas que responden a las expectativas de negocio de la empresa.

Un sistema OLAP está preparado para realizar informes complejos de una manera simple.

OLAP proporciona una vista de datos multidimensional. Los cubos proporcionan una vista de los datos multidimensional que se extiende más allá del análisis de dos dimensiones que puede proporcionar una simple planilla de cálculo utilizada como tal.

Los usuarios pueden cambiar fácilmente las filas, las columnas, y las páginas en informes de OLAP, pudiendo leer la información de la manera que se crea más conveniente para el análisis.

2.3.3. OLTP vs OLAP

Tabla 3: Diferencias entre las tecnologías OLTP Y OLAP

	OLTP	OLAP
Objetivo Principal	Asistir a aplicaciones específicas, y mantener integridad de los datos.	Asistir en el análisis del negocio, identificando tendencias, comparando períodos, gestiones, mercados, índices, etc. mediante el almacenamiento de datos históricos
Alineación de Datos	Están alineados por aplicación. Diferentes sistemas tienen distintos tipos de datos, los cuales son estructurados por aplicación. Se focaliza en el cumplimiento de requerimientos de una aplicación en especial o una tarea específica.	Están alineados por dimensión. Todos los tipos de datos integrados en un solo sistema. Los datos son organizados definiendo dimensiones del negocio (áreas temáticas o sujetos). Se focaliza en el cumplimiento de requerimientos del análisis del negocio.
Integración de datos	Los datos se encuentran típicamente no integrados, son calificados como datos primitivos o datos operacionales. Los mismos son estructurados independientemente uno de otros, pudiendo tener diferentes estructuras de claves y convenciones de nombres. Son usualmente almacenados en diferentes formatos de archivos, relacional, VSAM, archivos planos, etc.	Los datos deben estar integrados. Son conocidos como datos derivados o datos DSS dado que provienen de sistemas transaccionales o sistemas de archivos maestros preexistentes en las mismas organizaciones o de sistemas externos de información. El DW, con el objetivo de alinear los datos por

	<p>Incluso, si todos los datos están en formato relacional, los mismos pueden residir en diferentes plataformas de hardware y en distintas RDBMSs.</p>	<p>áreas temáticas, debe integrar datos operacionales estandarizando estructuras y convenciones de nombres (concepto de diccionario de datos).</p>
<p>Historia</p>	<p>Usualmente retienen datos para 60 a 90 días, después son resguardados por los administradores de base de datos en almacenamientos secundarios fuera de línea (cintas o en disco a nivel de back up). También es común que contengan sólo valores corrientes, el actual balance de cuentas para clientes y no valores históricos. Puede no incluir el tiempo como un componente de la clave. Por ejemplo, sólo el balance corriente de cuentas es almacenado, por lo tanto, no tiene sentido guardar el tiempo como parte de la clave de los datos.</p>	<p>Almacenan tanta historia como sea necesario para el análisis del negocio, típicamente dos a cinco años de datos históricos. Retienen valores para cada período (el atributo más atómico de la dimensión tiempo) en la base de datos. Es decir, que almacenan una serie de fotos instantáneas de datos operacionales, la frecuencia con la cual define el nivel de detalle es la que se indica en la correspondiente hoja de la dimensión tiempo. Toda esta cantidad y tipo de historia apunta a ayudar a la generación de reportes de comparación de tendencias y períodos de tiempo. Por otro lado, las bases de datos</p>

	<p>orientadas al análisis siempre contienen el tiempo como clave dado que una de las principales razones para la construcción del data warehouse es el almacenamiento de datos históricos y el análisis a lo largo del tiempo.</p>
<p>Acceso y Manipulación de los datos</p> <p>Realizan una manipulación de datos registro por registro con grandes cantidades de inserts, updates y deletes. Además, necesitan de rutinas de validación y transacciones a nivel registro (OLTP on-line TRANSACTION processing). Generalmente, poseen pequeñas cantidades de datos involucrados en un solo proceso o transacción y la puesta a punto de la base de datos para el procesamiento de transacciones, se focaliza en mecanismos de locking y asignación de recursos (tuning específico)</p>	<p>Tienen una carga y acceso masivo de datos, no se realizan inserts, updates o deletes. La carga y refresco es batch (lo que se conoce como proceso BULK COPY). La validación de datos se realiza antes o después de la carga (nunca a nivel registro o transacción). Principalmente, se realizan sentencias de SELECT sobre varios registros y tablas (OLAP on-line ANALYTICAL processing), teniendo grandes volúmenes de datos involucrados en un único proceso o análisis. Es por ello, que generalmente, no se respetan las formas normales tan necesarias en los sistemas operacionales</p>

		<p>clásicos. Las anomalías que tienden a subsanar estas reglas de normalización no se presentan en los sistemas OLAP donde la carga de la información está automatizada y puede permitirse el manejo de redundancia controlada da como punto para la mejora de los tiempos de respuesta de las consultas a la base de datos.</p>
Patrones de uso	<p>Mantienen un patrón de uso constante requiriendo grandes cantidades de recursos y consumiendo sólo el tiempo referido a la transacción.</p>	<p>Tienen un patrón de uso liviano con picos de usos eventuales en el tiempo (afectados por la disponibilidad de los datos y el flujo de trabajo del negocio). Los picos de uso suceden el mismo día de cada semana y el mismo día de cada mes (cuando los datos están por primera vez disponibles o cuando el negocio necesita por primera vez un reporte).</p>
Granularidad de los datos	<p>Se encuentran los datos a nivel detallado o nivel transaccional. Una transacción incluye a nivel atómico cada uno de los componentes de su estructura (fecha, hora,</p>	<p>La granularidad de los datos viene dada por el uso de los mismos. Si bien un data warehouse puede tener información a nivel transaccional, el</p>

	código de cliente, código de movimiento, importe, etc.)	objetivo de esta granularidad mínima está asociado con el deseo de realizar ciertos tipos de análisis que requieren que la información esté a ese nivel de detalle, pero no significa que veamos la información a nivel transaccional.
Perfil de Usuario	El perfil de usuario que interactúa con dichos sistemas se encuadra dentro de los empleados operacionales de una organización (comunidad operativa).	El perfil de usuario sobre este tipo de sistemas corresponde a la comunidad gerencial, la cual está a cargo de la toma de decisiones.

Fuente: Elaboración propia

2.3.4. Sistema de soporte de decisiones (DSS)

Es un sistema interactivo provisto de programas y herramientas, para ayudar a los responsables de la toma de decisiones a utilizar tecnologías de comunicaciones, datos, documentos, conocimiento y/o modelos para identificar y resolver problemas, para completar tareas del proceso de decisión, y para tomar decisiones. Es una amplia área de análisis que sirve para que la gente examine datos a fin de tomar decisiones, ya sean grandes o pequeños, sobre los negocios de sus compañías.

2.3.5. Data Warehouse

Es un repositorio de información extraída de otros sistemas corporativos, sean estos sistemas transaccionales, bases de datos departamentales, o Intranet de la compañía, a la que los hombres de negocios de la empresa

pueden acceder. Los sistemas Data Warehouse están orientados a procesos de consultas en contraposición con los procesos transaccionales, sus tablas pueden no estar normalizadas y se admite redundancia en los datos. Mejor dicho, la Data Warehouse es un sistema, no un producto, en el que se almacenan datos. Es una técnica para consolidar y administrar datos de variadas fuentes con el propósito de responder preguntas de negocios y tomar decisiones, de una forma rápida. Una Data Warehouse se vale de una base de datos relacional diseñada para el acceso rápido y análisis y no al proceso transaccional. La Data Warehouse separa la carga del análisis y normalmente contiene datos históricos derivados de datos transaccionales.

Sus principales características son:

- Orientado hacia información relevante de la organización.
- Datos integrados.
- Variable en el tiempo.
- No Volátil.

2.3.6. Datamart

Es una Data Warehouse solo que más pequeña; en otras palabras, es una Data Warehouse orientada a algún tema. Los Datamart suelen ser usados por un departamento o grupo de usuarios en una compañía, para un conjunto definido de tareas. Una Datamart se considera independiente, ya que recibe datos desde un Data Warehouse. Los Datamart aislados, es decir los que toman sus datos directamente desde sistemas transaccionales y no dependen de otros Data Warehouse, recién el nombre de “Datamart Independientes”.

2.3.7. Dimensiones

Son los diferentes puntos de vista por los que queremos analizar la información. Las dimensiones contienen los diversos atributos que queremos analizar, además se estructuran en forma jerárquica, conforme a diferentes niveles de detalle. Las tablas de dimensiones se construyen con todos los atributos que incluyen de una forma des-normalizada y con una clave que identifica el mínimo nivel de detalle. Podemos distinguir varios tipos de dimensiones:

Dimensiones normales: aquellas que agrupan diferentes atributos que están relacionados por el ámbito al que se refieren (todas las características de un cliente, los diferentes componentes de la dimensión tiempo, etc)

Dimensiones causales: son atributos que pueden causar cambios en los procesos de negocio (por ejemplo, la dimensión promoción en el proceso de negocio de ventas).

Dimensiones heterogéneas: que agrupan conjuntos heterogéneos de atributos, que no están relacionados entre sí.

Dimensiones roll-up: es un subconjunto de otra, necesarias para el caso en que tenemos tablas de hechos con diferente granularidad (ver la entrada anterior del blog).

Dimensiones Junk: agrupa indicadores de baja cardinalidad como pueden ser flags o indicadores.

Dimensiones role-playing: cuando una misma dimensión interviene en una tabla de hechos varias veces (por ejemplo, la fecha en una tabla de hechos

donde se registran varias fechas referidas a conceptos diferentes), es necesario reutilizar la misma dimensión, pues no tiene sentido crear tantas dimensiones como usos se hagan de ella. En consecuencia, se definen las dimensiones role-playing. Podemos crear vistas sobre la tabla de la dimensión completa que nos permiten usarla varias veces o jugar con los alias de tabla. La misma dimensión juega un rol diferente según el sitio donde se utiliza.

Dimensiones degeneradas: no tienen ningún atributo y, por tanto, no tienen una tabla específica de dimensión. Incluyen para ellas un identificador en la tabla de hechos, que identifica completamente a la dimensión (por ejemplo, un pedido de ventas). Nos interesa tener determinada la transacción (para realizar data mining, por ejemplo), pero los datos interesantes de este elemento los tenemos repartidos en las diferentes dimensiones (cliente, producto, etc).

Mini dimensiones o dimensiones Outrigger: conjunto de atributos de una dimensión que se extraen de la tabla de dimensión principal, pues se suelen analizar de forma diferente. El típico ejemplo son los datos socio demográficos asociados a un cliente (que se utilizan, por ejemplo, para el datamining).

2.3.8. Métricas

Son valores que recogen el proceso de una actividad o los resultados de la misma y son usados por los analistas en sus query's para medir la performance del comportamiento de un proceso o un objeto del negocio. Las medidas candidatas son los datos numéricos, pero no cada atributo numérico en una

medida candidata. Estas medidas proceden del resultado de la actividad de negocio.

Métricas de realización de actividad: miden la realización de una actividad. Por ejemplo, la participación de una persona en un evento.

Métricas de resultado de una actividad: recogen los resultados de una actividad. Por ejemplo, la cantidad de unidades vendidas.

2.3.9. Indicadores

Es una variable dimensional unitaria, expresada como un cociente, que correlaciona dos variables cualesquiera. A través de este concepto es posible relacionar diversas variables presentes en los procesos de una empresa. Ejemplo: N° Ciclos/Hora, $LtsComb$, $/Maquina$. La importancia de un indicador radica en la particularidad de informar al usuario el estado actual del sistema. Revelar el estado del sistema, nos permitirá tomar decisiones preventivas o correctivas de acuerdo a los resultados de las comparaciones entre el valor esperado y el valor obtenido del sistema.

2.3.10. Fact Table

Los hechos son los indicadores de negocio que dan sentido al análisis de las dimensiones. Las tablas de hechos incluyen los indicadores asociados a un proceso de negocio en concreto y las claves de las dimensiones que intervienen en dicho proceso, en el mínimo nivel de granularidad o detalle. Podemos tener varios tipos de tablas de hechos:

Transaction fact tables: representan eventos que suceden en un determinado espacio-tiempo. Se caracterizan por analizar los datos con el máximo detalle.

Reflejan las transacciones relacionadas con nuestros procesos de negocio (ventas, compras, inventario, contabilidad, etc.).

Factless fact tables: no tienen medidas y representan la ocurrencia de un evento determinado. Por ejemplo, la asistencia a un curso puede ser una tabla de hechos sin métricas asociadas.

Periodic snapshot fact tables: son tablas de hecho usadas para recoger información de forma periódica a intervalos de tiempo regulares sobre un hecho. Nos permiten tomar una foto de la situación en un momento establecido (por ejemplo, al final del día, de una semana o de un mes). Un ejemplo puede ser la foto del stock de materiales, al final de cada día.

Accumulating snapshot fact table: representan el ciclo de vida completo de una actividad o proceso, que tiene un principio y final. Suelen representar valores acumulados.

Consolidated fact tables: tablas de hechos construidas como la acumulación, en un nivel de granularidad o detalle diferente, de las tablas de hechos de transacciones.

2.3.11. Esquema Estrella.

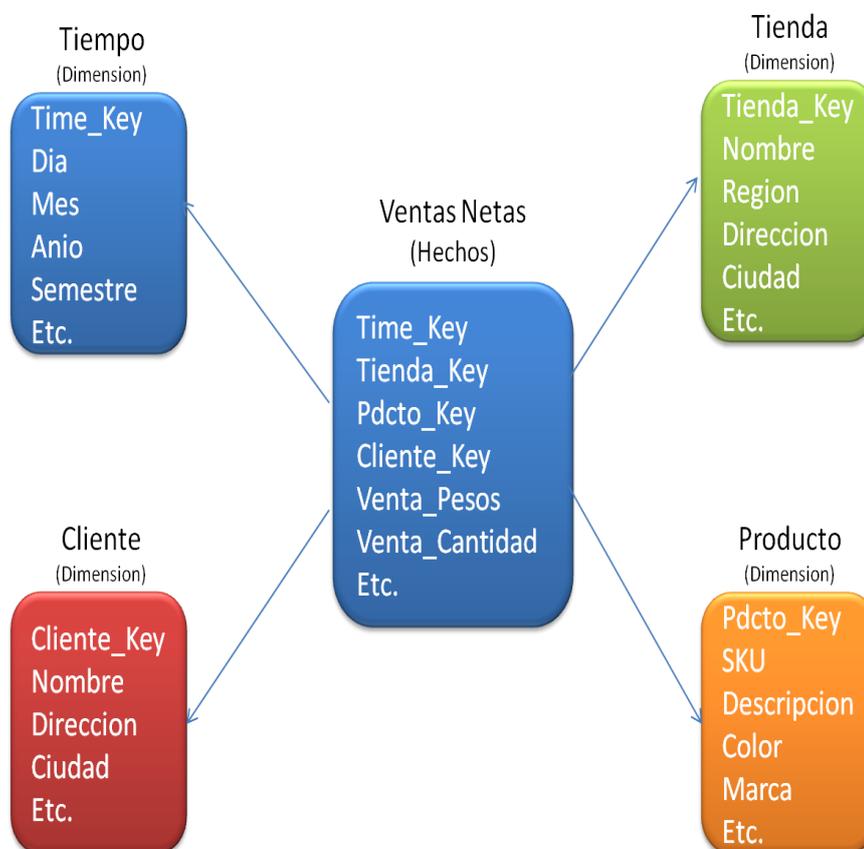


Figura 8: Esquema Estrella

Un modelo estrella es un modo de representar datos multidimensionales en una base de datos relacional. Las tablas dimensiones guardan información descriptiva acerca de sus miembros y sus relaciones. Mientras, que las tablas hechas almacenan datos de negocio. Los query's multidimensionales son hechos haciendo joins entre las tablas de hechos y de dimensiones Los sistemas OLAP que utilizan este tipo de modelo de base de datos son llamados sistemas ROLAP

2.3.12. Esquema Copo De Nieve

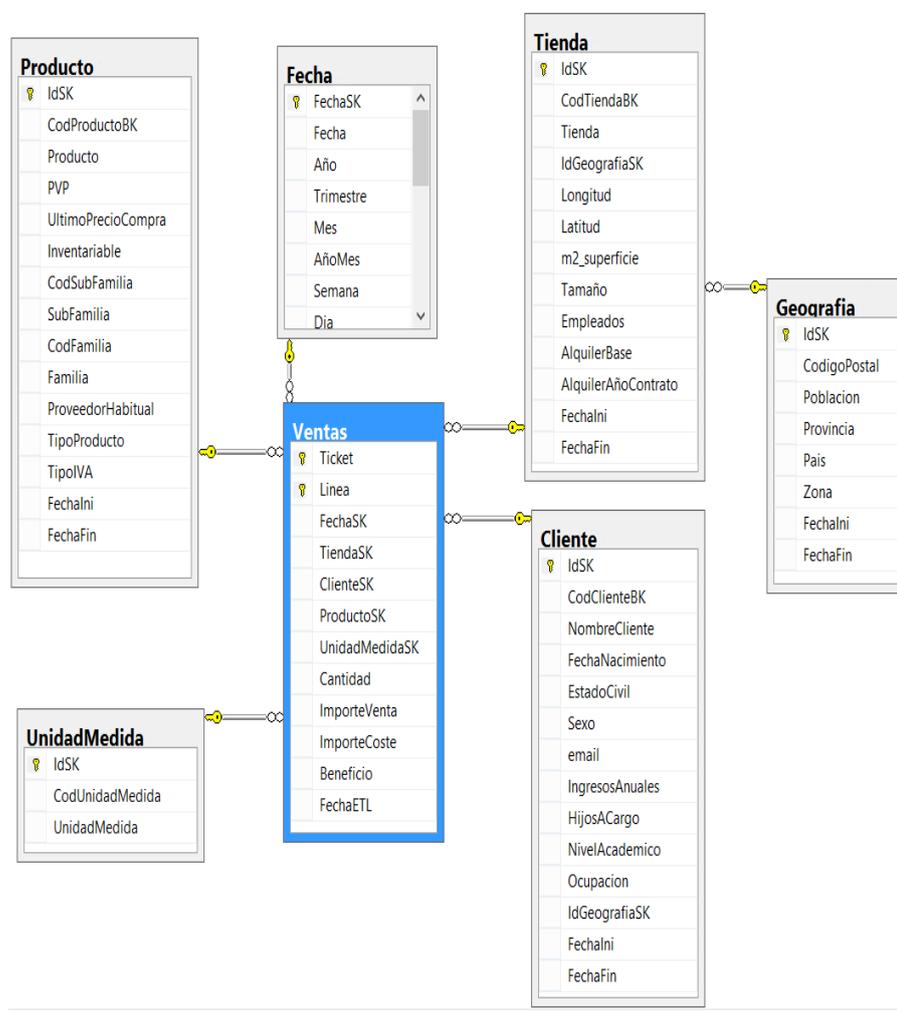


Figura 9: Esquema copo de nieve

Las bases de datos relacionales, a menudo emplean esquemas de copo de nieve para proporcionar los mejores tiempos de respuesta posibles, a las consultas complejas. Los esquemas de copos de nieve contienen una tabla de hechos centrales para el tema y numerosas tablas de dimensiones para la información descriptiva sobre las dimensiones del tema. La tabla hechos puede contener varios millones de filas. La información a la que se tiene acceso con más frecuencia se agrega previamente y se resume para mejorar

aún más, el rendimiento. Si bien el esquema copo de nieve se considera fundamentalmente una herramienta con la que el administrador de datos puede aumentar el rendimiento y simplificar el diseño del almacén de datos, también se utiliza para representar la información del almacén de datos, de modo tal, que tenga más sentido para los usuarios finales.

Suele aplicarse cuando diversos atributos caracterizan a los niveles más altos de jerarquía. Se puede elegir el normalizar solo algunas dimensiones y otras no, aumentando así, la complejidad del diseño y metadatos.

Ventajas:

No se necesita el atributo de nivel.

Fácil para definir jerarquías.

Mejora considerablemente, el rendimiento cuando un gran número de requisitos solicita datos agregados.

Los requerimientos escanean un reducido número de filas

Desventajas:

Aumenta la complejidad de mantener la meta información debido al aumento del número de tablas.

2.3.13. Enfoque “Bottom-up”

El enfoque Bottom-up, establece que el desarrollo de una solución de inteligencia de negocios debe partir por el desarrollo de los Data Marts y que los conjuntos de estos se constituyen en la Data Warehouse.

2.3.14. Enfoque “Top-down”

El enfoque Top-down, establece que el desarrollo de una solución de inteligencia de negocios debe partir por el desarrollo del Data Warehouse que deriva en la Data Marts.

2.3.15. Cubos OLAP

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.

A menudo se pensaba que todo lo que los usuarios pueden querer de un sistema de información se podría hacer de una base de datos relacional. No obstante, Codd fue uno de los precursores de las bases de datos relacionales, por lo que sus opiniones fueron y son respetadas

2.4 OPERACIONALIZACIÓN DE VARIABLES

Tabla 4: Operacionalización de variables

	VARIABLE INDEPENDIENTE		
	DIMENSIONES	INDICADORES	INSTRUMENTOS
Desarrollo del Datamart	Análisis de requerimientos	Satisfacción de los requerimientos	Nº de requerimientos identificados
		Estética	Grado de Atracción

Diseño del Datamart	Usabilidad	Facilidad de operación	
	Funcional	Facilidad de navegación	
Implementación del Datamart	Eficiencia para el mantenimiento y actualización de los procesos ETL	Facilidad de aprender	
	Eficiencia al crear consultas y cuadros dimensionales OLAP	Nivel de precisión	
	Disponibilidad de información que genera el sistema	Nº de usuarios satisfechos	
Evaluación del Datamart por el usuario	Aceptación del Datamart por parte de los usuarios	Nº de usuarios satisfechos	
VARIABLE DEPENDIENTE			
	DIMENSIONES	INDICADORES	INSTRUMENTOS
Eficiencia en la toma de decisiones en el supermercado Mia Market	Tiempo administrado en la toma de decisiones.	Comparación del tiempo en la cual se tomaban las decisiones	Entrevista a partir del uso del sistema de inteligencia de negocios
	Apoyo en la toma de decisiones	Análisis de datos y modelado de decisiones	Encuestas estadísticas
	Validación del sistema	Ayuda en el proceso de toma de decisiones	Encuestas estadísticas

Fuente: Elaboración propia

CAPITULO III

MATERIALES Y MÉTODOS

3.1 LOCALIZACIÓN DEL PROYECTO.

Supermercados Mía Market se encuentra ubicado en el distrito de Juliaca, Provincia de San Román, Departamento de Puno. Con dirección Jr. Jorge Chávez # 288 / Unión # 100 – Cercado.



Figura 10: Ubicación del establecimiento

3.2 MATERIAL DE ANÁLISIS

En el presente trabajo de investigación se tomó como material de análisis los reportes de la base de datos extraídos de SQL Server de la empresa mencionada; facilitados por el departamento de Informática registrados semanalmente durante el último año 2017.

3.2.1. Entrevistas

Para la investigación se pretende indagar sobre la optimización del proceso de toma de decisiones, para lo cual se usará la guía de entrevista del anexo 03, donde se analiza el uso y los beneficios de los procesos en la toma de decisiones.

3.2.2. Guía de observaciones

Como material de investigación usaremos las guías de observación, que estarán dirigidas a la gerencia y departamento de informática, quienes serán los usuarios del modelo de inteligencia de negocios. Se aplicará la guía de observación del anexo 4, con el fin de conocer como es la interacción entre el modelo de inteligencia de negocios y la gerencia.

3.3 MÉTODO DE RECOLECCIONES DATOS.

La base de datos fue extraída directamente de los archivos de la oficina de Informática, previa solicitud dirigida al Gerente y la autorización conjunta por motivos de seguridad y auditoria

3.3.1. Entrevistas

Las entrevistas a realizar, serán de tipo personalizada, acudiendo a las oficinas de los responsables de la toma de decisiones.

3.3.2. Guía de observaciones

Para la aplicación de este material, observaremos las reuniones de la alta dirección y departamento de informática para evaluar cuál es su sistema de trabajo.

3.4 POBLACIÓN

La población está conformada por todo el personal del supermercado Mia Market SRL - Juliaca.

3.5 DISEÑO DE LA MUESTRA

Tipo de muestra: Muestreo estratificado simple, para efectos de prueba se escogió a 9 encargados y directivos del área de gerencia de Mia Market SRL – Juliaca.

3.6 MÉTODO

Por lo cual el presente proyecto corresponde a la metodología de investigación cuantitativa debido a su naturaleza que presenta y ser desarrollado en forma objetiva.

3.7 MÉTODO DE ANÁLISIS DE DATOS

Para la recolección de datos se utilizaron los siguientes métodos y técnicas:

3.7.1 La entrevista

Las entrevistas implican que una persona calificada (entrevistador) aplica el cuestionario a los participantes; el primero hace las preguntas a cada entrevistado y anota las respuestas. Su papel es crucial, es una especie de filtro. (Hernández et al., 2010). Para lo cual se realizó un cuestionario previo que sirvió para guiar la entrevista cumpliendo el objetivo principal de la investigación, recabando información relevante para llevar a cabo la investigación. Las preguntas que se formularon están en el “Anexo A” y se detallará más adelante en la aplicación de la metodología la forma como se procedió.

3.7.2 Encuestas

Las encuestas fueron elaboradas por un conjunto de preguntas escritas por el investigador aplicadas a los involucrados en el desarrollo de la investigación, los cuales permitieron hacer las pruebas del sistema para su respectiva validación. El cuestionario que se utilizó para realizar estas encuestas está en el “Anexo H” y esto se desarrollará a detalle más adelante en la prueba de hipótesis.

3.8 METODO DE TRATAMIENTO DE DATOS

Para verificar la veracidad de la hipótesis, se aplicó la investigación experimental para medir el efecto de la variable independiente sobre la variable dependiente, teniendo la siguiente información:

- Recopilación de datos
- Construcción de tablas
- Interpretación de los resultados
- Prueba de hipótesis

La recolección de los datos para el análisis se realizó mediante dos cuestionarios uno aplicado antes de la construcción del sistema y el otro después de la construcción del sistema.

3.9 BASE DE DATOS TRANSACCIONAL EN SQL SERVER

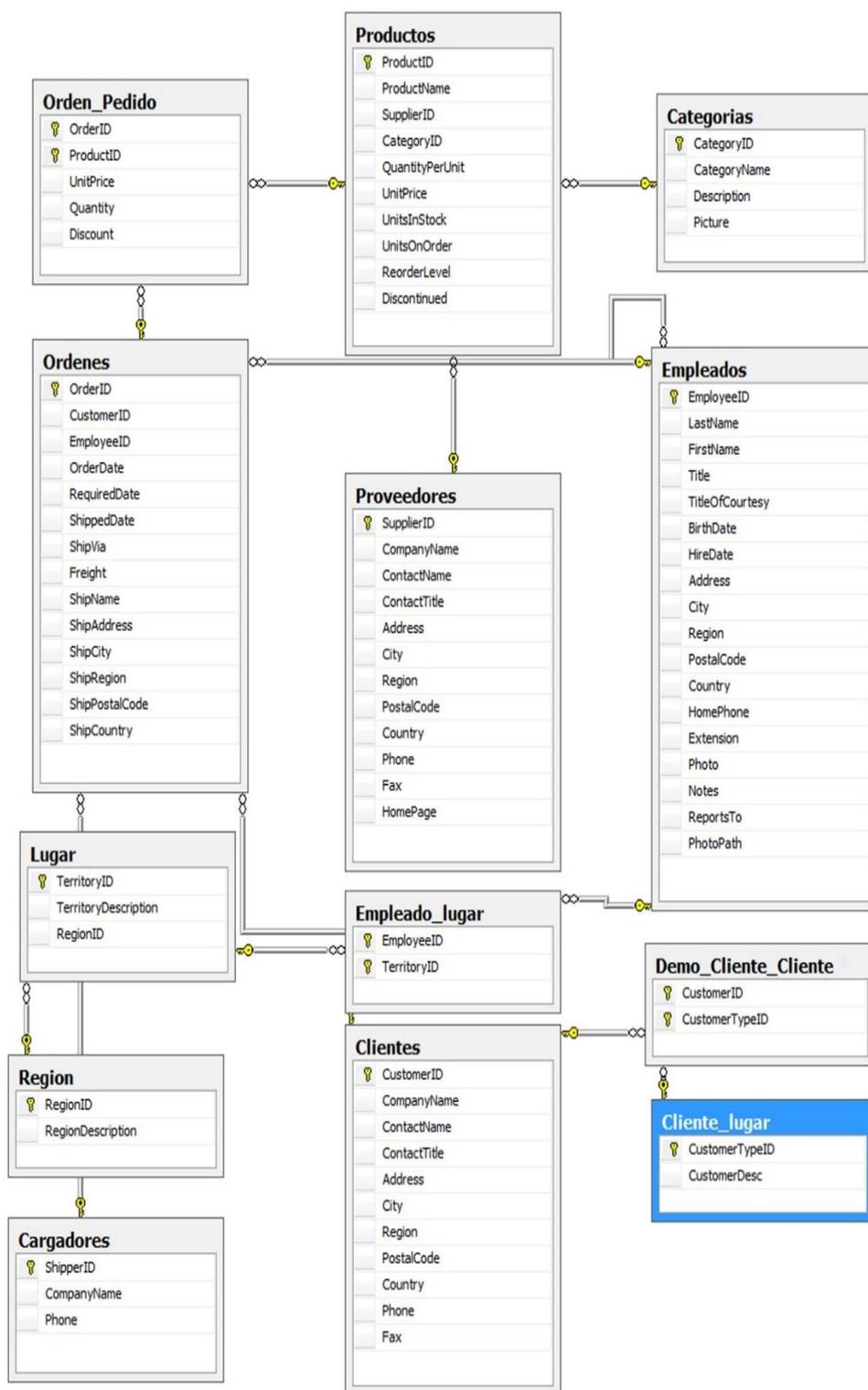


Figura 11: Base de datos transaccional en SQL Server

3.10 DICCIONARIO DE LA BASE DE DATOS GENERAL

Tabla 5: Descripción de la base de datos transaccional SQL Server

TABLA	DESCRIPCIÓN
Date	Se registra el día, semana, mes, trimestre, semestre y año de la base de datos
Productos	Se registran todos los productos que llegan de los proveedores solicitados para su inmediato registro en la base de datos
Categorías	Se registran el tipo de productos para tener una fácil selección para su posterior separación de acuerdo al tipo de producto
SalesTerritory	Se registran el orden de pedido de todos los productos
Proveedores	Se registran nuestros proveedores de acuerdo al tipo de producto solicitado con las fechas establecidas de entrega
Empleados	Se registran el número de personas que trabaja en la empresa con sus datos personales respectivamente
Geographics	Se registra el lugar de los productos dentro de la empresa ordenados por tipo de establecimiento
FactInternetSales	Se registra la cantidad de ventas
Currency	Se registra el tipo de moneda en este caso Soles moneda nacional
Region	Se registra el lugar del cliente, lugar donde vive que en este caso vendría ser domicilio
Clientes	Se registra la lista de clientes que adquirieron productos de nuestro supermercado

Fuente: Elaboración Propia.

3.11 BASE DE DATOS TRANSACCIONAL.

3.11.1 Descripción de la tabla DATE

Tabla 6: Descripción de la tabla DATE

Nombre columna	Tipo de dato	Null option	Descripción
DateKey	Int	not null	Número o clave de la fecha
FullDateAlternateKey	Date	not null	Número o clave de la fecha alternativa
DayNumberOfWeek	Tinyint	not null	Nombre del día de la semana

SpanishDayNameOfWeek	nvarchar (10)	not null	Nombre de la semana
DayNumberOfMonth	tinyint	not null	Día y nombre del mes
DayNumberOfYear	smallint	not null	Día y nombre del año
WeekNumberOfYear	tinyint	not null	Número de la semana y año
SpanishMonthName	nvarchar(10)	not null	Nombre del mes
MonthNumberOfYear	tinyint	not null	Mes nombre y año
CalendarQuarter	tinyint	not null	Trimestre
CalendarYear	smallint	not null	Año
CalendarSemester	tinyint	not null	Semestre
FiscalQuarter	tinyint	not null	Trimestre Fiscal
FiscalYear	smallint	not null	Año fiscal
FiscalSemester	tinyint	not null	Semestre fiscal

Fuente: Elaboración Propia

3.11.2 Descripción de la tabla PRODUCTOS

Tabla 7: Descripción de la tabla PRODUCTOS

Nombre columna	Tipo de dato	Null option	Descripción
Productid	Int	Not null	Número del producto
Productname	Nvarchar (40)	Not null	Nombre del producto
Supplierid	Int	Null	Número del proveedor
Categoryid	Int	Null	Número de categoría
Quantityperunit	Nvarchar (20)	Null	Precio unitario por cantidad
Unitprice	Money	Null	Precio unitario
Unitsinstock	Smallint	Null	Unidades en stock
Unitsonorder	Smallint	Null	Orden de unidades
Reorderlevel	Smallint	Null	Reordenar nivel
Discontinued	Bit	Not null	Descuento

Fuente: elaboración propia.

3.11.3 Descripción de la tabla categorías

Tabla 8: Descripción de la tabla categorías

Nombre columna	Tipo de dato	Null option	Descripción
CategoryID	int	not null	Número de categoría
CategoryName	nvarchar (15)	not null	Nombre de categoría
Description	Ntext	null	Descripción
Picture	image	null	Imágen

Fuente: Elaboración propia

3.11.4 Descripción de la tabla FactInternetSales

Tabla 9: Descripción de la tabla FACTINTERNETSALES

Nombre columna	Tipo de dato	Null option	Descripción
ProductKey	int	not null	Número o clave del producto
OrderDateKey	int	null	Número de la fecha de pedido
DueDateKey	int	null	Número de fecha de vencimiento
ShipDateKey	int	null	Número de fecha de envío
CustomerKey	int	null	Número del cliente
PromotionKey	int	null	Número de promoción
CurrencyKey	int	nul	Número de moneda
SalesTerritoryKey	int	null	Número del territorio de ventas
SalesOrderNumber	nvarchar(20)	null	Número de pedido de ventas
SalesOrderLineNumber	tinyint	null	Número de línea de orden de venta
RevisionNumber	tinyint	null	Número de revision

OrderQuantity	smallint	null	Cantidad de pedido
UnitPrice	money	null	Precio unitario
ExtendedAmount	money	null	Monto extendido
UnitPriceDiscountPct	float	null	Precio unitario de descuento
DiscountAmount	float	null	Importe de descuento
ProductStandardCost	money	null	Costo del producto estándar
TotalProductCost	money	null	Costo total de producción
SalesAmount	money	null	Cantidad de ventas
TaxAmt	money	null	Monto del impuesto
Freight	money	null	Carga
CarrierTrackingNumber	nvarchar (25)	not null	Número del seguimiento del operador
CustomerPONumber	nvarchar (25)	not null	Número del pedido del cliente

Fuente: Elaboración propia

3.11.5 Descripción de la tabla PROVEEDORES

Tabla 10: Descripción de la tabla PROVEEDORES

Nombre columna	Tipo de dato	Null option	Descripción
SupplierID	int	not null	Número de Proveedor
CompanyName	nvarchar (40)	not null	Nombre de la compañía
ContactName	nvarchar (30)	null	Nombre del contacto
ContactTitle	nvarchar (30)	null	Título del contacto
Address	nvarchar (60)	null	Dirección
City	nvarchar (15)	null	Ciudad
Region	nvarchar (15)	null	Región
PostalCode	nvarchar (10)	null	Número de Código

Country	nvarchar (15)	null	País
Phone	nvarchar (24)	null	Número Teléfono
Fax	nvarchar (24)	Null	Número de fax
HomePage	Ntext	null	Region de envío

Fuente: Elaboración propia

3.11.6 Descripción de la tabla EMPLEADOS

Tabla 11: Descripción de la tabla EMPLEADOS

Nombre columna	Tipo de dato	Null option	Descripción
EmployeeID	int	not null	Número de Empleado
LastName	nvarchar (20)	not null	Apellidos
FirstName	nvarchar (10)	not Null	Nombres
Title	nvarchar (30)	null	Título o cargo
TitleOfCourtesy	nvarchar (25)	null	Título o cargo de cortesía
BirthDate	datetime	null	Fecha de nacimiento
HireDate	datetime	nul	Fecha de contratación
Address	nvarchar (60)	null	Dirección
City	nvarchar (15)	null	Lugar o ciudad
Region	nvarchar (15)	null	Región
PostalCode	nvarchar (10)	null	DNI o código de verificación
Country	nvarchar (15)	null	País
HomePhone	nvarchar (24)	null	Número de Teléfono o celular
Extension	nvarchar (4)	null	Extensión de país si es extranjero
Photo	image	null	Fotografía
Notes	Ntext	null	Notas
ReportTo	Int	null	Reportar a
PhotoPath	nvarchar (255)	null	Ruta de la Fotografía

Fuente: Elaboración propia

3.11.7 Descripción de la tabla CURRENCY

Tabla 12: Descripción de la tabla CURRENCY

Nombre columna	Tipo de dato	Null option	Descripción
CurrencyKey	int	not null	Número código de la moneda

CurrencyAlternateKey	nchar (3)	not null	Código alternative
CurrencyName	nvarchar (50)	not null	Nombre de la moneda

Fuente: Elaboración propia

3.11.8 Descripción de la tabla REGION

Tabla 13: Descripción de la tabla REGION

Nombre columna	Tipo de dato	Null option	Descripción
RegionID	Int	Not null	Número ID de la región
RegionDescription	nchar (50)	Not null	Descripción de la region

Fuente: Elaboración Propia

3.11.9 Descripción de la tabla SalesTerritory

Tabla 14: Descripción de la tabla SALESTERRITORY

Nombre columna	Tipo de dato	Null option	Descripción
SalesTerritoryKey	int	null	Número del territorio de ventas
SalesTerritoryAlternateKey	int	not null	Número alternativo del territorio de ventas
SalesTerritoryRegion	nvarchar(50)	null	Region del territorio de ventas
SalesTerritoryCountry	nvarchar(50)	null	Pais del territorio de ventas
SalesTerritoryGroup	nvarchar(50)	not null	Gripo del territorio de ventas

Fuente: Elaboración Propia

3.11.10 Descripción de la tabla GEOGRAPHICS

Tabla 15: Descripción de la tabla GEOGRAPHICS

Nombre columna	Tipo de dato	Null option	Descripción
----------------	--------------	-------------	-------------

GeographyKey	int	not null	Clave geográfica
City	nvarchar(30)	null	Ciudad
StateProvinceCode	nvarchar (3)	null	Código de la provincia
StateProvinceName	nvarchar(50)	null	Nombre de la provincia
CountryRegionCode	nvarchar(3)	null	Código de la región país
SpanishCountryRegionName	nvarchar(50)	null	Nombre de la región y país
PostalCode	nvarchar(15)	null	Código postal
SalesTerritoryKey	int	null	Clave de ventas en el territorio

Fuente: Elaboración Propia

3.11.11 Descripción de la tabla CLIENTES

Tabla 16: Descripción de la tabla CLIENTES

Nombre columna	Tipo de dato	Null option	Descripción
CustomerID	nchar (5)	not null	Número de cliente
CompanyName	nvarchar (40)	not null	Nombre de la compañía
ContactName	nvarchar (30)	null	Nombre del contacto
ContactTitle	nvarchar (30)	null	Título o cargo del contacto
Address	nvarchar (60)	null	Dirección
City	nvarchar (15)	null	Ciudad
Region	nvarchar (15)	null	Región
PostalCode	nvarchar (10)	null	DNI o código de verificación
Country	nvarchar (15)	null	País
Phone	nvarchar (24)	null	Teléfono o celular
Fax	nvarchar (24)	null	Número de fax

Fuente: Elaboración Propia

3.12 MODELADO DIMENSIONAL

Definir una vista de origen de datos en un proyecto de “Analysis Services”

3.12.1. Crear Un Proyecto De Analysis Services

Ejecutar SQL Server Business Intelligence Development Studio ubicado en inicio, todos los programas, Microsoft SQL Server 2008 (esta ubicación puede variar dependiendo de la versión de Windows). Generar un nuevo Proyecto de Analysis Services al cual se le llamará, “Supermercado_Mia_Market”. (Archivo, Nuevo, Proyecto...)

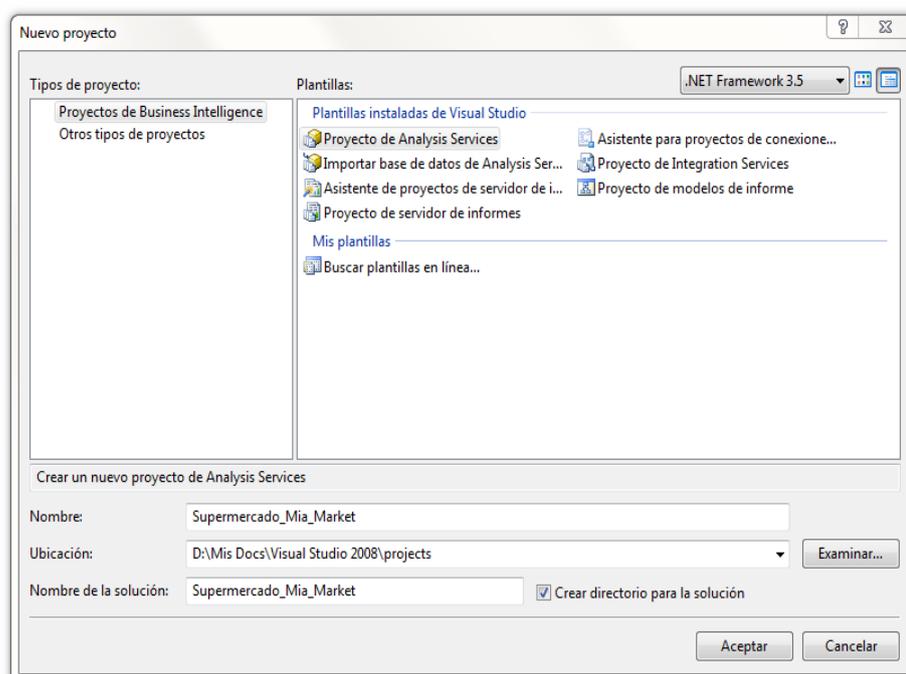


Figura 12: Crear nuevo proyecto en Analysis Service

Hacer clic en Aceptar. Ahora ya está creado el Proyecto de Analysis Services.

3.12.2. Definir Un Origen De Datos

Hacer clic con el botón derecho sobre “Orígenes de Datos” y seguidamente en “Nuevo Origen de Datos”.



Figura 13: Definir un nuevo origen de datos en Analysis Service

Seleccionar la base de datos de donde se leerán los datos, en este caso “Supermercado_Mia_Market”.

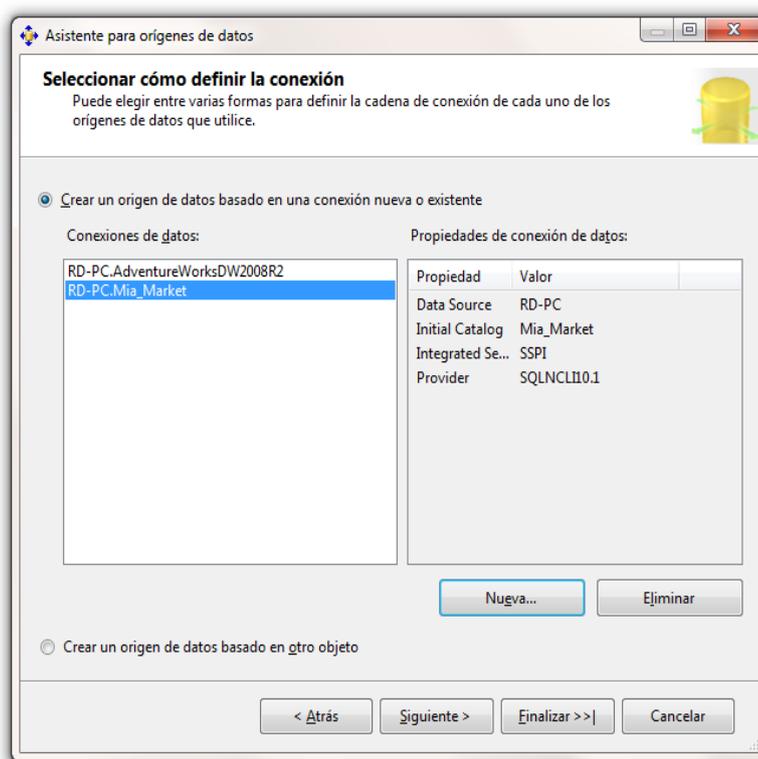


Figura 14: Definir la conexión con la base de datos

Proporcionar los datos con los cuales se conectará a la base de datos (en este caso se utilizará la cuenta de servicio)

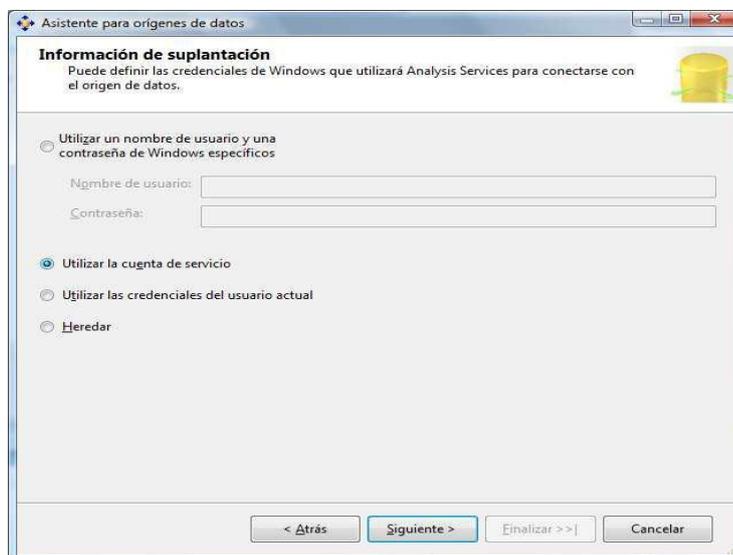


Figura 15: Asistente de origen de datos

Asignar un nombre al origen de datos (Supermercado_Mia_Market) y finalizar.

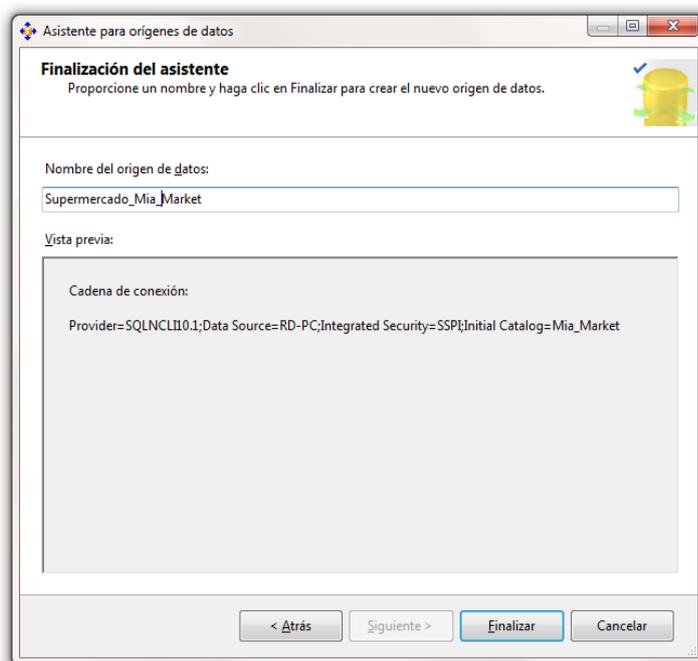


Figura 16: Finalización del Asistente de origen de datos

Con esto ya se encuentra listo el origen de datos.

3.12.3 Definir Una Vista De Origen De Datos

Realizar clic con el botón derecho sobre “Vistas del Origen de Datos” y a continuación clic sobre “Nueva Vista del Origen de Datos”.

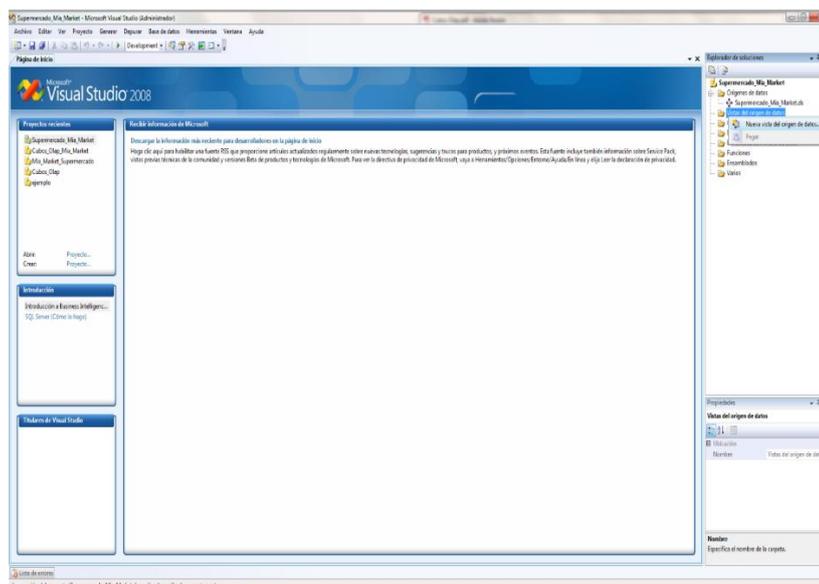


Figura 17: Definir Vista de origen de datos

Seleccionar el origen de datos (creado con anterioridad) y realizar clic en siguiente.

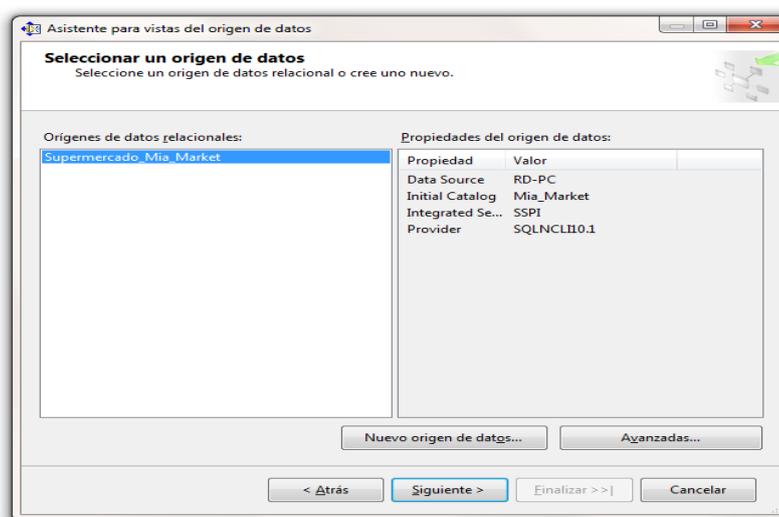


Figura 18: Asistente para vistas de origen de datos

En la lista Objetos disponibles, seleccione los siguientes objetos. Para seleccionar varias tablas, haga clic en cada una de ellas mientras mantiene presionada la tecla CTRL:

DimCustomer (dbo)

DimGeographic (dbo)

DimProduct (dbo)

DimDate (dbo)

DimInternetSales (dbo)

Una vez realizado esto hacer clic en siguiente. Asignar un nombre a la vista del origen de datos (AdventureWorksDW2008) y finalizar.

La vista del origen de datos Spermercado_Mia_Market aparece en la carpeta Vistas del origen de datos del Explorador de soluciones. El contenido de la vista del origen de datos también se muestra en el Diseñador de vistas del origen de datos de Business Intelligence Development Studio. Este diseñador contiene los elementos siguientes:

El panel Diagrama, en el que las tablas y sus relaciones se representan gráficamente.

El panel Tablas, en el que las tablas y los elementos de esquema se muestran en una vista de árbol.

El panel Organizador de diagramas, en el que puede crear subdiagramas de modo que pueda ver los subconjuntos de la vista de origen de datos.

Una barra de herramientas específica del Diseñador de vistas de origen de datos.

3.12.4. Definir e Implementar Un Cubo

Definir una dimensión, En la tarea siguiente, usará el Asistente para dimensiones con objeto de generar una dimensión de tiempo. Para definir una dimensión En el Explorador de soluciones, haga clic con el botón secundario en Dimensiones y, a continuación, en Nueva dimensión. En la página Asistente para dimensiones, haga clic en Siguiente. En la página Seleccionar método de creación, compruebe que está seleccionada la opción Usar una tabla existente y, a continuación, haga clic en Siguiente. En la página Especificar información de origen, compruebe que está seleccionada la vista del origen de datos Supermercado_Mia_Market. En la lista Tabla principal, seleccione Date. Haga clic en Siguiente.

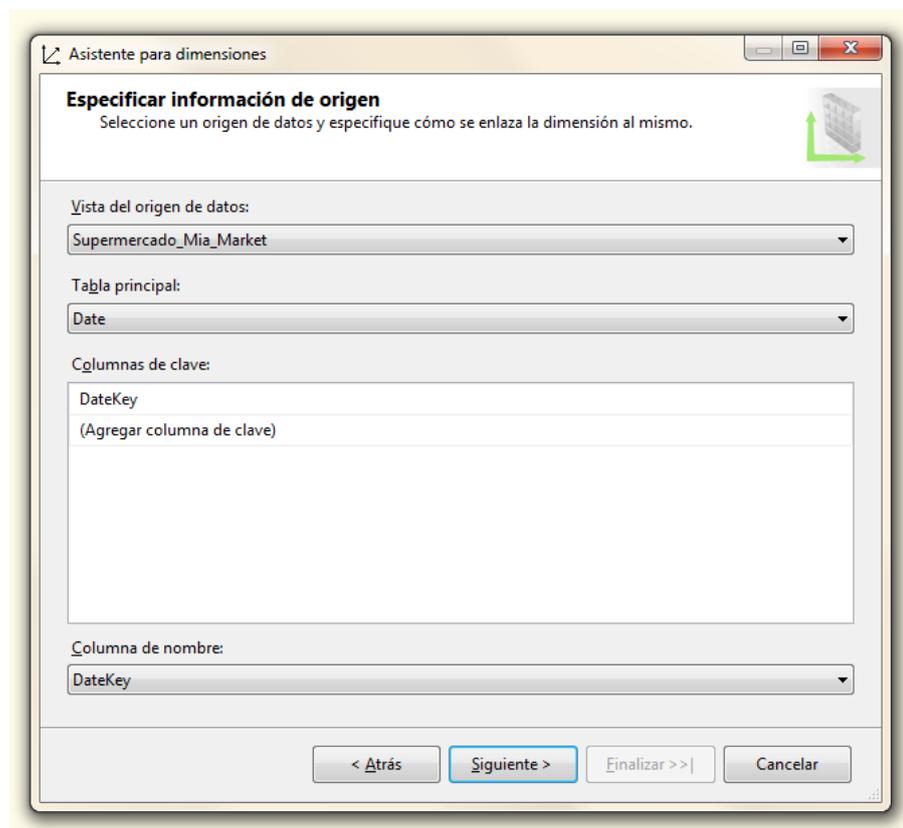


Figura 19: Asistente de dimensiones

En la página Seleccionar los atributos de la dimensión, seleccione las casillas situadas junto a los siguientes atributos:

DateKey

Full Date Alternate Key

Spanish Month Name

Calendar Quarter

Calendar Year

Calendar Semester

Cambie el valor de la columna Tipo de atributo del atributo Full Date Alternate Key de Regular a Date. Para ello, haga clic en Regular en la columna Tipo de atributo. A continuación, haga clic en la flecha para expandir las opciones. Por último, haga clic en fecha > Calendario > Fecha. Haga clic en Aceptar. Repita estos pasos para cambiar el tipo de atributo de los siguientes atributos como se indica a continuación:

Spanish Month Name a Mes

Calendar Quarter a Trimestre Fiscal

Calendar Year a Año

Calendar Semester a Semestre

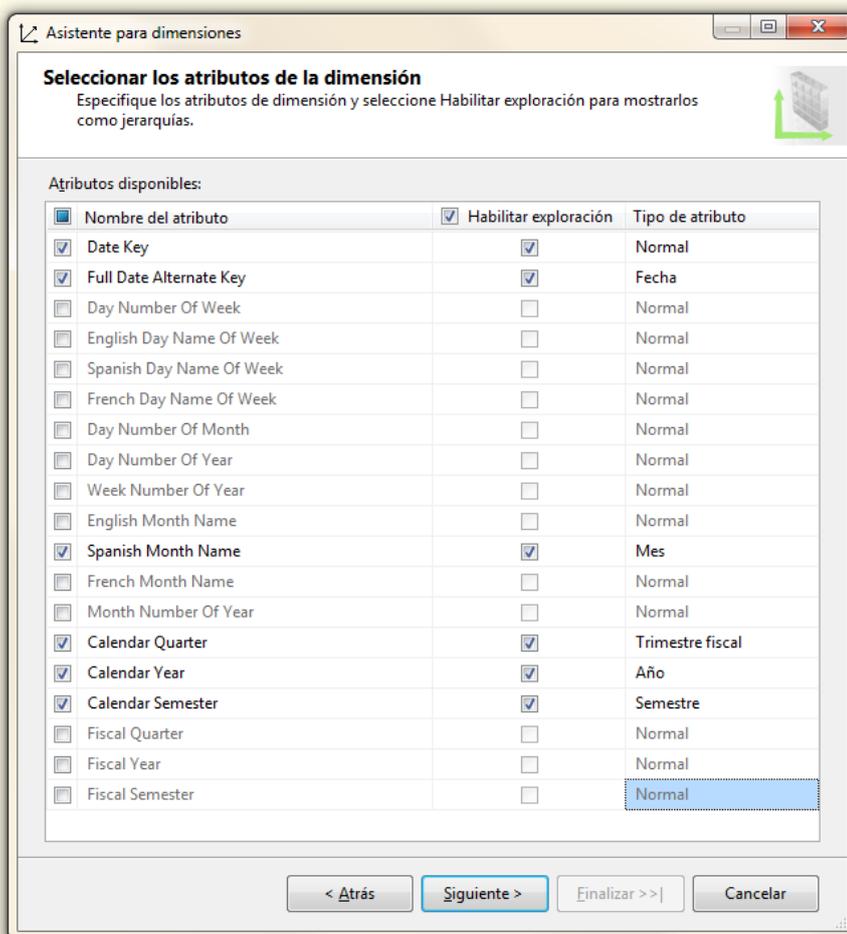


Figura 20: Selección de dimensiones de la tabla DATE

3.12.5 Definir Un Cubo

El Asistente para cubos le ayuda a definir los grupos de medida y las dimensiones de un cubo. En la tarea siguiente, usará el Asistente para cubos para generar un cubo. Para definir un cubo y sus propiedades, en el Explorador de soluciones, haga clic con el botón secundario en Cubos y, a continuación, haga clic en Nuevo cubo.

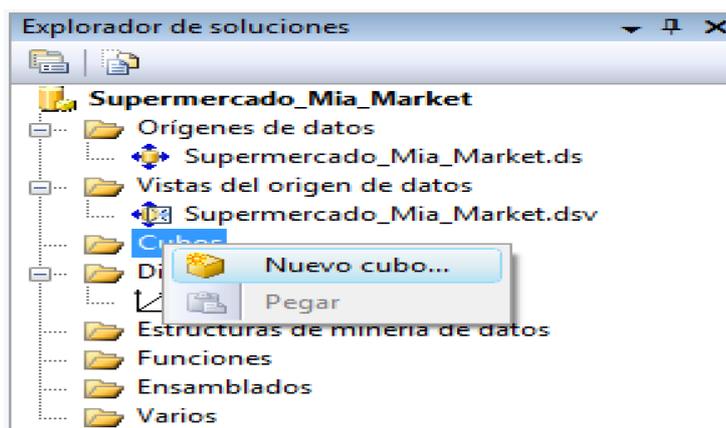


Figura 21: Explorador de soluciones para nuevo cubo

En la página Asistente para cubos, haga clic en Siguiente.

En la página Seleccionar método de creación, compruebe que la opción Usar tablas existentes está seleccionada y, a continuación, haga clic en Siguiente.

En la página Seleccionar tablas de grupo de medida, compruebe que la vista del origen de datos Supermercado_Mia_Market está seleccionada. Haga clic en Sugerir para que el Asistente para cubos sugiera las tablas que se deben usar para crear los grupos de medida. El asistente examinará las tablas y sugerirá Ventas como tabla de grupos de medida. Las tablas de grupos de medida, también denominadas tablas de hechos, contienen las medidas que son de su interés, como el número de unidades vendidas.

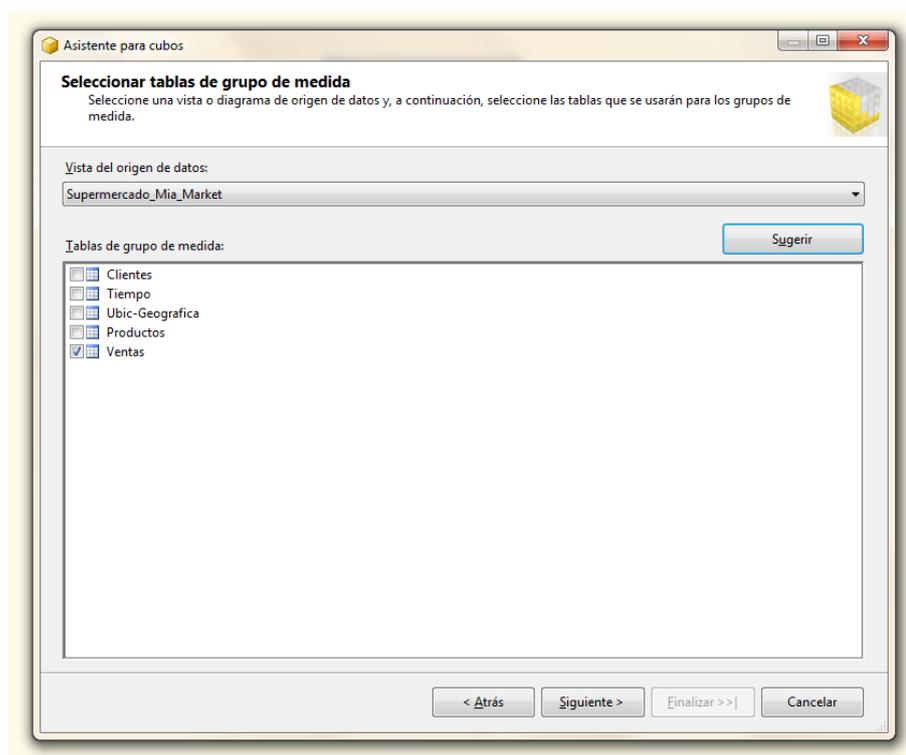


Figura 22: Asistente para definir cubos

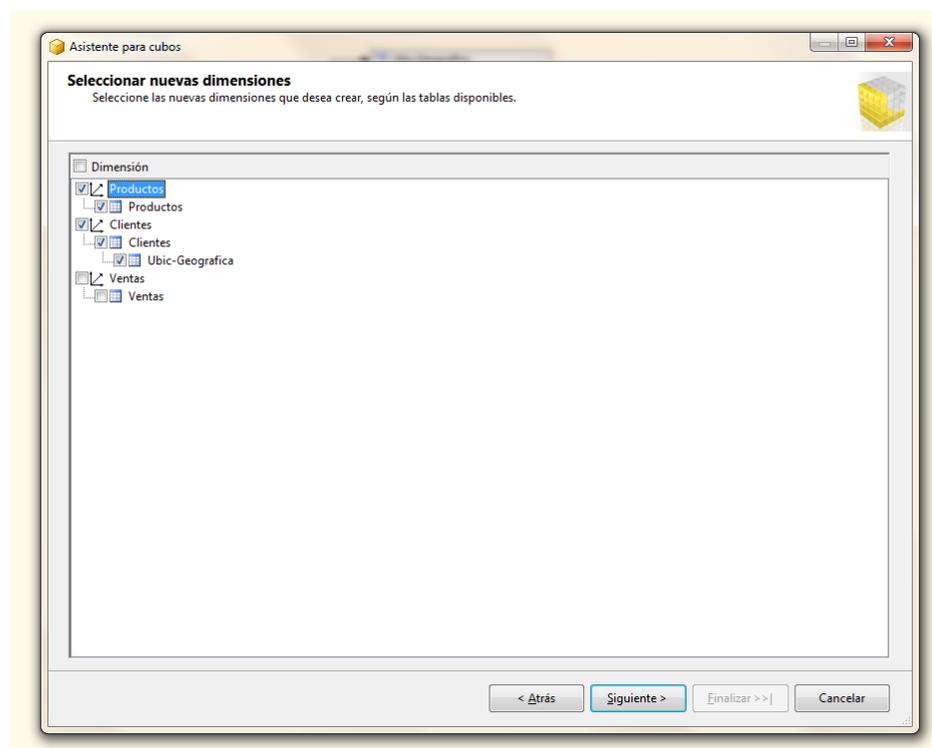


Figura 23: Asistente para definir cubos seleccionando nuevas dimensiones

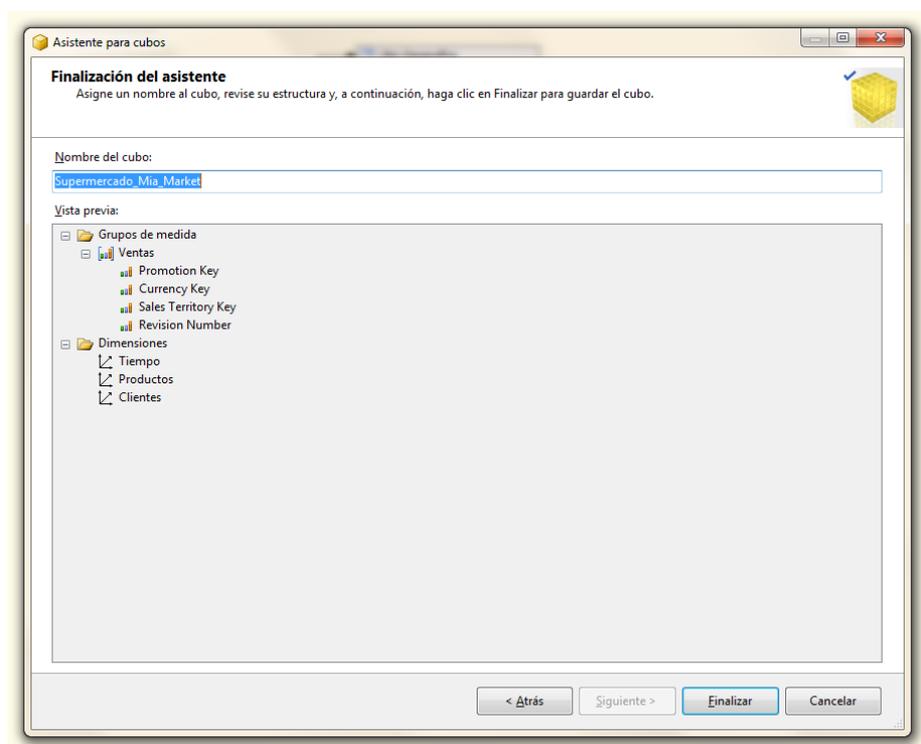


Figura 24: Finalizar el asistente renombrando el cubo

Haga clic en Finalizar para completar el asistente. En el Explorador de soluciones, en el proyecto “Supermercado_Mia_market” aparece en la carpeta Cubos, y las dimensiones de base de datos aparecen en la carpeta Dimensiones. Adicionalmente, en el centro del entorno de desarrollo, el Diseñador de cubos muestra el cubo “Supermercado_Mia_Market”.

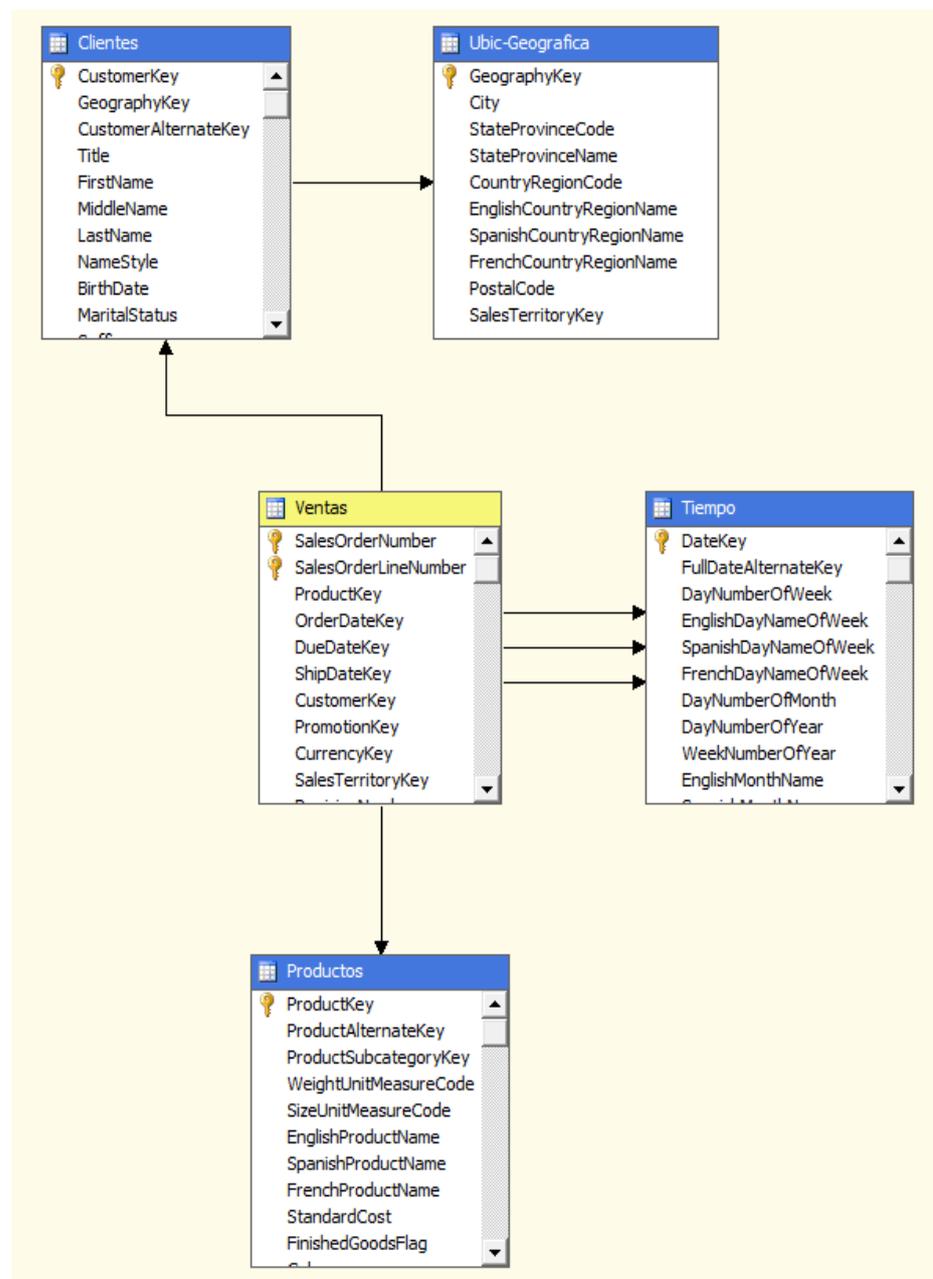


Figura 25: Esquema con implementación completa

3.12.6. Agregar Atributos a Dimensiones

En las tareas siguientes, usará el Diseñador de dimensiones para agregar atributos a las dimensiones Clientes y Productos. Agregar atributos a la dimensión Clientes

Para agregar atributos Abra el Diseñador de dimensiones para la dimensión Clientes. Para ello, haga doble clic en la dimensión Clientes del nodo Dimensiones del Explorador de soluciones. En el panel Atributos, observe los atributos Customer Key y Geography Key creados mediante el Asistente para cubos. En la barra de herramientas de la ficha Estructura de dimensión, utilice el icono Zoom para ver las tablas del panel Vista de origen de datos al 100 por cien. Arrastre las columnas siguientes de la tabla Clientes del panel Vista de origen de datos al panel Atributos:

<i>BirthDate</i>	<i>SpanishEducation</i>
<i>MaritalStatus</i>	<i>SpanishOccupation</i>
<i>Gender</i>	<i>HouseOwnerFlag</i>
<i>EmailAddress</i>	<i>NumberCarsOwned</i>
<i>YearlyIncome</i>	<i>Phone</i>
<i>TotalChildren</i>	<i>DateFirstPurchase</i>
<i>NumberChildrenAtHome</i>	<i>CommuteDistance</i>

Arrastre las columnas siguientes de la tabla Ubic. Geografica del panel Vista de origen de datos al panel Atributos:

<i>City</i>
<i>StateProvinceName</i>
<i>SpanishCountryRegionName</i>
<i>PostalCode</i>

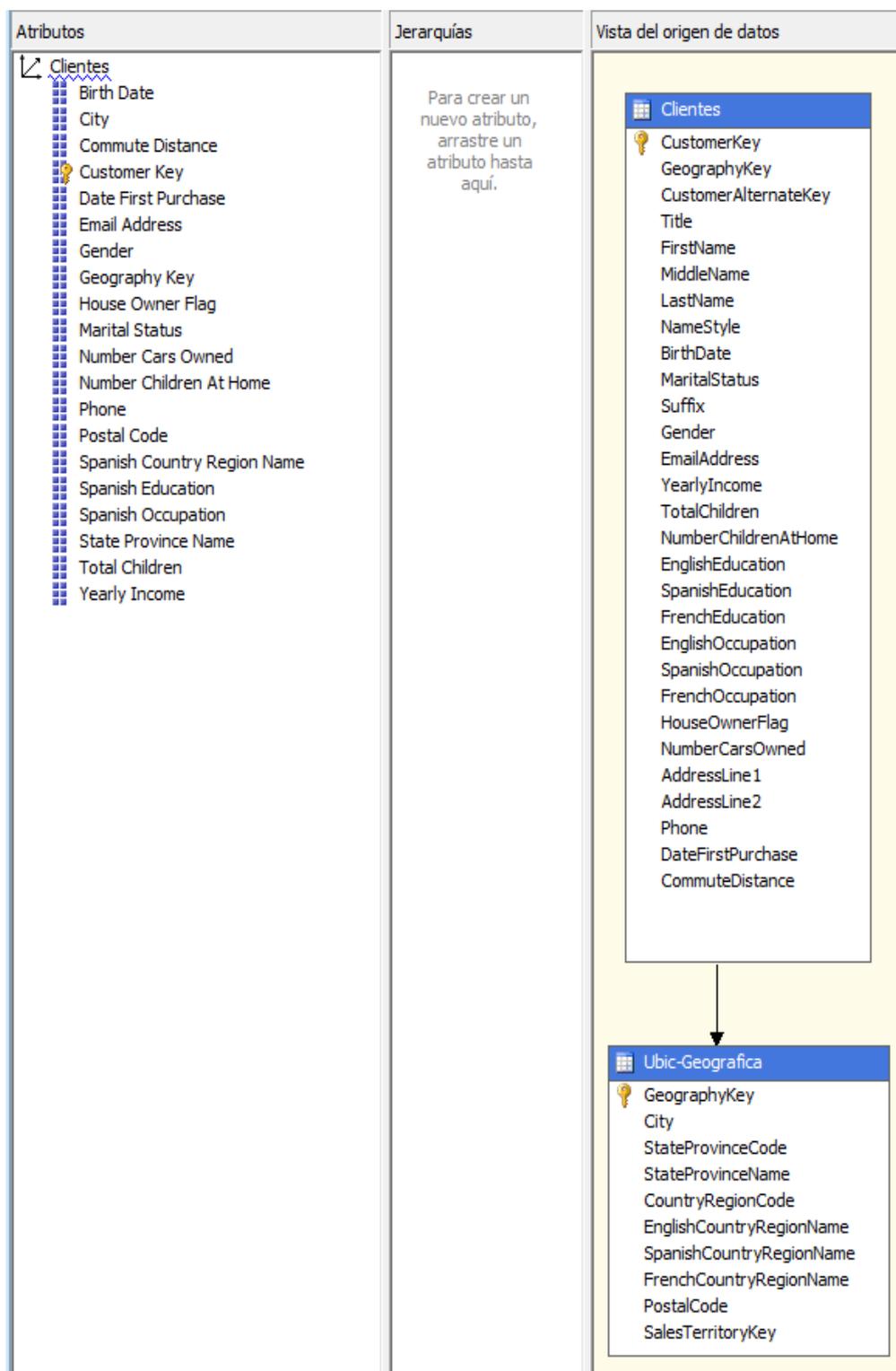


Figura 26: Implementación de la dimensión clientes

En el menú Archivo, haga clic en Guardar todo.

3.12.7. Agregar atributos a la dimensión Productos

Para agregar atributos Abra el Diseñador de dimensiones para la dimensión Productos. En el panel Atributos, observe el atributo Product Key creado mediante el Asistente para cubos. En la barra de herramientas de la ficha Estructura de dimensión, utilice el icono Zoom para ver las tablas del panel Vista de origen de datos al 100 por cien. Arrastre las columnas siguientes de la tabla Productos del panel Vista de origen de datos al panel Atributos:

<i>StandardCost</i>	<i>DaysToManufacture</i>
<i>Color</i>	<i>ProductLine</i>
<i>SafetyStockLevel</i>	<i>DealerPrice</i>
<i>ReorderPoint</i>	<i>Class</i>
<i>ListPrice</i>	<i>Style</i>
<i>Size</i>	<i>ModelName</i>
<i>SizeRange</i>	<i>StartDate</i>
<i>Weight</i>	<i>EndDate</i>
<i>Status</i>	

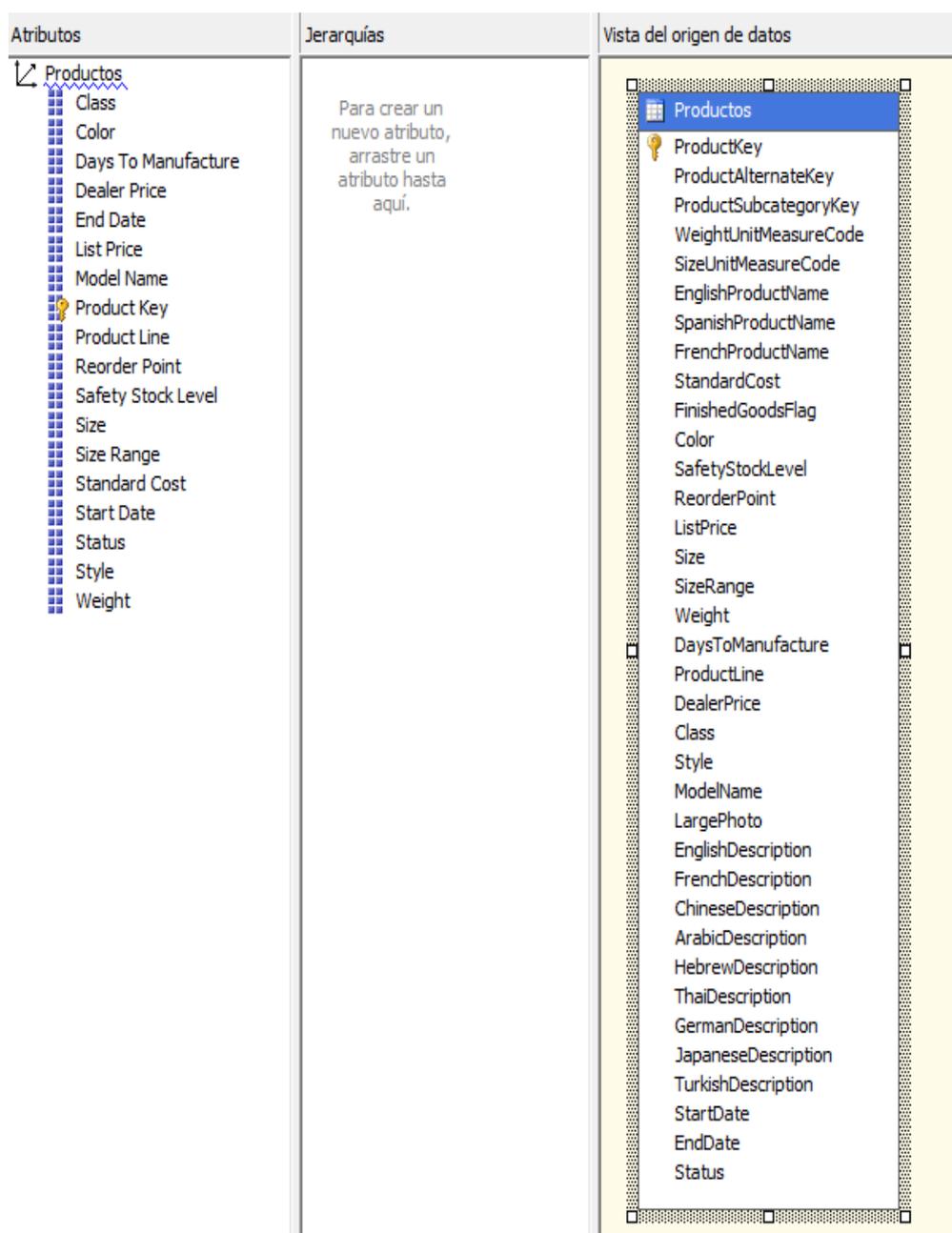


Figura 27: Implementación de la dimensión Productos

En el menú Archivo, haga clic en Guardar todo.

Revisar las propiedades de cubo y dimensión

Después de definir un cubo, se pueden revisar los resultados mediante el Diseñador de cubos. En la tarea siguiente, revisará la estructura del cubo del proyecto “Supermercado_Mia_Market”.

Para revisar las propiedades del cubo y de las dimensiones en el Diseñador de cubos

Para abrir el Diseñador de cubos, haga doble clic en el cubo Supermercado_Mia_Market en el nodo Cubos del Explorador de Soluciones

Definir medidas del cubo

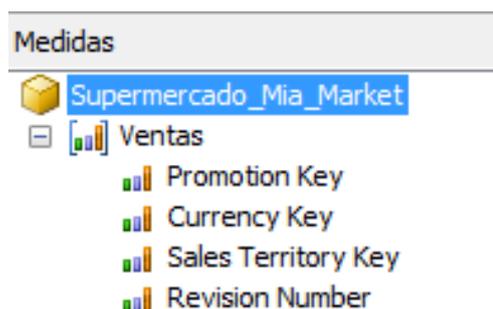


Figura 28: Definir medidas del cubo

En el panel Medidas de la ficha Estructura de cubo del Diseñador de cubos, expanda el grupo de medida Ventas para ver las medidas definidas. Si desea cambiar el orden de dichas medidas, arrástrelas para obtener el orden que desee. El orden afectará al modo en que determinadas aplicaciones ordenarán las medidas. El grupo de medida y cada una de sus medidas tienen propiedades que pueden modificarse en la ventana Propiedades. En el panel Dimensiones de la ficha Estructura de cubo del Diseñador de cubos, revise las dimensiones de cubo que se encuentran en el cubo “Supermercado_Mia_Market”.

Uso de dimensiones

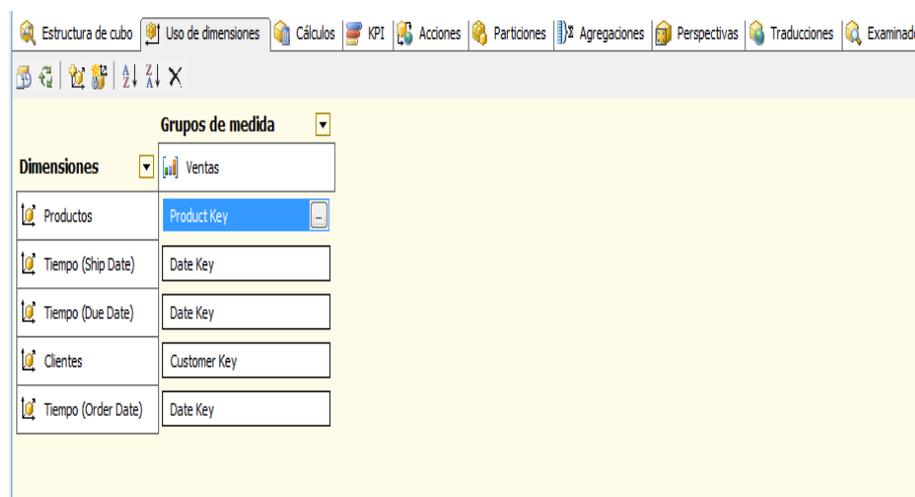


Figura 29: Uso de dimensiones

Haga clic en la ficha Particiones.

Particiones

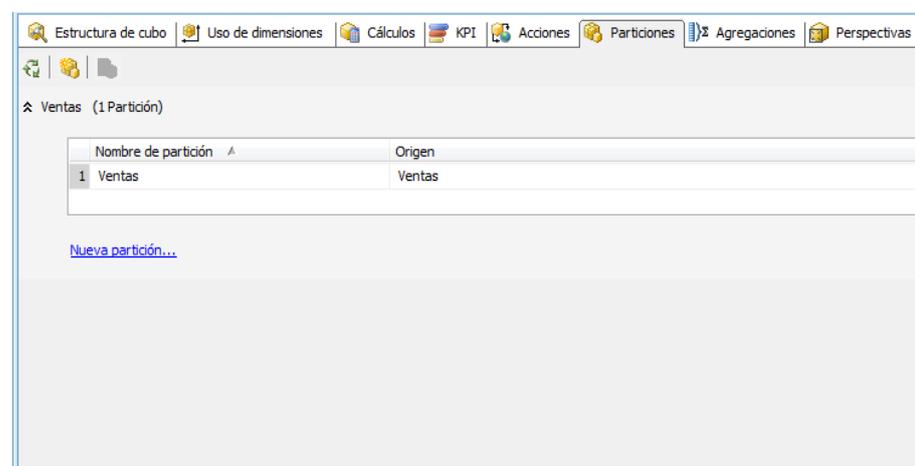


Figura 30: Particiones

El Diseñador de cubos definió una única partición para el cubo, utilizando el modo de almacenamiento de procesamiento analítico en línea multidimensional (MOLAP) sin agregaciones. Con MOLAP, todos los datos de nivel hoja y todas las agregaciones se almacenan en el cubo para maximizar el rendimiento. Las agregaciones son resúmenes de datos pre-

calculados que mejoran el tiempo de respuesta de las consultas ya que tienen las respuestas preparadas antes de que se planteen las preguntas. Puede definir particiones adicionales, parámetros de almacenamiento y parámetros de reescritura en la ficha Particiones. Haga clic en la ficha Examinador, Observe que el cubo no puede examinarse porque todavía no se ha implementado en una instancia de Analysis Services. En este punto, el cubo del proyecto “Supermercado_Mia_Market” es simplemente una definición de un cubo, que puede implementar en cualquier instancia de Analysis Services. Cuando implementa y procesa un cubo, puede crear los objetos definidos en una instancia de Analysis Services y rellenar los objetos con datos de los orígenes de datos subyacentes.

3.12.8. Implementar un proyecto de Analysis Services

Para ver los datos de dimensión y de cubo de los objetos del cubo “Supermercado_Mia_Market” del proyecto Supermercado_Mia_Market”, debe implementar el proyecto en una instancia determinada de Analysis Services y luego procesar el cubo y sus dimensiones. Al implementar un proyecto de Analysis Services se crean y definen objetos en una instancia de Analysis Services. Cuando se procesan los objetos en una instancia de Analysis Services, se copian los datos de los orígenes de datos subyacentes en los objetos del cubo. Para implementar el proyecto de Analysis Services En el Explorador de soluciones, haga clic con el botón secundario en el proyecto “Supermercado_Mia_Market” y, a continuación, haga clic en Propiedades. Aparece el cuadro de diálogo Páginas de propiedades de

“Supermercado_Mia_Market”, en el que se muestran las propiedades de configuración de Active(Development). Puede definir varias configuraciones, cada una con distintas propiedades. Por ejemplo, es posible que un programador desee configurar el mismo proyecto para implementarlo en distintos equipos de implementación y con distintas propiedades de implementación, como nombres de base de datos o propiedades de procesamiento. Fijese en el valor de la propiedad Ruta de acceso de los resultados. Esta propiedad especifica la ubicación en la que se guardan las secuencias de comandos de implementación XMLA cuando se crea un proyecto. Estas son las secuencias de comandos que se utilizan para implementar los objetos del proyecto en una instancia de Analysis Services. En el nodo Propiedades de configuración del panel de la izquierda, haga clic en Implementación. Revise las propiedades de implementación del proyecto. De forma predeterminada, la plantilla del proyecto de Analysis Services configura un proyecto de Analysis Services para implementar de forma incremental todos los proyectos en la instancia predeterminada de Analysis Services en el equipo local, crear una base de datos de Analysis Services con el mismo nombre que el proyecto y procesar los objetos después de la implementación utilizando la opción de procesamiento predeterminada. Temas relacionados: Configurar las propiedades de un proyecto de Analysis Services. Haga clic en Aceptar.

En el Explorador de soluciones, haga clic con el botón secundario en el proyecto “Supermercado_Mia_Market” y, a continuación, haga clic en Procesar.

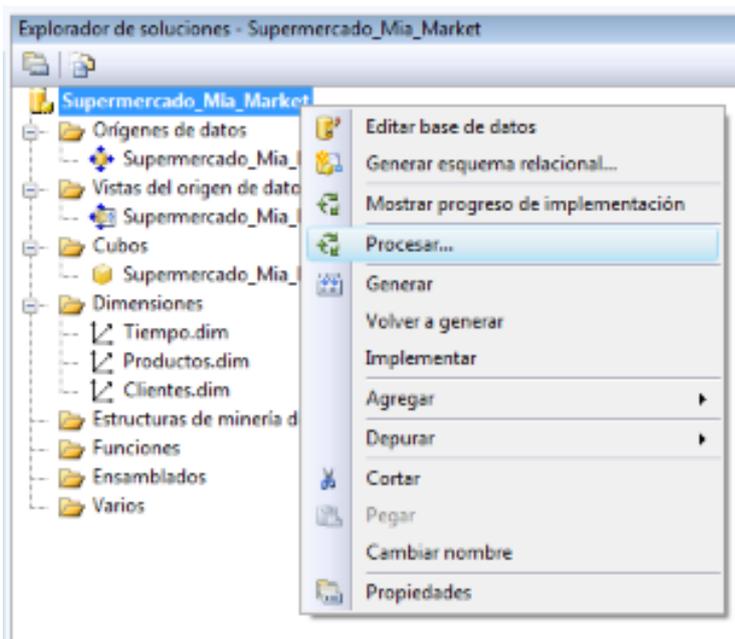


Figura 31: Proceso y ejecución del cubo 1

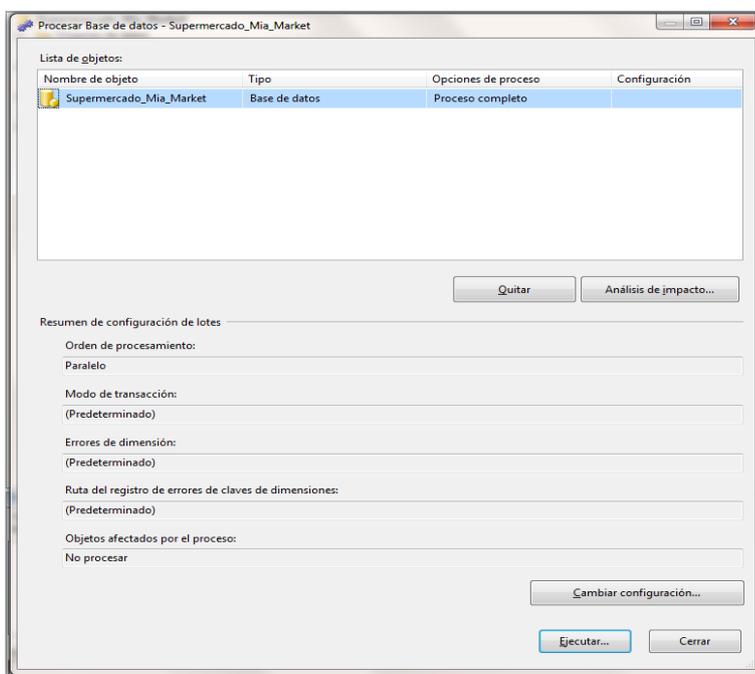


Figura 32: Proceso y ejecución del cubo 2

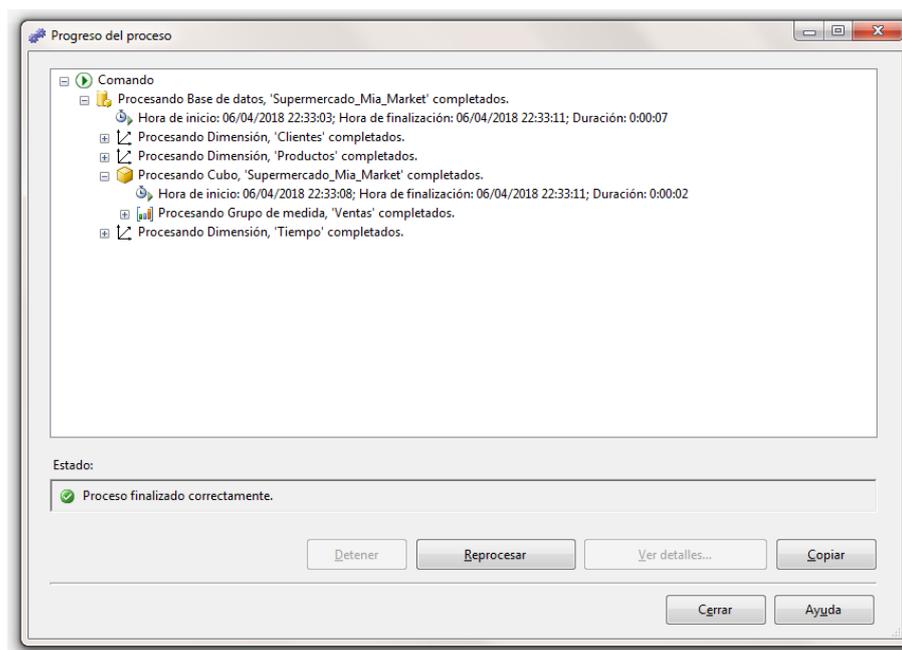


Figura 33: Proceso y ejecución del cubo 3

Ha implementado correctamente el cubo “Supermercado_Mia_Market” en la instancia local de Analysis Services y luego lo ha procesado.

3.12.9 Examinar el cubo

Una vez que se ha implementado un cubo, los datos de éste pueden verse en la ficha Examinadorb, para examinar el cubo implementado, cambie al Diseñador de dimensiones para la dimensión Productos en Business Intelligence Development Studio. Para ello, haga doble clic en la dimensión Productos del nodo Dimensiones del Explorador de soluciones.

Haga clic en la ficha Examinador para mostrar el miembro All de la jerarquía de atributo Product Key.

Examinador del cubo implementado

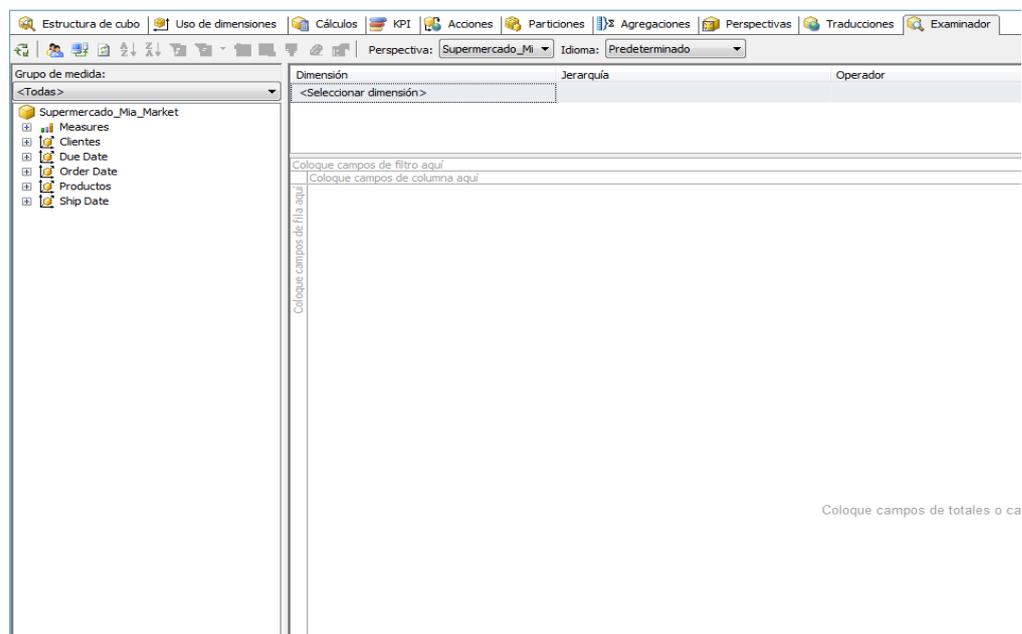


Figura 34: Examinador del cubo implementado

La idea es ver el cubo en trabajados con datos, para ello, vamos a realizar una prueba de cómo el cubo ver los datos a partir del cubo.

Un cubo es para ver un indicador de gestión, esa es la base, por lo cual nuestro indicador a buscar será “Ventas por país y detalladas por modelos de productos”.

Para eso debemos primero insertar una nueva medida a la tabla de hechos, que es la “Sales Amount”. Para ello vamos a la pestaña “Estructura de cubo” e insertamos “Sales Amount”

Cubo implementado

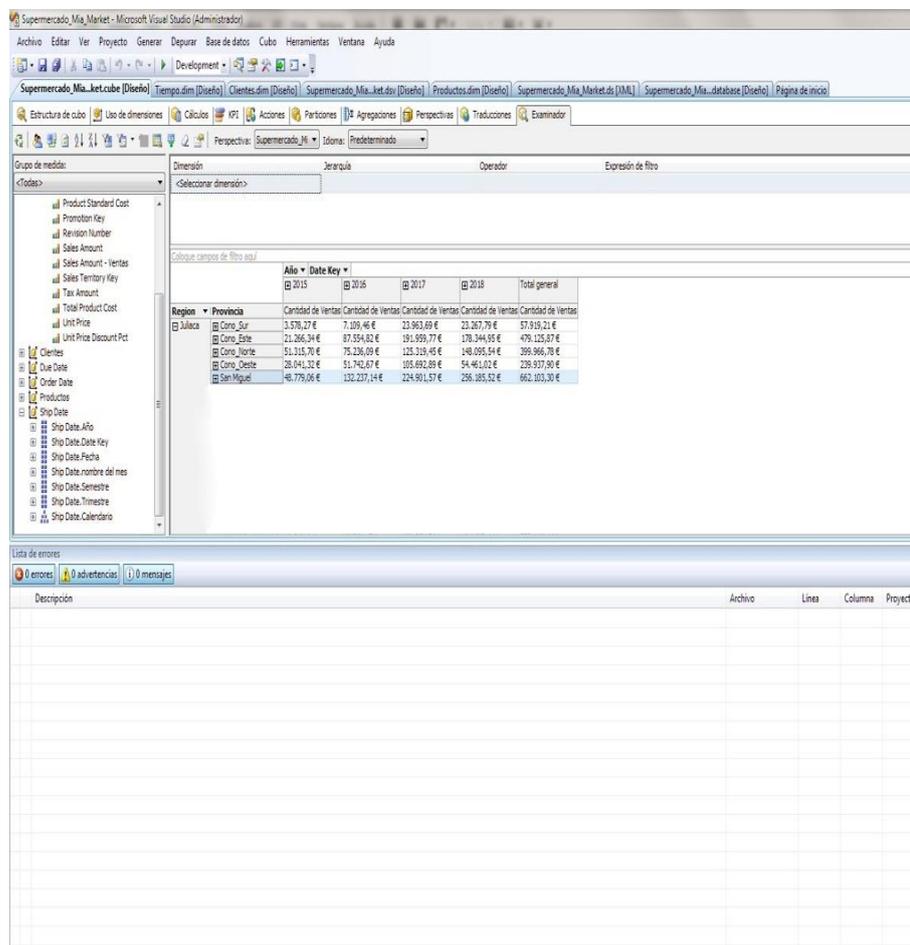


Figura 35: Cubo Implementado

Realizar la conexión

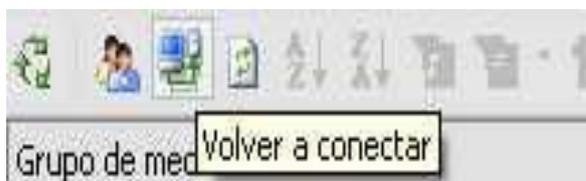


Figura 36: Realizar Conexión

Expandir el cubo implementado

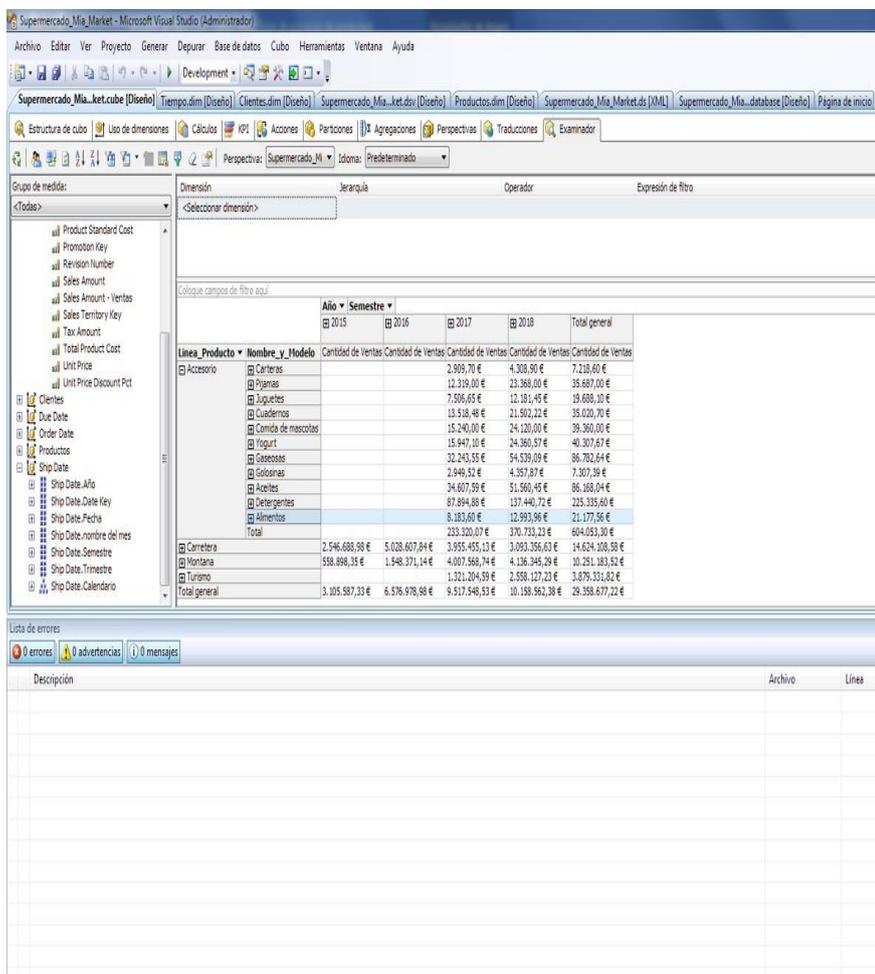


Figura 37: Expandir cubo implementado

CAPITULO IV

RESULTADOS Y DISCUSIÓN

4.1 IDENTIFICACIÓN DE REQUERIMIENTOS

La captura de requerimientos se realiza a través de entrevistas a los encargados de los componentes que son caso de estudio como son ingenieros informáticos y de sistemas de la oficina de informática y desarrollo de nuevos proyectos, generalmente asociados con información y reportes del movimiento económico de la empresa.

Los requerimientos están asociados a las necesidades de los directivos de la empresa empezando desde el gerente general y el cuerpo de directivos socios y jefes de área del supermercado.

4.2 REQUERIMIENTOS FUNCIONALES

Se lista los requerimientos que el sistema debe satisfacer para cumplir con los objetivos planteados para resolver el problema que atraviesa el gerente y el personal autorizado, jefes de área y directivos de la empresa.

Tabla 17: *Requerimientos Funcionales*

REQUERIMIENTO FUNCIONAL	PRIORIDAD	DIFICULTAD
El sistema permitirá generar un reporte en el que se muestre los productos por nombres y modelos al momento de ser consultados.	2	1
El sistema permitirá generar indicadores del número total de productos con sus precios.	1	2
El sistema permitirá reflejar la cantidad total productos por	1	3

año, semestre, trimestre mensual, semanas y por días.		
El sistema permitirá comparar el estado de la cantidad de ventas.	1	2
El sistema permitirá comparar el estado de cada producto y modelo con los precios respectivos.	1	2

Fuente: Elaboración Propia

Leyenda de Prioridad

Tabla 18: Leyenda de prioridad

Calificación	Prioridad
1	Importante
2	Regular
3	Prescindible

Fuente: Elaboración Propia

Leyenda Dificultad

Tabla 19: Leyenda de dificultad

Calificación	Prioridad
1	Fácil
2	Regular
3	Difícil

Fuente: Elaboración Propia

4.3 REQUERIMIENTOS NO FUNCIONALES

Exportación de reportes a archivos de formatos estándar como archivo de texto (.txt) o Excel (.xls). Mostrar la información del Datamart en base a los siguientes 3 tipos: Reportes tradicionales (basados en columnas), Tablas pivote (Matriz de doble entrada) y Gráficos.

Entre otros requerimientos no funciones de la herramienta se tiene:

Disponibilidad: El sistema al ser de consultas y análisis no impacta a las operaciones diarias que se realizan en el supermercado por ello que su disponibilidad puede ser menor a 24 horas los 7 días de la semana. Además, el sistema no se encontrar disponible los días en los que se realice la actualización de datos del Datamart.

Utilidad: El sistema será apropiado para el uso de los distintos perfiles, de manera que sirvan como soporte de las labores de todas personas que interactúen con él.

Escalabilidad: Las herramientas de negocios permiten cualquier tipo de consultas y reportes que más adelante se podrían solicitar.

4.4 ANÁLISIS DE REQUERIMIENTOS

Para el análisis de requerimientos se siguió la metodología de diseño dimensional en 4 etapas, para luego hacer el modelo dimensional.

Proceso de diseños dimensional de 4 etapas.

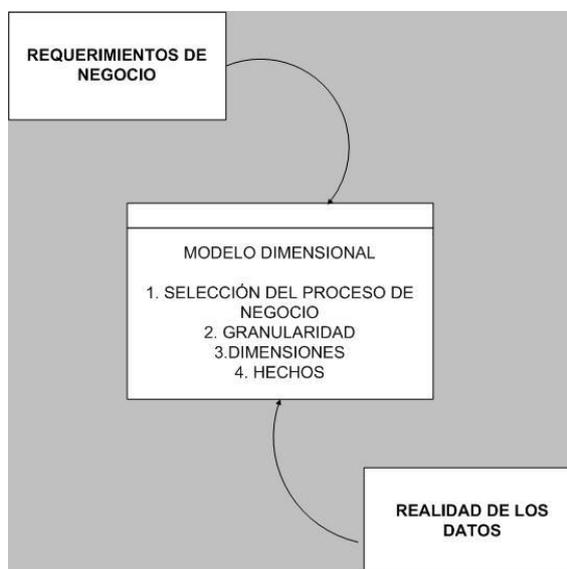


Figura 38: Proceso de diseño de 4 etapas

a) Selección del proceso del negocio

Después de haber realizado la entrevista con el personal que labora en la gerencia Mia Markert se decide: El sistema de inteligencia de negocios realizará seguimiento al proceso de negocio relacionados con toma de decisiones. El objetivo es tener un mejor conocimiento de la evolución de las áreas antes mencionadas, para poder analizar los resultados comparando la evolución con lo planificado y así tomar las medidas correctivas.

b) Granularidad

A partir de los requerimientos de información se debe definir la granularidad. ¿Qué nivel de detalle debería estar disponible en el modelo dimensional?

Este paso es sumamente importante ya que de él dependerá el modelo de datos y la implementación del Datamart.

c) Dimensiones

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

d) Hechos

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

Todos los hechos candidatos deben ser fieles al grano definido en b). Los hechos que pertenecen a un grano diferente deberán estar en una tabla de hechos separada.

4.5 DISEÑO Y DESARROLLO DE LA PRESENTACIÓN DE DATOS

En esta etapa se realizó la extracción, transformación y la carga de los datos que son necesarios para nuestro Data Warehouse con la herramienta Pentaho Data Integration (Kettle).

Arquitectura de la plataforma de negocios

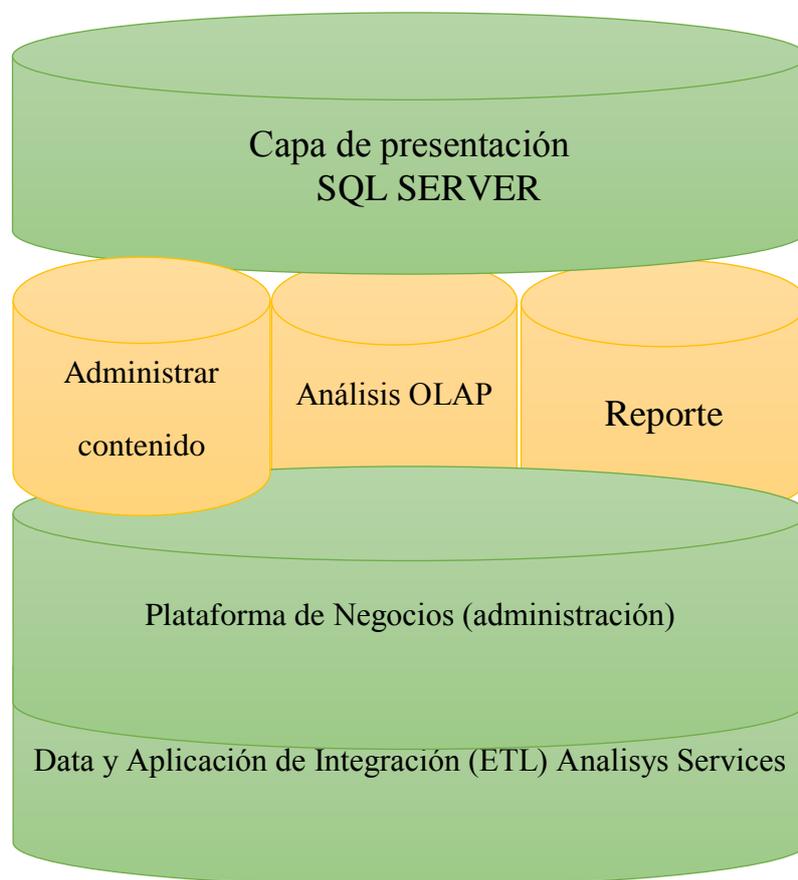


Figura 39: Arquitectura de la plataforma de negocios

4.6 CONSTRUCCIÓN DE PROCESOS ETL

Seguidamente se muestra los procesos ETL que se realizó

a) ETL Dimensión Tiempo

El proceso va a generar todos los tiempos desde el 1 de enero del 2010 hasta el 31 diciembre del 2018.

b) ETL demás dimensiones

El proceso ETL de las demás dimensiones, menos complicado puesto que simplemente se va a extraer tablas con sus respectivos datos de base de datos original y se almacenan en el Data Warehouse.

Para este proceso se crea un JOB que contienen todas transformaciones necesarias para poblar el Data Warehouse. Para este fin se tiene que realizar la conexión a la base de datos de la entidad.

Una vez realizada la conexión se realizó las transformaciones para extraer las tablas necesarias de la base de datos relacional de la entidad para luego seleccionarlas y finalmente ingresar esos datos en el Data Warehouse.

Para que el proceso de carga de datos de las diferentes tablas se realice de forma automática se crea el JOB que contiene todas las transformaciones.

c) ETL Carga de Hechos

La carga de datos en las tablas de Hechos del Data Warehouse se realizó de forma similar simplemente se cambió los parámetros de datos de origen y datos de salida de las transformaciones.

4.7 DISEÑO DE LA ARQUITECTURA TÉCNICA

En esta parte de la metodología se establece la arquitectura de la plataforma de Inteligencia de Negocios

Capa de presentación mediante ella los usuarios finales en este caso las personas que conforman el gerente y los directivos del supermercado, los cuales acceden a los datos para su análisis y luego elegir la creación de reportes, vistas de datos y administrar contenidos.

La plataforma de inteligencia de negocios es la integración de herramientas se Server Business Intelligence la cual contiene la herramienta para el análisis de cubos OLAP, administración de contenidos y generación de reportes.

4.8 IMPLEMENTACIÓN DE NEGOCIOS

Siguiendo con la metodología, se realiza la implementación de la plataforma de negocios. A continuación, se presenta el desarrollo concerniente a la etapa de implementación de la plataforma.

IMPLEMENTACIÓN

Consola de Usuario. - En esta fase se pone en funcionamiento toda la plataforma dirigida a los usuarios finales.

Primero se ingresa a la herramienta a través de SQL Server, ya que esta herramienta es una consola de consultas que contiene toda la base de datos con el usuario de Windows que es dado por defecto al momento de la instalación o para mayor seguridad asignar como usuario el nombre de la PC y su propia contraseña

Conectando servidor con SQL Server 2008 R2



Figura 40: Conectando servidor con SQL Server 2008 R2

Segundo se ingresa al Analisis Services de SQL server y Intelligence Bussines de forma directa.

4.9 PRUEBA DE VALIDEZ Y CONSISTENCIA DE DATOS.

El sistema de inteligencia de negocios está desarrollado sobre la plataforma de Business Intelligence Analisis Services, el cual es una herramienta de calidad, tiene características de datos, las cuales son características propias de esta herramienta.

La confiabilidad y consistencia de la información obtenida, como el correcto procesamiento de los datos, son cuestiones que han sido evaluadas en investigaciones similares a la presente en las que se ha utilizado como herramienta la plataforma SQL server 2008 y que, usando técnicas para los resultados, la confirmación de la consistencia de datos. Para la presente investigación se ha decidido usar un sistema de pruebas, el cual implica la operación o aplicación del mismo a través de condiciones controladas y la consiguiente evaluación de la información. Las condiciones controladas deben incluir tanto situaciones normales como anormales. El objetivo del

sistema de pruebas es encontrar un error para determinar situaciones en donde algo pasa cuando no debe de pasar y viceversa.

En la presente investigación se hizo la prueba estadística como las encuestas y así comprobar si son los correctos.

4.10 CONTRASTACIÓN DE HIPÓTESIS

La contrastación de la hipótesis se realizó de acuerdo al método propuesto Pre-Test–Post-Test, que nos permitirá aceptar o rechazar la hipótesis. Por lo cual se realizó una prueba por cada indicador en las cuales se usarán las siguientes fórmulas.

Prueba T-Student

$$\bar{D} = \frac{\sum_{i=1}^n Di}{n} \dots\dots\dots (1)$$

$$\sigma_D^2 = \frac{n \sum_{i=1}^n Di^2 - (\sum_{i=1}^n Di)^2}{n(n-1)} \dots\dots\dots (2)$$

$$t = \frac{\bar{D}\sqrt{n}}{\sqrt{S_D}} \dots\dots\dots (3)$$

Dónde:

\bar{D} : Diferencia de promedio

σ_D^2 : Desviación estándar de las diferencias

INDICADORES

AD : Nivel de eficiencia en la toma de decisiones antes de la implementación del Datamart.

DD : Nivel de eficiencia en la toma de decisiones después de la implementación del Datamart.

HIPÓTESIS ESTADÍSTICA

HIPÓTESIS NULA

H₀: El desarrollo de un Datamart, no optimizará el proceso de toma de decisiones en la gerencia del supermercado Mia Market Juliaca.

HIPÓTESIS ALTERNA

H_a: El desarrollo de un Datamart, optimizará el proceso de toma de decisiones en la gerencia del supermercado Mia Market Juliaca.

$$H_0 : AD - DD \leq 0$$

$$H_a : AD - DD > 0$$

NIVEL DE SIGNIFICANCIA

El nivel de significancia será del 5%. Es decir, $\alpha=0.05$ y $n-1=8$ grados de libertad; se tiene el valor crítico de T de Student:

$$t_{8,0.05}=1.86$$

Como $\alpha=0.05$ y $n-1=8$ grados de libertad, la región de rechazo consiste en aquellos valores de t mayores que: $t_\alpha = 1.86$

4.11 SITUACIÓN ACTUAL

Se muestra la situación actual sobre la eficiencia a la hora de realizar las consultas respectivas para tomar decisiones acertadas.

Tabla 20: Pre-Test del sistema actual para tomar decisiones del supermercado Mia Market

PREGUNTA	CALIFICACIÓN					Puntaje Total	Puntaje Promedio
	E	B	R	M	D		
	5	4	3	2	1		
1.- El desarrollo del DataMart, ¿Está cumpliendo los requisitos solicitados?	0	2	3	3	1	24	2.67
2.- En la toma de decisiones, ¿cómo califica usted la rapidez de respuesta de la herramienta a la hora de solicitar información?	0	1	4	3	1	23	2.56
3.- ¿Cómo es la interacción del Datamart con la solución de inteligencia de negocios?	0	0	5	4	0	23	2.56
4.- La solución de inteligencia de negocios ¿Mejora el proceso de decisiones?	0	1	5	3	0	25	2.78

Fuente: Elaboración propia

Tabla 21: Post-Test del sistema actual para tomar decisiones del supermercado Mia Market

PREGUNTA	CALIFICACIÓN					Puntaje Total	Puntaje Promedio
	E	B	R	M	D		
	5	4	3	2	1		
1.- El desarrollo del DataMart, ¿Está cumpliendo los requisitos solicitados?	1	3	4	1	0	31	3.44
2.- En la toma de decisiones, ¿cómo califica usted la rapidez de respuesta de la herramienta a la hora de solicitar información?	3	3	2	1	0	35	3.89
3.- ¿Cómo es la interacción del Datamart con la solución de inteligencia de negocios?	4	3	1	1	0	37	4.11
4.- La solución de inteligencia de negocios ¿Mejora el proceso de decisiones?	5	3	1	0	0	40	4.44

Fuente: Elaboración propia

Acrónimos de calificativos

Tabla 22: Abreviaturas de calificación

Abreviatura	Descripción
E	Excelente
B	Bueno
R	Regular
M	Malo
D	Deficiente

Fuente: Elaboración propia

Definición de variables

Tabla 23: Definición de variables

Abreviatura	Descripción
AD_i	Puntuación del sistema actual
DD_i	Puntuación con el Datamart desarrollado
D_i	$(AD_i - DD_i)$
D_i^2	$(AD_i - DD_i)^2$

Fuente: Elaboración propia

Tabla 24: Contrastación entre el Pre-test y Post-test

Pregunta	AD_i	DD_i	D_i	D_i^2
1.- El desarrollo del DataMart, ¿Está cumpliendo los requisitos solicitados?	2.67	3.44	-0.78	0.60
2.- En la toma de decisiones, ¿cómo califica usted la rapidez de respuesta de la herramienta a la hora de solicitar información?	2.56	3.89	-1.33	1.78
3.- ¿Cómo es la interacción del Datamart con la solución de inteligencia de negocios?	2.56	4.11	-1.56	2.42
4.- La solución de inteligencia de negocios ¿Mejora el proceso de decisiones?	2.78	4.44	-1.67	2.78

Fuente: Elaboración propia

4.12 RESULTADOS DE LA HIPÓTESIS ESTADÍSTICA

DIFERENCIA DE PROMEDIO.

Reemplazando en la formula (1)

$$\bar{D} = \frac{\sum_{i=1}^n Di}{n} = \frac{-5.33}{9} = -0.59$$

DESVIACIÓN ESTÁNDAR

Reemplazando en la formula (2)

$$\sigma_D^2 = \frac{n \sum_{i=1}^n Di^2 - (\sum_{i=1}^n Di)^2}{n(n-1)} = \frac{9(7.58) - (-5.33)(-5.33)}{9(9-1)} = 0.55$$

EL VALOR DE LA T CALCULADA

Reemplazando en la formula (3)

$$t = \frac{\bar{D} \sqrt{n}}{\sqrt{\sigma_D}} = \frac{-0.59 \sqrt{9}}{\sqrt{0.74}} = -2.06$$

CONCLUSIÓN

Dado que $t_c = -1.86$, es menor que Valor estando este dentro de la región de rechazo, se concluye que $AD - DD < 0$, se rechaza la hipótesis nula H_0 , y se acepta la hipótesis alterna H_a , entonces se prueba la valides de hipótesis con un nivel de significancia $\alpha=0.05$, por lo tanto, se puede afirmar que el desarrollo de un Datamart optimizará el proceso de toma de decisiones con la aceptación mayoritaria en la gerencia del supermercado Mia Market Juliaca.

Grafica de T-Student

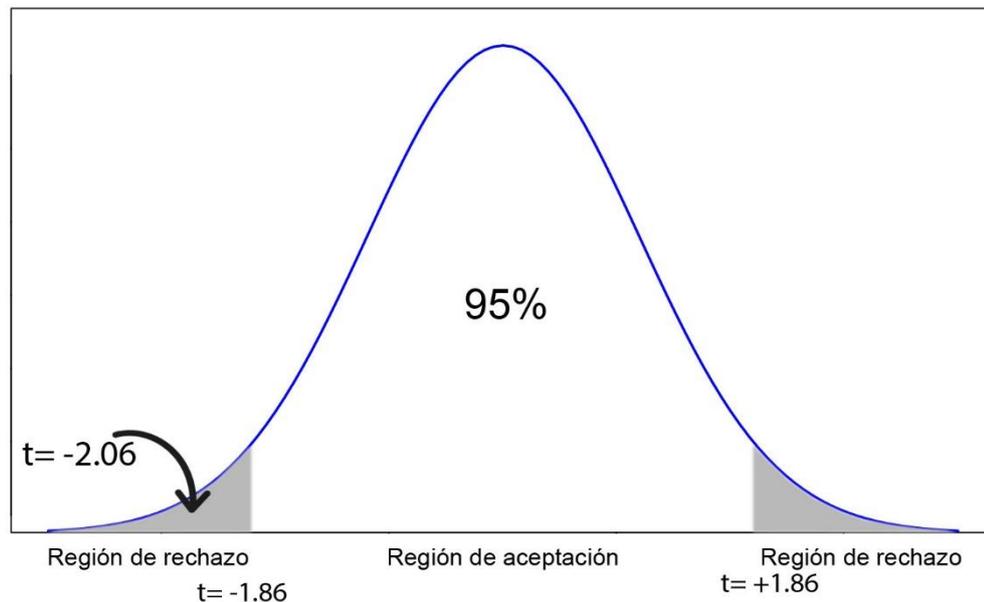


Figura 41: Grafica de T-Student

DISCUSIÓN

A partir de los hallazgos encontrados, aceptamos la hipótesis alterna que establece que existe relación de dependencia entre la eficiencia en la toma de decisiones con el sistema desarrollado en Mia Market de la ciudad de Juliaca.

Estos resultados guardan relación con lo que sostiene **Gutierrez Colmenares (2016)** en usuarios resaltando las ventajas de los avanzados en construcción de un sistema de apoyo a la toma de decisiones para el área de gerencia del Hospital (Post-Grados). Que a su vez guarda relación con **Nuñez Bejarano (2010)** y **Fernandez Bermudez (2010)** que realizó un análisis, diseño e implementó una solución de inteligencia de negocios para el área de Finanzas en la municipalidad de Lima. Donde los usuarios en estudio también realizan los procedimientos de reportes de forma manual. Estos autores expresan que las personas tienen

dificultades a la hora de realizar las consultas con información lenta y poco fiable, Ello es acorde con lo que en este estudio se halla.

Pero, en lo no concuerda el estudio de los autores referidas con el presente, es que ellas mencionan que su muestra de estudio son cantidades más grandes y utilizan otras metodologías **Lopez Dilmar (2007)**, **Illbay Maydana (2009)** utiliza más de tres variables de estudio. debido a que las áreas también cuentan con mayor personal por tratarse de entidades públicas. En las herramientas de uso también cuentan con una base de datos mucho mayor como son los Datawarehouses y gestores de base de datos como Oracle. En este estudio, se utilizó SQL Server R2 y no se encuentran estos resultados.

En lo que respecta a la relación entre el desarrollo de un sistema de inteligencia de negocios y la variable eficiencia este estudio todos encuentran relación alguna. Pero, **Villanueva Fuentes (2008)** y **Guillen Quilca (2017)** implementan tecnologías que generan reportes de forma personaliza en minutos y segundos tratándose de Datawarehouse para su correcta toma de decisiones en asuntos de mayor relevancia.

CONCLUSIONES

PRIMERO:

Con el desarrollo del Datamart para la gerencia del supermercado Mia Market; se logró identificar la información requerida, haciendo posible la extracción de la base de datos transaccional, para determinar cuál es la cantidad de productos que ingresan y salen de la empresa, los cuales son factores determinantes para el desarrollo de la herramienta necesaria y tomar decisiones consolidando la información relevante.

SEGUNDO:

Con el desarrollo del Datamart para la gerencia del supermercado Mia Market; se logró analizar el modelamiento multidimensional de las distintas áreas que apoyan a gerencia mediante los informes de los reportes de forma rápida, proceso que simplifica las consultas de las dimensiones que proporcionan un único origen de datos.

TERCERA:

Con el desarrollo del Datamart para la gerencia del supermercado Mia Market; se diseñó la herramienta que realiza un modelado multidimensional representando las actividades comerciales y empresariales; haciendo posible la extracción, transformación y carga de toda la base de datos, para su correcta interpretación y tomar decisiones más fiables que beneficien a la empresa.

CUARTO:

Con el desarrollo del Datamart para la gerencia del supermercado Mia Market; se implementó la herramienta logrando el rendimiento de las consultas rápida mediante interfaz de usuario en los datos empresariales. Se pueden anotar y ampliar para admitir construcciones de consultas complejas. Los desarrolladores crean cubos para admitir

tiempos de respuesta rápida desplegando la información con mayor facilidad en el menor tiempo posible, Otra ventaja importante del uso del Datamart de datos multidimensionales de Analysis Services es la integración con las herramientas de informes BI utilizadas habitualmente, como Excel, Reporting Services y PerformancePoint, así como las aplicaciones personalizadas.

Finalmente, se concluye que el Datamart se desarrolló cumpliendo paso a paso las especificaciones y su correcto funcionamiento con un interfaz confiable para su manipulación y ejecución para que cumpla con su labor.

RECOMENDACIONES

PRIMERO:

Seguir con la construcción y tener acceso completo al Datawarehouse por limitaciones de política de la empresa para tener un mayor control y monitoreo. En el presente proyecto las bases están construidas para seguir extendiendo la Datamart en las demás sucursales pronto a funcionar.

SEGUNDO:

Construir una herramienta propia de la organización, para la visualización del Datamart, haciendo que esto cuente con inteligencia de negocios.

TERCERO:

Se recomienda realizar entrevistas para una adecuada toma de requerimientos, conocimiento del giro de negocio, estar interactuando con los empleados y clientes para el diseño de una adecuada estructura de base de datos del sistema.

CUARTO:

Se recomienda documentarse bien en el uso de las herramientas y realizar pruebas antes de iniciar el uso de “producción” de estas. Se puede conocer muy bien el proceso a desarrollar, pero si las herramientas no son utilizadas de la manera correcta entonces llevará al fracaso del producto final.

REFERENCIAS BIBLIOGRÁFICAS

LIBROS

- Hernández, R., Fernández, C., & Baptista, M. (2010). *Metodología de la Investigación (5 ed.)*. Mexico DF, Mexico: McGram-Hill.
- Howson, C. (2009). *Bussines Intelligence - Estrategias para una implementación exitosa (1 ed.)*. Mexico, Mexico: The McGraw-Hill Companies, Inc.
- Kimball, R. (2004). *The Data WareHousetagin toolkit*. London, Inglaterra: IDG Books.
- Kimball, R. (2008). *Data WareHouse Lifecycle Tolkit*. London, Inglaterra: LEA.
- Kimball, R., & Ross, M. (2012). *The Data Warehouse Toolkit (2 ed.)*. EE.UU: John Wiley & Sons, Inc.
- Díaz, J. (2010). *Introducción al Business Intelligence (1ra ed)*. Barcelona, España: UOC.
- Larissa, M., & Atre, S. (2013). *Business Intelligence Roadmap: The Complete Project Lifecycle for Decision Support Applications (3 ed.)*. Boston, EE.UU: Pearson Education.
- Lluís, J. (2008). *BussinesIntelligence: Competir con Información*. Barcelona: Dataprix.
- Ralph, K., & Ross, M. (2012). *The data WareHouse toolkit: The complete to dimensional modeling*. New York, EE.UU: WILEY.
- Sánchez, L. (2014). Análisis de Información y Toma de decisiones para Administración de Negocios. *Tesis de Bachiller publicada*. Universidad Nacional Autónoma de México, México.

TESIS

- Chávez, D. (2015). Sistema de Soporte a la Toma de Decisiones basado en Inteligencia de Negocios para Mejorar los Procesos Comerciales del Importador Peruano. *Tesis de maestría publicada*. Universidad Católica Santo Toribio de Mogrovejo, Chiclayo, Perú.
- Fernández, E. (2010). Análisis, Diseño e Implementación de un Data Mart de clientes para el área de Marketing de una Entidad Aseguradora. *Tesis de Bachiller publicada*. Pontificia Universidad Católica del Perú, Lima, Perú.
- López, C. (2007). Análisis, diseño e implementación de un Datamart para la dirección financiera y recursos humanos de la Escuela Politécnica para una toma de desicion efectiva. *Tesis de Pre-grado*. Escuela Politécnica del Ejercito, Sagolqui, Ecuador.
- Moreno, R. (2013). *Análisis, Diseño e Implementación de Datamarts para las áreas de Ventas y Recursos Humanos de una Empresa dedicada a la exportación e importación de Productos Alimenticios*. Pontificia Universidad Católica del Perú, Lima, Perú.
- Núñez, G. (2008). Análisis, Diseño e Implementación de una Solución de Inteligencia de Negocio para el área de Finanzas de la Municipalidad de Lima. *Tesis de Bachiller publicada*. Pontificia Universidad Católica del Perú, Lima, Perú.
- Villanueva, Á. (2008). Análisis, Diseño e Implementación de un Data Warehouse de soporte de decisiones para un Hospital del Sistema de Salud Público. *Tesis de Bachiller publicada*. Pontificia Universidad Católica del Perú, Lima, Perú.

FUENTES ELECTRONICAS

Bernabeu, D. (2008). *Data Warehousing & Data Warehouse*. Obtenido de DaraPrix:

<http://www.dataprix.com/data-warehousing-y-metodologia-hefesto/i-data-warehousing-investigacion-y-sistematizacion-conceptos-0>

Bernabeu, D. (2016). *DataPRIX Knowledge is the Goal Bussiness Intelligence*

Funcionalidades de Business Intelligence. Obtenido de Bussiness Intelligence:

<http://www.dataprix.com/281-funcionalidad-business-intelligence>

BI. (2018). *Business Intelligence Fácil. Microstrategy*. Obtenido de Microstrategy:

<http://www.businessintelligence.info/productos/microstrategy.html>

Cientifica, R. (2017). *Cómo diseñar grandes variables en bases de datos*

multidimensionales. Obtenido de Inteligencia de negocios:

<https://www.uv.es/=buso/gv/gv.html>

Datawarehouse. (2018). *Almacen de datos*. Obtenido de Datawarehouse:

<http://datawarehouse4u.info/>

Espinosa, R. (2018). *El Rincon del BI. Construccion procesos ETL utilizando Kettle*.

Obtenido de Inteligencia de negocios:

<http://churriwifi.wordpress.com/2010/05/10/16-3-construccion-procesos-etl-utilizando-kettle-pentaho-data-integration/>

Gravitar. (2018). *Herramientas BI*. Obtenido de Desarrollo de Datamart:

<http://www.gravitar.biz/index.php/herramientas-bi/>

HD, R. (2016). *Creacion de un datamart paso a paso*. Obtenido de

<https://raynerhd.com/2008/10/22/creacion-de-un-datamart-paso-a-paso/>

Pérez, M. (2017). *Gestión de la información*. Obtenido de Datamart:

<http://glossarium.bitrum.unileon.es/Home/gestion-de-la-informacion>

Rojas, C. (2018). *Datawarehousing*. Obtenido de Creacion de Datawarehouse y Cubos

OLAP: <http://html.rincondelvago.com/datawarehousing.html>

Shisubha, H. (2017). *ETL SERVICES*. Obtenido de Base de Datos Transaccional:

<http://www.srisubha.co.za/solutions/etl-services/>

Sinnexus. (2017). *Business Intelligence - Informática Estratégica*. Obtenido de Inteligencia de negocios:

http://www.sinnexus.com/business_intelligence/sistemas_soporte_decisiones.aspx

Software, E. (2018). *Cubos OLAP de información para la toma de decisiones*. Obtenido de

Creacion de cubos: <http://www.evaluandosoftware.com/cubosolap-informacion-la-toma-decisiones/>

Wolff, C. (2017). *Modelamiento Multidimensional*. Obtenido de Inteligencia de negocios:

<http://www.inf.udec.cl/revista/edicion4/cwolff.htm>

Zen. (2017). *Bring your data to life Bussines Intelligence*. Recuperado el Consultado el 30 de Enero de 2018, de Data to life: <http://www.zenlatam.com/business-intelligence-pentaho/>

ANEXOS

ANEXO A: Guía de Entrevista N° 01

Entrevistado: Tec.	Fecha:
Entrevistador: Alain Yucra Halanoca	Área: Área de Informática
<p>Objetivo:</p> <p>Conocer la situación actual del área de informática para obtener reportes</p>	
<p>Dirigido a:</p> <p>Ing. del módulo de informática</p>	
<p>Preguntas:</p> <p>¿Cuáles son los pasos que se requiere para solicitar un reporte del área de informática?</p> <p>¿Cuáles son los reclamos que se presentan en cuanto a la solicitud de reportes?</p>	
<p>Resumen:</p> <p>Podemos hallar dos situaciones, PRIMERO, el reporte se puede encontrar realizado con anterioridad y solicitarlo directamente con el jefe de área o el encargado para su verificación. SEGUNDO, es cuando el reporte aún no se encuentra disponible y está en proceso; entonces para que podamos proporcionarlo en archivos distintos como Excel o pdf al solicitante.</p> <p>Los reclamos más comunes son el tiempo debido que en algunos casos el jefe de área o encargado se encuentra en reuniones, fallas en sistema o mantenimiento de ordenadores; y eso hace que la espera sea demasiado para obtenerlo.</p>	

ANEXO B: Guía de Entrevista N° 02

Entrevistado: Secr.	Fecha :
Entrevistador: Alain Yucra Halanoca	Área : Área de Gerencia
Objetivo: Conocer la situación actual del área de Gerencia para tomar decisiones	
Dirigido a: Gerente de Mia Market SRL	
Preguntas: ¿Cuál es la función de gerencia en el supermercado Mia Market SRL? ¿Cuál de sus funciones están directamente relacionadas al momento de tomar decisiones con el apoyo de otras áreas? ¿Cómo califica usted al sistema de gestión administrativa? ¿Cómo se realiza los reportes en el supermercado?	
Resumen: <p>Gerencia es el encargado de efectuar la dirección, evaluación y ejecución de las actividades con ayuda de áreas relacionadas con los procesos de presupuesto, contabilidad y tesorería, así como elaborar e interpretar los estados financieros y presupuestales de la empresa. Está a cargo de un gerente.</p> <p>Dentro de nuestras funciones de Gerencia las que están más ligadas a la Alta Dirección son: “Dirigir, evaluar y ejecutar los procesos que permitan presentar a la Alta Dirección la información económica y financiera oportuna para la toma de decisiones,</p>	

con contabilidad formular el presupuesto institucional en coordinación con las unidades orgánicas, emitir los informes de disponibilidad presupuestaria para la adquisición o contratación de bienes, servicios que requiera la entidad, formular la consolidación mensual de gastos y efectuar la consolidación con los saldos de balance y presentar a la alta dirección los estados financieros y presupuestales del pliego”.

Mia Market, actualmente es un sistema que, si cumple con nuestras necesidades diarias, pero existe un déficit en cuanto a la generación de reportes, debido a la espera que existe para poder obtenerlos.

Para solicitar los reportes existe un menú dentro del sistema, en el cual yo como encargada/secretaria de gerencia, puedo seleccionar los reportes que necesito, y este sistema me los proporciona en Excel o un archivo pdf.

Lo mismo sucede cuando la Alta Dirección está en reunión y desean obtener reportes en ese momento. Ellos interactúan con el sistema con ayuda del área de informática para realizar estos reportes y tienen que esperar un tiempo para poder obtener la información. Esto se constituye en una queja constante, por parte de la alta dirección, pues tienen que esperar demasiado tiempo para obtener la información.

ANEXO C: Funciones de la Gerencia de Mía Market SRL

funciones del área de gerencia son las siguientes:

1. Formular, ejecutar y evaluar los planes de gerencia.
2. Organizar, programar y ejecutar los procesos que permitan presentar a la alta dirección la información económica y financiera oportuna para la toma de decisiones.
3. Formular el presupuesto institucional en coordinación con las unidades orgánicas.
4. Elaborar los requerimientos de información para la ejecución de nuevos proyectos como presupuestaria, consistentes en la programación mensual de ingresos y gastos y programación de asignación trimestral, para la aprobación de los calendarios de compromisos.
5. Elaborar la formalización mensual del presupuesto institucional y gestionar los pedidos de mayores recursos que requiera la entidad.
6. Emitir los informes de disponibilidad presupuestaria para la adquisición o contratación de bienes, servicios u obras que requiera la entidad.
7. Emitir opinión técnica en materia presupuestal y absolver las consultar que incidan en el aspecto presupuestal, que formulen las unidades orgánicas de la empresa.
8. Ejecutar las operaciones del sistema de contabilidad del pliego presupuestal de la institución, controlando las operaciones contables y financieras en los libros correspondientes y en el sistema integrado de administración financiera (SIAF), verificando su sustentación.

9. Formular la consolidación mensual de gastos y efectuar la consolidación con los saldos de balance.
10. Presentar a la Alta Dirección los estados financieros y presupuestales del pliego, así como los anexos respectivos en el marco de la normativa legal vigente.
13. Verificar la realización del registro contable de los gastos comprometidos en la ejecución presupuestal, de acuerdo a las normas vigentes.
11. Ejecutar el pago de las obligaciones y efectuar las conciliaciones bancarias respectivas.
12. Administrar los recursos financieros de la empresa, en el marco de la normativa legal vigente.
13. Verificar la documentación que sustenta las operaciones financieras de ingresos y egresos de fondos.
14. Controlar y custodiar los recursos y valores a cargo de la institución.
15. Manejar cuentas y sub cuentas bancarias, por toda fuente de financiamiento y ejecutar las conciliaciones correspondientes.
16. Evaluar y resolver los expedientes y/o documentos correspondientes a su competencia funcional y efectuar su seguimiento, con observancia de las políticas, normal y procedimiento establecidos.
17. Proponer las normas y procedimientos orientados a mejorar las actividades de la Gerencia.
18. Mantener actualizada la información en los sistemas informáticos implementados, en el ámbito de su competencia.

- 19.** Cumplir con las disposiciones contenidas en los procedimientos internos, así como los encargados legales asignados.
- 20.** Realizar las demás funciones que le asigne el gerente general.

ANEXO D: Entrevista de evaluación de resultados

Entrevistado: Secr.	Fecha:
Entrevistador: Alain Yucra Halanoca	Area: Área de Gerencia
Objetivo:	
Dirigido a: Alta dirección	
<p>Preguntas : Seleccione calificación</p> <p>1.- El desarrollo del DataMart, ¿Está cumpliendo los requisitos solicitados?</p> <p>A.- Deficiente B.- Malo C.- Regular D.- Buena E.- Excelente</p> <p>2.- En la toma de decisiones, ¿cómo califica usted la rapidez de respuesta de la herramienta a la hora de solicitar información?</p> <p>A.- Deficiente B.- Malo C.- Regular D.- Buena E.- Excelente</p> <p>3.- ¿Cómo es la interacción del Datamart con la solución de inteligencia de negocios?</p> <p>A.- Deficiente B.- Malo C.- Regular D.- Buena E.- Excelente</p> <p>4.- La solución de inteligencia de negocios ¿Mejora el proceso de decisiones?</p> <p>A.- Deficiente B.- Malo C.- Regular D.- Buena E.- Excelente</p>	

ANEXO E: Guía de Observación

Personal:		Lugar y Fecha:			
Evaluador:		Registro de cumplimiento			OBSERVACIONES
N°	Acciones a evaluar	SI	NO	NA	
1	Utiliza la solución de inteligencia de negocios en la toma de decisiones.				
2	La solución de inteligencia de negocios agiliza el proceso de toma de decisiones.				
3	La solución de inteligencia de negocios permite realizar búsquedas múltiples de información.				
4	La solución de inteligencia de negocios responde de manera adecuada ante las necesidades.				
5	La solución de inteligencia de negocios presenta una interfaz de fácil uso.				

6	son entendibles los resultados mostrados en la solución de inteligencia de negocios.				
7	La solución de negocios cumple con sus beneficios principales de rapidez e información confiable.				
8	Existen fallas en la solución de inteligencia de negocios, al momento de buscar la información requerida.				
9	Cuando se presentan fallas, el encargado del módulo de finanzas logra solucionar el problema.				