

ESCUELA SUPERIOR POLITECNICA DEL LITORAL



Facultad de Ingeniería en Electricidad y Computación

TRABAJO DE TITULACIÓN

**“IMPLEMENTACIÓN DE DATAMART Y PANELES DE INFORMACIÓN
DESTINADA A COMPLEMENTAR EL MÓDULO DE VENTAS DE UN
SISTEMA ERP DESARROLLADO POR UNA EMPRESA DE SOFTWARE
ECUATORIANA, USANDO HERRAMIENTAS DE INTELIGENCIA DE
NEGOCIOS.”**

Previa a la obtención del título de:

MAGISTER EN SISTEMAS DE INFORMACIÓN GERENCIAL

PRESENTADO POR:

LCDO. ANGEL PATRICIO JIMÉNEZ PROCEL

GUAYAQUIL – ECUADOR

2020

AGRADECIMIENTO

A Dios, quien provee de bendiciones cada día de nuestras vidas, en especial para con mi familia; por ese apoyo incondicional durante este camino de crecimiento intelectual y profesional.

A mi esposa e hijo, por la paciencia y comprensión de haber estado ausente todas esas horas de trabajo y estudio, que se verán recompensados con el fruto de nuestro trabajo y nuestro amor.

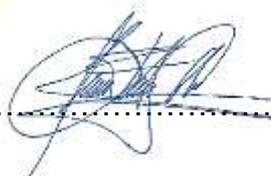
A todos quienes conforman este prestigioso establecimiento, que nos han brindado su ayuda y su conocimiento para hacer de nosotros; mejores personas y mejores profesionales.

A handwritten signature in blue ink, consisting of a large, stylized initial 'A' followed by a series of loops and a vertical stroke, likely representing the author's name.

DEDICATORIA

A Dios, mi creador, a mis padres; quienes cada día de mi vida están en mis pensamientos y a mi esposa e hijo; que son mi razón de vivir.

TRIBUNAL DE SUSTENTACIÓN



MSIG. Lenin Freire Cobo

DIRECTOR MSIG



MSIG. Omar Maldonado Dañín

DIRECTOR DEL PROYECTO DE GRADUACIÓN



MSIG. Ronny Santana Estrella

MIEMBRO DEL TRIBUNAL

RESUMEN

El objetivo principal de este proyecto fue la construcción de una solución de inteligencia de negocios que incluya la construcción de un Datamart y el diseño e implementación de paneles de información, para complementar el módulo de reportes de ventas de un ERP de una empresa desarrolladora de software ecuatoriana, con el fin de automatizar y agilizar el procesamiento de la información para sus clientes.

Para la planificación de este proyecto se inició con el inventario de los recursos disponibles dentro de la empresa, además de un repaso de literatura sobre los temas de inteligencia de negocios y Datamart que incluyen como componentes los recursos internos encontrados, con la finalidad de establecer los beneficios e identificar la subutilización de las herramientas existentes para darles un mejor uso.

Para determinar la necesidad de datos que se requieren visualizar y tener una solución que entregue información útil y valiosa a los usuarios, se contó con el apoyo de personal clave como son los clientes que usan el ERP; con quienes se mantuvo sesiones de trabajo para obtener la información que nos permita diseñar los paneles de información y el Datamart.

El resultado final para la empresa dueña del ERP debería poder evaluarse por el personal que la dirige y así poder observar la utilidad de esta herramienta que puede abrir nuevos mercados como parte de la cartera de servicios a ofrecer a sus clientes, y a su vez; estos poder disponer de un servicio que pueda ser de apoyo en la toma de decisiones mediante la automatización y consolidación de la información.

ÍNDICE GENERAL

AGRADECIMIENTO	i
DEDICATORIA	ii
TRIBUNAL DE SUSTENTACIÓN	iii
RESUMEN.....	iv
ÍNDICE GENERAL	vi
ABREVIATURAS Y SIMBOLIGÍAS.....	x
ÍNDICE DE FIGURAS.....	xii
ÍNDICE DE TABLAS.....	xiv
INTRODUCCIÓN.....	xv
CAPÍTULO 1 GENERALIDADES	1
1.1 Antecedentes.....	1
1.2 Descripción del Problema	4
1.3 Solución Propuesta.....	5
1.4 Objetivo general.....	7
1.5 Objetivos específicos.....	8
1.6 Metodología	8
1.6.1 FASE 1: Levantamiento de información.....	9
1.6.2 FASE 2: Análisis y definiciones de la solución	9
1.6.3 FASE 3: Diseño y construcción de la solución	10
1.6.4 FASE 4: Implementación	10
CAPÍTULO 2 MARCO TEÓRICO	12
2.1 Definición de términos técnicos	12
2.1.1 Datos.....	12

2.1.2	Información	13
2.1.3	Conocimiento	14
2.1.4	Query	15
2.1.5	Inteligencia de Negocios (BI)	15
2.2	Recursos de hardware y software a utilizar	17
2.2.1	AWS.....	17
2.2.2	Microsoft SQL Server.....	18
2.2.3	Microsoft SQL Server Integration Services	19
2.2.4	SQL Server Data Tools para visual studio	19
2.2.5	Power BI de Microsoft.....	20
2.3	Metodologías para construcción de un Datamart	21
2.3.1	Metodología Kimball.....	21
2.3.2	Metodología Inmon	22
2.4	Herramientas de inteligencia de negocios	23
2.4.1	Componentes de una solución de BI	23
2.4.2	Datawarehouse.....	24
2.4.3	Datamart	26
2.4.4	Diferencia entre Datawarehouse y Datamart	35
2.4.5	ETL	37
2.4.6	Sistemas de Información.....	41
CAPÍTULO 3 DEFINICIÓN DE LA SITUACIÓN ACTUAL Y REQUERIMIENTOS		45
3.1	Definición de la situación actual.....	45
3.1.1	Hardware	45
3.1.2	Software.....	46

3.1.3	Reportes disponibles.....	47
3.1.4	Fuentes de información.....	49
3.2	Levantamiento de información de procesos actuales	53
3.3	Definición de actores y casos de uso.....	55
3.4	Definición del proceso mejorado.....	58
3.5	Levantamiento de Requerimientos Funcionales	59
3.6	Alcance del Proyecto	61
CAPÍTULO 4 ANÁLISIS Y DISEÑO.....		63
4.1	Análisis de requerimientos.....	63
4.2	Arquitectura	64
4.3	Diseño del Datamart	65
4.3.1	Medidas	65
4.3.2	Dimensiones	66
4.4	Definición de métodos y extracción de información, construcción del ETL.....	67
4.4.1	Creación Datamart.....	67
4.4.2	Creación y preparación del proyecto SQL Integration Services	77
4.4.3	Creación del Flujo de datos base para el proceso ETL de las tablas de dimensión 82	
4.4.4	Creación del flujo de datos para proceso ETL de la dimensión Dim_Producto.....	84
4.4.5	Creación del flujo de datos para proceso ETL de la dimensión Dim_Vendedor	88
4.4.6	Creación del flujo de datos para proceso ETL de la dimensión Dim_Sucursal	90
4.4.7	Creación del flujo de datos para proceso ETL de la dimensión Dim_Tiempo	93
4.4.8	Creación del flujo de datos para proceso ETL de la dimensión Dim_Tipo.....	97
4.4.9	Creación del flujo de datos para proceso ETL de la tabla de hechos Fact_ventas ..	99
4.5	Diseño y creación de los paneles de información.....	104

4.5.1	Diseño de paneles	104
4.5.2	Creación de Paneles	106
CAPÍTULO 5 IMPLEMENTACIÓN Y ANÁLISIS DE RESULTADOS		117
5.1	Plan de pruebas	117
5.1.1	Pase para implementación en ambiente de pruebas	119
5.2	Resultados de las pruebas	123
5.3	Plan de implementación.....	129
5.4	Análisis de resultados	132
CONCLUSIONES Y RECOMENDACIONES.....		135
BIBLIOGRAFÍA.....		137

ABREVIATURAS Y SIMBOLIGÍAS

AS:	Analysis Services (Servicios de Análisis).
AWS:	Amazon Web Services (Servicios Web de Amazon).
BI:	Business Inteligence (Inteligencia de Negocios).
CEO:	Chief Executive Officer (Director Ejecutivo).
DBA:	Database Administrator (Administrador de Base de Datos).
DM:	Datamart.
ERP:	Enterprise Resource Planning (Planeamiento de los recursos empresariales).
ETL:	Extract, Transform and Load (Extracción, Transformación y Carga).
IBM:	International Business Machines Corporation.
IESS:	Instituto Ecuatoriano de Seguridad Social.
ISV:	Vendedor independiente de software.
OLAP:	On-Line Analytical Processing (Procesamiento Analítico en Línea).
OLTP:	On-Line Transaction Processing (Procesamiento de Transacciones en Línea).

QA:	Control de Calidad.
RS:	Servidor de Reportes.
SQL:	Structured Query Language (Lenguaje de consulta estructurada).
SSIS:	Server SQL Integration Services.
TI:	Tecnologías de la Información.

ÍNDICE DE FIGURAS

Figura 1.1: Solución propuesta	6
Figura 1.2: Metodología del proyecto	8
Figura 2.1: Metodología Kimball	22
Figura 2.2: Metodología Inmon	22
Figura 2.3: Componentes de una solución BI	23
Figura 2.4: Modelo Estrella	32
Figura 2.5: Modelo Copo de nieve	34
Figura 2.6: ETL	37
Figura 2.7: Tipos de sistemas de información	41
Figura 3.1: Módulos ERP	47
Figura 3.2: Reportes módulo ventas	48
Figura 3.3: Ejemplo de reporte	48
Figura 3.4: Proceso consulta reporte	54
Figura 3.5: Proceso reportes por un recurso externo	55
Figura 3.6: Proceso mejorado	58
Figura 4.1: Arquitectura	64
Figura 4.2: Modelo dimensional Datamart Ventas	67
Figura 4.3: Crear proyecto Integration Services (1)	78
Figura 4.4: Crear proyecto Integration Services (2)	78
Figura 4.5: Paquetes SSIS	79
Figura 4.6: ETL Definición Origen de datos DB Security	80
Figura 4.7: ETL Definición origen de datos	81
Figura 4.8: ETL Definición destino de datos	81
Figura 4.9: Tareas de flujo de datos de dimensiones	82
Figura 4.10: Estructura básica del flujo de datos	83

Figura 4.11: Flujo de datos Dim_Producto	85
Figura 4.12: Combinación mezcla Dim_Producto.....	86
Figura 4.13: Flujo de datos Dim_vendedor.....	88
Figura 4.14: Combinación de mezcla Dim_vendedor	89
Figura 4.15: Flujo de datos Dim_sucursal	91
Figura 4.16: Combinación de mezcla Dim_sucursal	92
Figura 4.17: Flujo de datos Dim_Tiempo.....	93
Figura 4.18: Combinación de mezcla dim_tiempo.....	95
Figura 4.19: Flujo de datos Dim_tipo	97
Figura 4.20: Combinación de mezcla Dim_tipo	98
Figura 4.21: Flujo de datos para Fact_ventas	100
Figura 4.22: Parámetros en Borrar Fact_ventas.....	100
Figura 4.23: Flujo de datos para Fact_ventas	101
Figura 4.24: Conexión Power BI Datamart	107
Figura 4.25: Campos en Power BI	107
Figura 4.26: Panel de ventas por categoría, local y vendedor	110
Figura 4.27: Panel ventas perdidas por producto y sucursal.....	112
Figura 4.28: Panel ABC Productos	114
Figura 4.29: Panel ABC rentabilidad.....	116
Figura 5.1: Implementación paquete SSIS en pruebas	120
Figura 5.2: Evidencia de cuadro de datos total de ventas	126
Figura 5.3: Evidencia de cuadro de datos clasificación	127
Figura 5.4: Evidencia cuadro de datos rentabilidad.....	127
Figura 5.5: Evidencia cuadro de datos, más vendidos	128
Figura 5.6: Comparativo más vendidos	129
Figura 5.7: Implementación paquete SSIS en producción	130

ÍNDICE DE TABLAS

Tabla 1: Productos de la empresa	3
Tabla 2: Diferencias entre Datawarehouse y Datamart	36
Tabla 3: Tablas y campos del ERP	49
Tabla 4: Actores	55
Tabla 5: CU-001 Caso de uso	57
Tabla 6: Requerimiento funcional del sistema	59
Tabla 7: Transformación de campos en Dim_tiempo	96
Tabla 8: Búsquedas para Fact_ventas	103
Tabla 9: Diseño de Paneles	105
Tabla 10: Campos en BI - Panel ventas por categorías, local y vendedor	108
Tabla 11: Campos en BI Panel Ventas perdidas por producto y sucursal	111
Tabla 12: Campos en BI Panel Curva ABC de productos	113
Tabla 13: Campos en BI Panel Curva ABC de rentabilidad	115
Tabla 14: Actores responsables en pruebas	117
Tabla 15: Archivos implementación para pruebas	120
Tabla 16: Cadena de conexión servidor de pruebas	121
Tabla 17: Job DM Ventas SSIS pruebas	122
Tabla 18: Job DM ventas SSIS pruebas 2	123
Tabla 19: Resultados de pruebas	124
Tabla 20: Archivos de implementación para producción	130
Tabla 21: Cadena de conexión servidor de producción	131
Tabla 22: Job DM Ventas SSIS producción	131

INTRODUCCIÓN

En la actualidad, donde la tecnología es una herramienta indispensable para el desarrollo de los negocios y el comercio se mueve cada vez más al rededor del mundo digital, surge como consecuencia un insumo que permite tener una mayor ventaja competitiva sobre las demás empresas y se trata de los datos; que tratados adecuadamente se convierten en una poderosa herramienta de apoyo en la toma de decisiones.

Una empresa ecuatoriana desarrolladora de software que es dueña de un ERP desarrollado con tecnología Microsoft destinado a las PYMES del mercado retail, aloja la aplicación y bases de datos en un servidor virtual en la nube teniendo como proveedor a la empresa Amazon, haciendo uso de sus servicios de hosting, potenciando así los beneficios de todos sus clientes que usan esta modalidad sobre su software.

Si bien el uso de la tecnología del ERP ha brindado gran ayuda a sus usuarios en mantener sus negocios dentro de la competencia del mercado, pero al pasar del tiempo es necesario ir evolucionando de acuerdo a las nuevas tecnologías, es así que para seguir manteniendo esa competencia surgen nuevas necesidades que agiliten aún más el proceso de toma de decisiones.

El tiempo que antes nos parecía adecuado en el proceso de tratamiento de los datos para obtener una información consolidada ya no es el mismo en la actualidad, debido al avance de la tecnología que cada día hace que estos procesos sean más cortos, haciendo que las empresas con mayores recursos para invertir en cada nueva tecnología que aparece en el mercado siempre esté a la vanguardia del mercado.

Las tecnologías de Datawarehouse sumadas a inteligencia de negocios o Business Intelligence BI, son una combinación que nos permite manejar grandes cantidades de datos de diferentes fuentes para ser transformados teniendo como resultado una información depurada, integra y organizada en un almacén, agilizando los procesos para el diseño y creación de reportes de tipo gerencial.

Esta empresa desarrolladora aun no hace uso del conocimiento y de los recursos para poder a disposición de sus clientes estas novedosas herramientas, como consecuencia aquellos que conocen de estas tecnologías buscan soluciones de terceros y aquellos que no, basan sus decisiones en la experiencia del día a día de sus negocios o simplemente no aplican estas técnicas de inteligencia de negocios.

CAPÍTULO 1

GENERALIDADES

1.1 Antecedentes

Una empresa ecuatoriana desarrolladora de software que brinda soluciones tecnológicas en las áreas farmacéutica, salud, y retail siendo una de las empresas más importantes en automatizar e implementar tecnología para el sector farmacéutico, lo que la ubica como pionera en esta área en el Ecuador. Como empresa se ha posicionado y dado a conocer por el desarrollo de su marca Neptuno, la cual contiene en sus versiones Neptuno Medical, Neptuno Retail y Neptuno Farmacias la mayor participación de sus ingresos, además; desarrolla software que están fuera de la marca como son GestorX, FactuDire y @pedidos.

Esta empresa no solo se ha limitado al mercado ecuatoriano y ha logrado ingresar con sus productos a otros países tales como: Panamá, Perú, Colombia, Venezuela, Bolivia y República Dominicana.

Tuvo sus inicios en el año 1995 y pasó a ser parte de un importante grupo de empresas con el fin de desarrollar y diseñar sistemas a la medida para las empresas que conformaban dicho grupo y así cubrir los requerimientos inmediatos para lograr posicionarse en el mercado; es así que como parte del crecimiento del grupo y la aparición de nuevas tecnologías, desde el 2004 pasan a hacer socios de negocios en Microsoft en el segmento ISV Solutions y en 2011 partner de IBM, ganando en el mismo año el premio como mejor socio de Microsoft en el sector de la salud en Latinoamérica y el Caribe. Bajo estas diversas asociaciones estratégicas que esta empresa ha realizado con las antes mencionadas, se ha centrado en desarrollar todas sus aplicaciones con tecnología Microsoft.

A partir del 2017 se separa del grupo de empresas al cual pertenecía, con la visión de ofrecer sus soluciones tecnológicas libremente a cualquier empresa que lo requiera, de esa manera impulsando mayormente la penetración en el área de retail con su sistema

Neptuno Retail, de salud con Neptuno Medical y facturación electrónica con FactuDires. A continuación, mencionaremos las principales características de sus productos:

Tabla 1: Productos de la empresa

PRODUCTOS DE LA EMPRESA	
PRODUCTO	CARACTERÍSTICAS
<p>Neptuno Farmacias</p> <p>Neptuno Retail</p>	<p>Desarrollado sobre tecnología Microsoft.Net y Microsoft SQL. Sistema Cliente/Servidor.</p> <ul style="list-style-type: none"> ○ Neptuno Cadena Modelo de Hosting. ○ Neptuno Cadena Infraestructura Propia. ○ Neptuno Standalone.
<p>Neptuno Medical</p>	<p>Desarrollado sobre tecnología Microsoft.Net y Microsoft SQL. Sistema Web.</p> <ul style="list-style-type: none"> ○ Neptuno Medical IESS. ○ Neptuno Medical Privados.
<p>GestorX</p>	<p>Desarrollado sobre tecnología Microsoft.Net y Microsoft SQL. Sistema Web, Cliente/Servidor y Móvil. Sistema creado para laboratorios y sus visitantes médicos.</p>
<p>@Pedidos</p>	<p>Desarrollado sobre tecnología Microsoft.Net y Microsoft SQL. Aplicativo Móvil utilizado para que los clientes hagan pedidos a sus proveedores.</p>

1.2 Descripción del Problema

Los clientes de una empresa ecuatoriana desarrolladora de software que usan su ERP, se ven limitados en la toma de decisiones y control de su negocio en el área de ventas debido a que el módulo actual de reportes que usan, no brinda información estadística e histórica de manera procesada, esto causa que no sean fácilmente interpretados ante la necesidad del cliente de realizar comparativos de costos, ganancias entre productos e incluso entre diferentes unidades de tiempo; por lo que necesitan pasar por un proceso de depuración y tratamiento de la información que efectúan terceras personas y luego poder presentarles este tipo de reportes estadísticos más acordes a sus necesidades de control y toma de decisiones, pero esto toma su tiempo ya que depende de la disponibilidad de aquellas personas y esperar hasta que tengan lista la información requerida. Además, por ser pequeñas empresas, no tienen los recursos necesarios para mantener una infraestructura propia que permita soportar personal experto destinado exclusivamente a la generación de esta información, ni los equipos de cómputo que cumplan con requisitos de hardware para alojar una aplicación de inteligencia de negocios como Microsoft PowerBI o Reporting Services. La empresa en la actualidad no ofrece un servicio de paneles de información para sus

clientes, lo que provoca que busquen alternativas con terceros afectando el rendimiento del ERP para los demás usuarios.

1.3 Solución Propuesta

Diseñar e implementar una solución de inteligencia de negocios estándar para los clientes que usan el módulo de ventas del ERP, que comprenderá la creación de un Datamart con su respectivo proceso ETL, utilizando los recursos que disponen en la infraestructura tecnológica, en servidores hosting y personal experto de la empresa desarrolladora de software a través de la metodología Kimball. Los paneles propuestos son:

- Ventas por categorías por local y vendedor.
- Oportunidades de ventas perdidas detallado por productos y por local.
- Curva ABC de productos (Top 10 de productos con mayor y menor rotación).
- Curva ABC de rentabilidad (Top 10 de productos vendidos con mayor y menor rentabilidad).

La empresa desarrolladora de software brindará a sus clientes la información correspondiente a sus ventas de manera rápida,

automática y fácil de interpretar, con la posibilidad de acceder a ella sin tener que contar con personal experto en inteligencia de negocios o tecnología para extraer, procesar y crear los paneles de información.

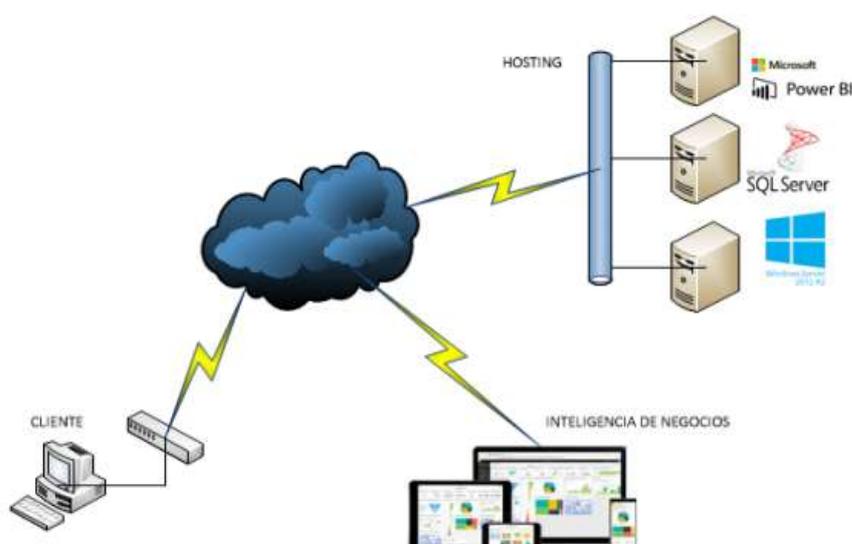


Figura 1.1: Solución propuesta

Dentro de los recursos existentes en la infraestructura actual que vamos a utilizar en esta propuesta están: uno de los servidores en Hosting con los componentes licenciados de Windows Server 2012 R2, que nos servirá como contenedor del Datamart, teniendo como motor de base de datos SQL SERVER 2017. El proceso de extracción, transformación y carga de datos se lo realizará con las herramientas de SQL Server Data Tools y finalmente para realizar la construcción y visualización de los reportes, se utilizará Reporting

Services y Microsoft Power BI. Todos los servidores hosting mantienen conectividad vía internet y enlaces de datos, con esto lograremos que los servicios estén disponibles en todo momento y desde cualquier lugar para los clientes. Los beneficios de centralizar la información y las operaciones en el Datamart dentro del hosting serán: poder trabajar de manera mucho más rápida y eficiente, facilitando el levantamiento de información, definición de requerimientos, diseño de paneles, implementación, mantenimiento, soporte y evaluación de resultados. Se contará con respaldo y disponibilidad de la información ante daños o desastres que dejen los equipos del cliente fuera de servicio. Además, de que la construcción de este Datamart puede ser replicado a los demás segmentos y clientes de esta empresa.

1.4 Objetivo general

Desarrollar una aplicación que mediante el uso de paneles de información, permita al área de ventas de las empresas que usan el ERP, optimizar el proceso de toma de decisiones.

1.5 Objetivos específicos

- Realizar el levantamiento de información para la construcción del Datamart, que es la base para la posterior implementación de los paneles de información.
- Analizar los datos disponibles a partir del levantamiento de información para la creación del Datamart y el respectivo proceso ETL.
- Diseñar un Datamart y un proceso ETL de acuerdo a la infraestructura actual de la empresa desarrolladora.
- Implementar un Datamart y un proceso ETL que sea funcional para los requerimientos actuales de los clientes, utilizando los recursos disponibles.
- Desarrollar paneles de información de acuerdo a las necesidades de los clientes y que ayuden a la toma efectiva de decisiones estratégicas.

1.6 Metodología

Para el desarrollo de este proyecto se manejarán las siguientes fases:



Figura 1.2: Metodología del proyecto

1.6.1 FASE 1: Levantamiento de información

- En esta fase se realiza el levantamiento de información en base a los requerimientos recibidos en mesa de ayuda, de parte de los usuarios sobre la información que necesitan en los reportes personalizados, y así crear un estándar que pueda ser utilizado por todos los clientes.
- Se realiza el inventario y un análisis de los recursos que dispone la empresa y que serán utilizados para este proyecto.
- Se establece parámetros básicos para elegir un cliente para realizar las pruebas piloto.
- Se elige un cliente que cumpla con los parámetros establecidos.

1.6.2 FASE 2: Análisis y definiciones de la solución

En esta fase se realizan los siguientes procesos:

- Se entrevista al cliente escogido para realizar las pruebas piloto con el fin de conocer su necesidad sobre los reportes actuales.
- Se analiza la información recabada en mesa de ayuda y entrevista con el cliente para determinar si todos los datos requeridos por los usuarios existen dentro de las bases de datos.
- Esquematación de los campos que deben contener los reportes:
 - Ventas por categorías por local y vendedor.

- Oportunidades de ventas perdidas detallado por productos y por local.
- Curva ABC de productos (Top 10 de productos con mayor y menor rotación).
- Curva ABC de rentabilidad (Top 10 de productos vendidos con mayor y menor rentabilidad).
- Definición de la metodología para la construcción de Datamart.
- Definición de los orígenes de los datos.

1.6.3 FASE 3: Diseño y construcción de la solución

- Diseño y construcción del proceso ETL y Datamart considerando los recursos que tiene la empresa dueña del ERP y los datos obtenidos de los análisis previos.
- Diseño y construcción de los paneles de información.
- Aprobación de los diseños.

1.6.4 FASE 4: Implementación

Esta fase comprende inicialmente el periodo de pruebas con el cliente que se designó para hacer el piloto, ejecutando los siguientes pasos:

- Configuración de ETL sobre base del cliente.

- Extracción de datos y creación de Datamart.
- Verificación de datos.
- Configuración de paneles de BI.
- Revisión de paneles de información con el cliente.

Una vez culminado este proceso se procede a realizar los ajustes que sean necesarios, en caso de tener observaciones en la revisión realizada con el cliente. Concluidas las revisiones finales, se procede a realizar la puesta en producción.

CAPÍTULO 2

MARCO TEÓRICO

2.1 Definición de términos técnicos

2.1.1 Datos

Un dato se define como la unidad mínima semántica o conjunto discreto de valores, que por sí solos pudieran no tener ninguna relevancia para un análisis, toma de decisiones o idea sobre algún hecho específico. También se puede indicar que corresponde al elemento primario de información que sin un contexto o propósito, no sirve para considerarlo en un proceso de toma de decisiones. [1]

Según Davenport y Prusak (1999): “Los datos describen únicamente una parte de lo que pasa en la realidad y no proporcionan juicios de valor o interpretaciones, y por lo tanto no son orientativos para la

acción. La toma de decisiones se basará en datos, pero estos nunca dirán que hacer. Los datos no dicen nada acerca de lo que es importante o no. A pesar de todo, los datos son importantes para las organizaciones, ya que son la base para la creación de información.”

[2]

2.1.2 Información

Conjunto de datos ordenados y asociados que, a diferencia de estos, cobran significado convirtiéndose en mensaje que puede ser interpretado, ya que contiene relevancia y propósito siendo útil para la toma de decisiones, debido a la capacidad de impactar sobre el juicio y comportamiento del receptor. Al añadir valor en varios sentidos a los datos, estos se convierten en información:

- Corrigiendo inconsistencias y eliminando errores.
- Resumiendo, de manera más concisa los datos.
- Categorizando, es decir conociendo las unidades de análisis de los componentes principales.
- Calculando: Los datos pueden haber sido procesados y analizados matemática o estadísticamente.
- Contextualizando: Saber para qué propósito y contexto se están utilizando los datos.

2.1.3 Conocimiento

El conocimiento es el producto posterior y surge a partir de la información. Para que la información cobre sentido y se convierta en conocimiento, debe ser interpretada y aplicada por un conocedor ya que es la mezcla de experiencia, la incorporación de la información recibida y el análisis, que permite identificar los elementos constituyentes de lo que se está conociendo y sus relaciones como partes del todo; esto se produce gracias a la búsqueda de conexiones de los datos y la información, comparación con otros elementos dentro del mismo contexto tratando de predecir sus consecuencias bajo determinados comportamientos. [3]

Los datos son la clave del negocio, pero estos están distribuidos a través de múltiples sistemas o fuentes de información, por lo que se necesita que estos datos sean procesados para convertirlos en información útil para la toma de decisiones, sin consumir demasiado tiempo ni expuestos a error que puedan distorsionar la realidad de los hechos que se está buscando; es por eso que es necesario ayudarse con el uso de herramientas y procesos de extracción y tratamiento de los datos, para convertirlos en información que brinde el conocimiento para la toma acertada de decisiones sobre la empresa.

2.1.4 Query

En un contexto de base de datos o inteligencia de negocios, es una consulta o búsqueda que se realiza en el lenguaje de la herramienta que se está utilizando, para extraer información almacenada en las tablas de una base de datos de una manera rápida.

2.1.5 Inteligencia de Negocios (BI)

Hace referencia al proceso de transformación de datos brutos que generan las empresas en todas sus áreas, en información que tiene significado y es útil para el análisis e interpretación, de esta forma pueden ser aprovechados en el proceso de toma de decisiones. La información que se obtiene puede tener distintos alcances, tales como:

- Nivel operativo: En este rubro es utilizada para la toma de decisiones diarias acerca de las transacciones que se realizan al llevar a cabo las operaciones de la empresa.
- Nivel táctico: Aporta información para los mandos medios en análisis y decisiones mensuales, que son de utilidad para revisiones de seguimiento y toma de acciones.

- Nivel estratégico: A este nivel las decisiones son de mayor impacto en la compañía, siendo utilizada la información por la alta dirección.

Beneficios

- Incremento de la eficiencia: Al contar con los datos de manera accesible y ágil, se puede generar información de valor centralizada la cual se podrá visualizar en una única plataforma para aprovecharla de manera óptima para realizar análisis y tomar decisiones informadas y a tiempo.
- Respuestas rápidas a situaciones de negocio: Para poder tomar decisiones en el momento indicado es importante contar con la información a la mano de manera sencilla y no perder tiempo en buscar y consolidar datos. Gracias al BI, se puede tener las respuestas en minutos de manera clara y concisa por medio de reportes de indicadores y tableros de datos.
- Control de las áreas funcionales de la empresa: En todas las áreas de la empresa se genera información de valor día a día, se puede aprovechar de la mejor manera para conocer tendencias, proyectar datos, analizar escenarios, etc.
- Mejorar el servicio al cliente: Al contar con la información más importante y en tiempo real, se puede ofrecer a los clientes un servicio de mayor calidad desde el pedido hasta el servicio post-

venta al conocer más acerca de ellos y sus necesidades. Se puede analizar hábitos de compra, reconocer los productos más vendidos, etc.

- Presenta información por medio de tableros de indicadores para una comunicación más simple y directa de la situación de la empresa. Al tener la posibilidad de crear distintos tableros para control, se puede enfocar en los datos más relevantes en vez de mostrar sin necesidad, grandes cantidades de información. [4]

2.2 Recursos de hardware y software a utilizar

2.2.1 AWS

Por sus siglas en inglés, Amazon Web Services; ofrece servicios de hospedaje web en la nube a diferentes tipos de organizaciones con y sin fines de lucro. Esos servicios pueden ser desde infraestructura de cómputo, almacenamiento, bases de datos, hasta elementos tecnológicos más complejos como inteligencia artificial o análisis de internet de las cosas. [5]

Para este proyecto se utiliza una de las máquinas virtuales a la que Amazon llama instancia, que se mantiene en arrendamiento en este servicio en la que se alojan las bases de datos y aplicación donde se conecta un grupo de clientes del segmento retail.

Esta instancia tiene entre sus características principales:

- Sistema Operativo Windows Server 2012 R2.
- 16 GB en memoria ram.
- CPU.
- 500 gb almacenamiento.
- 750 Mbps de ancho de banda.

Entre los beneficios de utilizar este tipo de tecnología están:

- Bajo Costo.
- Agilidad y elasticidad instantánea.
- Accesibilidad y flexibilidad.
- Seguridad.

2.2.2 Microsoft SQL Server

Es un sistema desarrollado por la empresa Microsoft que sirve para la gestión de bases de datos relacionales, el cual se ejecuta bajo el lenguaje llamado transact-sql. Las características principales son:

- Permite trabajar en modo cliente-servidor.
- Además, permite administrar información de otros servidores de datos.
- Soporte de transacciones.

- Escalabilidad, estabilidad y seguridad.
- Soporta procedimientos almacenados.

2.2.3 Microsoft SQL Server Integration Services

Es una plataforma para construir soluciones de integración y transformación de datos a nivel empresarial. Resuelve problemas comerciales complejos copiando o descargando archivos, cargando almacenes de datos, limpiando y minando datos, y administrando objetos y datos de SQL Server.

Integration Services ayuda a la extracción y transformación de datos desde diversos orígenes, tales como bases de datos, archivos de tipo XML o planos entre otros, para luego ser depositados en uno o más destinos. [6]

2.2.4 SQL Server Data Tools para visual studio

Es un software de desarrollo útil para gestionar bases de datos relacionales de SQL Server, SQL de Azure, además de paquetes de IS, modelos de datos de AS e informes de (RS).

Con la ayuda de SSDT, se puede estructurar y poner en marcha gran variedad de contenido para SQL Server teniendo las mismas ventajas que se puede obtener al desarrollar aplicaciones en Visual Studio. [7]

2.2.5 Power BI de Microsoft

Power BI es una solución de análisis empresarial que permite visualizar los datos y compartir información con toda la organización, o insertarla en su aplicación o sitio web. Permite conectarse a cientos de orígenes de datos y dar vida a los datos con los paneles e informes dinámicos. [8]

Entre sus beneficios más importantes están:

- Conectar datos, estén donde estén para luego poder explorarlos con impresionantes visualizaciones interactivas.
- Publicar informes y paneles, además de colaborar con un equipo y compartir información dentro y fuera de la organización.
- Acceder sin problemas a la información en su oficina como fuera de ella con las aplicaciones de Power BI creadas por Microsoft.

2.3 Metodologías para construcción de un Datamart

Se analizaron dos metodologías para la construcción de una solución de inteligencia de negocios como son la de Ralph Kimball y Bill Inmon.

2.3.1 Metodología Kimball

Para Ralph Kimball un Datawarehouse se explica como “el enlace de diferentes Datamarts de una o varias organizaciones siguiendo una estructura definida previamente para sus dimensiones, atributos, organización jerárquica, hechos de negocio establecidos que se desea analizar. Estos Datamarts, se unen por lo que se define como bus structure, lo que permite poder ejecutar consultas unificadas sobre los diferentes Datamarts.

Bajo este criterio se contempla que, al interior de un Datawarehouse, se unen los diferentes Datamarts enlazados por bus structure. A esto se le denomina Bottom-up.” [9]

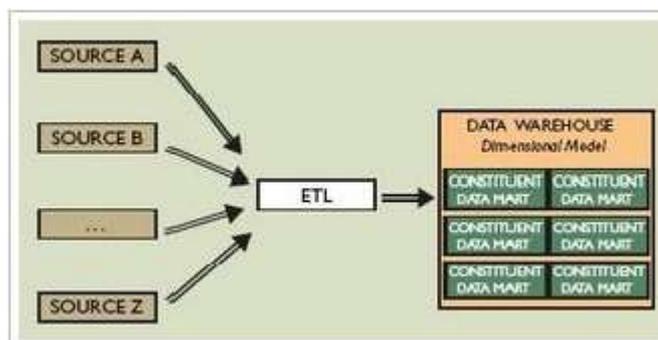


Figura 2.1: Metodología Kimball
Fuente: <http://mundobi.com.ar/?p=614>

2.3.2 Metodología Inmon

Bill Inmon en cambio presenta la necesidad de transferir toda la información de los diferentes Sistemas Transaccionales de las organizaciones a un repositorio común donde los datos puedan ser utilizados para el análisis con características como: no volátil, integrado, histórico y orientado a temas. [10]

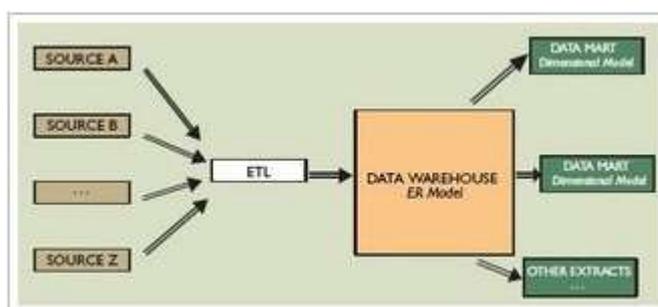


Figura 2.2: Metodología Inmon
Fuente: <http://mundobi.com.ar/?p=614>

Otra definición que puede ser aplicada es que un Datawarehouse se describe como un almacén, que concentra los datos de una organización los cuales han sido extraídos de diferentes fuentes

pasando por un proceso de tratamiento y dando como resultado que estos datos están ordenados, libres de errores e inconsistencias listos para ser convertidos en información y conocimiento.

2.4 Herramientas de inteligencia de negocios

Las herramientas de inteligencia de negocios permiten a las empresas extraer, procesar, visualizar, analizar, y compartir información clave para la organización de una manera automática y rápida.

2.4.1 Componentes de una solución de BI

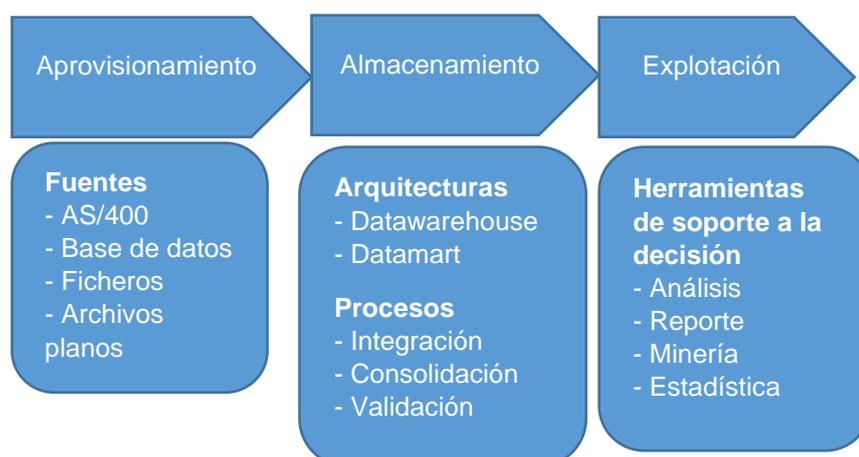


Figura 2.3: Componentes de una solución BI

2.4.2 Datawarehouse

Un Datawarehouse es un repositorio o almacén unificado para todos los datos que recogen los diversos sistemas de una empresa. El repositorio puede ser físico o lógico y hace hincapié en la integración y captura de datos de diversas fuentes sobre todo para fines analíticos y de acceso, mediante reportes; considerándose como un componente fundamental de la inteligencia de negocios. Normalmente, un Datawarehouse se aloja en un servidor corporativo o en la actualidad cada vez más común en la nube. Los datos de diferentes aplicaciones de procesamiento de transacciones Online (OLTP) y otras fuentes, se extraen selectivamente para su uso por aplicaciones analíticas y de consultas por usuarios. Datawarehouse es una arquitectura de almacenamiento de datos que permite a los ejecutivos de negocios organizar, comprender y utilizar sus datos para tomar decisiones estratégicas. [11]

Según Oracle: “Un Datawarehouse es una colección de datos orientada a temas específicos, integrada, no volátil, variante en el tiempo y organizada para apoyar las necesidades de administración”

Según Informix: “Un Datawarehouse no es un producto que puede ser comprado en el mercado, sino más bien un concepto que debe ser constituido”.

Beneficios

Debido al gran valor que poseen los datos, el manejo de estos se vuelve vital para el desarrollo de las empresas, es por esto que la correcta gestión y análisis muchas de las veces determinan el camino al éxito.

- **Ventaja competitiva.** - Bajo el análisis de los datos se puede identificar y predecir patrones de compra o tendencias, que harán que se pueda estar un paso delante sobre los competidores, creando estrategias que involucren cualquier área de la empresa, además que pueden brindar apoyo y aumentar el valor en las gestiones de relaciones con los clientes y proveedores.
- **Fácil acceso.** - Los datos se encuentran dispersos en diferentes sistemas, fuentes o departamentos de la empresa por lo que tener un solo repositorio con datos integrados y precisos, significa mejor y más rápido acceso y comunicación entre todas las partes involucradas, además de la flexibilidad de su uso de acuerdo a los requerimientos de cada uno de los actores.

- Mejor toma de decisiones. - Al tener la información procesada, ordenada y consistente, se cuenta con análisis más precisos lo que respalda y facilita la toma de decisiones que beneficien a la empresa.
- Aumenta productividad de la empresa. - Las relaciones con los clientes y proveedores mejoran, porque se puede predecir tendencias de consumos, entregándole al cliente lo que desea y demás productos o servicios que complementan sus necesidades, además de poder pronosticar rotaciones de inventarios afinando las políticas de abastecimiento y distribución para la empresa. Como punto adicional, al ser accesible y flexible la información, todos los departamentos dentro de la empresa podrán tener la información en corto tiempo haciendo que sus actividades sean más productivas. [12]

2.4.3 Datamart

Un Datamart se puede definir como un subconjunto de datos o base de datos departamental que pueden formar parte de un Datawarehouse, pero en este caso orientado a un área o departamento más específico de una empresa. Los Datamarts se caracterizan por tener una estructura óptima de datos que permite el análisis de la información desde cualquiera de las perspectivas de

los procesos de esa área. El acceso a la información del Datamart es limitado a ciertos usuarios del área indicada para cumplir con un propósito y función específica. Un Datamart puede ser alimentado externamente por un Datawarehouse o alimentarse de distintas fuentes de información, que es lo que Kimball denomina bus structure. A continuación, se detallan algunas ventajas de los Datamarts:

- Ahorro de espacio debido al poco volumen de datos.
- Menor tiempo en consultas.
- Validación directa de la información.
- Facilidad para el manejo de datos históricos.

OLAP. - Diseñado para reflejar la gestión. Las Herramientas OLAP se basan en presentar cubos que se construyen agregando las dimensiones e indicadores según los requerimientos de cada departamento; en síntesis, ofrece la misma información desde una visión multidimensional.

Características:

- Datos con criterio de agrupación.
- Datos normalizados y no normalizados.
- Contiene históricos.

- Estable.
- Consultas planificadas.
- Modelo multidimensional.
- Datos estáticos.

OLTP. - Diseñada para reflejar operaciones. Pueden tener como referencia un resumen del Datawarehouse, pero, normalmente se realiza mejoras en su desempeño (los filtros y nuevas asignaciones son las acciones más comunes) sacando provecho de los atributos propios de cada una de las áreas de la empresa. Las tablas report son normalmente las estructuras más frecuentes, que vienen a ser pequeñas fact-tables (que añaden las dimensiones necesarias), y las vistas materializadas, que a su vez se arman a partir de la misma estructura, teniendo como nuevo objetivo de aprovechar la nueva adaptación de un query. [13]

Características:

- Datos al máximo detalle.
- Datos normalizados.
- Poco historial.
- Volátil.
- Consultas continuas y no planificadas.

- Modelo entidad – relación.
- Datos dinámicos.

Beneficios

Su implementación es esencial para que una empresa pueda integrar y combinar los datos obtenidos desde las diferentes fuentes con el fin de descubrir oportunidades y amenazas que no son visibles en las bases de datos operacionales.

- Es un procedimiento de rápida resolución que ayuda a la consulta de información previo a realizar cambios en un ambiente empresarial.
- La producción de datos canalizados a perfiles de negocio. Debido a la información almacenada de manera sencilla, la transmisión de datos es rápida y óptima, asimismo útil para el proceso de toma de decisiones vitales para la organización.
- Disminuye las tareas de trabajo del personal. Los empleados pueden acceder a la información de manera más ágil a través de consultas, en cualquier momento y lugar.
- Los profesionales que tienen a cargo la administración de la información pueden utilizarla para mejorar los resultados en el uso de herramientas analíticas.[14]

Tipos

Se describen las siguientes clases de Datamarts: dependientes, independientes e híbridos en base a relación con el Datawarehouse y las fuentes de datos que se emplean para construir el sistema.

- **Datamarts dependientes.** - Datamart dependiente se construye a partir de un Datawarehouse empresarial que ya existe. Es una visión que implica iniciar desde arriba a abajo acumulando todos los datos de las operaciones de la empresa en una ubicación para luego extraer los datos que se necesitan de acuerdo a los requerimientos de las áreas que realizan su análisis.

Para crear un almacén de datos, se añade un grupo de datos concretos; a partir de este almacén, los datos se ordenan y se tratan para luego ser cargados al Datamart, en el cual se realizan las consultas. Pueden ser un subconjunto físico del almacén o una visión lógica:

- Subconjunto físico: Es la extracción de datos desde el origen para construir una base de datos apartada físicamente.
- Visión lógica: Es una tabla o vista virtual apartada lógicamente (no físicamente) del almacén de datos.

Los datos granulares del almacén de datos son la única relación que sirven para crear todos los Datamarts dependientes necesarios.

- **Datamarts independientes.** - Se define como un sistema autónomo, sin utilizar ningún almacén para ser creado y tiene como objetivo un departamento o área de la empresa. La extracción de los datos se produce desde las fuentes de origen para luego ser tratados y cargados al Datamart donde permanecen hasta ser requeridos para el análisis deseado. Estos Datamarts son fáciles de diseñar y desarrollar, ideales para objetivos a corto plazo, pero al crecer las necesidades de la empresa pueden ocasionar que la gestión se torne confusa y complicada cuando se tiene recursos que no conocen de las estructuras iniciales del Datamart o a su vez generar cambios para generar otro tipo de información.
- **Datamarts híbridos.** – Este tipo de Datamart combina para su creación datos existentes en un almacén con otras fuentes de datos operativas. Además de enfocar todos los beneficios hacia el usuario final también mantiene las ventajas de utilizar el método de arriba hacia abajo. [15]

Modelos de Datamart

- Modelo Estrella.** - Toma su nombre porque se basa en una tabla de hechos central donde se representan las medidas y a su alrededor están asociadas con las tablas de dimensiones y cada dimensión consta de una sola tabla, manteniendo siempre como eje central la tabla de hechos y las tablas de dimensiones no se relacionan entre sí, es por esto que se lo conoce como multidimensional sin jerarquías.

En este modelo, no es necesario que el usuario conozca el modelo entidad relación de las bases transaccionales.



Figura 2.4: Modelo Estrella

Fuente: <http://mundodb.es/disenio-data-warehouse-hechos-y-dimensiones-modelo-estrella-vs-copo-de-nieve>

En la Figura 2.4 se observa la tabla central llamada ventas, que es la tabla de hechos que contiene las medidas o indicadores del negocio; y su alrededor las tablas producto, tienda, tiempo y cliente que vienen a ser las dimensiones que aportan con información a la tabla central.

- **Modelo Copo de nieve.** - Tiene el mismo concepto del modelo estrella, pero con una estructura un poco más compleja debido a que una dimensión puede estar conformada o se normaliza por varias tablas o dimensiones, por eso se puede definir como multidimensional con jerarquías. A diferencia del modelo estrella, la tabla de hechos deja de ser la única que se relaciona con otras tablas, este modelo fue creado para facilitar el mantenimiento de las dimensiones, pero hace que la estructura sea más compleja y el proceso de extracción más laborioso. En este modelo si es necesario que el usuario conozca el modelo entidad relación en las bases transaccionales. [16]

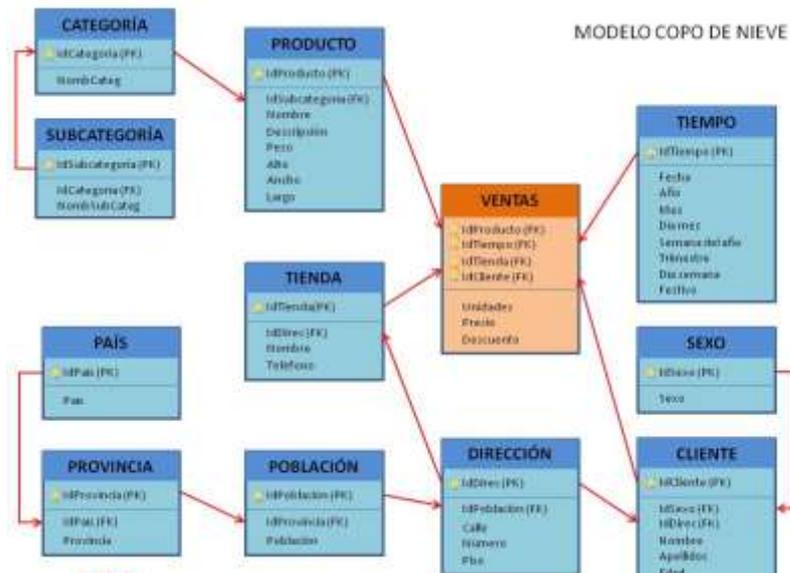


Figura 2.5: Modelo Copo de nieve

Fuente: <http://mundodb.es/disenio-data-warehouse-hechos-y-dimensiones-modelo-estrella-vs-copo-de-nieve>

En la Figura 2.5 se observa como la tabla central de hechos (Ventas) no es la única que se relaciona con más de una tabla de dimensiones y las tablas de dimensiones comienzan a relacionarse unas con otras, como ejemplo la dimensión cliente, no solo se relaciona con la tabla de hechos ventas, sino con las dimensiones sexo y dirección directamente.

Tabla de hechos

Se denomina hechos a los indicadores de negocios. Estos indicadores son datos cuantitativos que son filtrados y agrupados por las condiciones definidas en las tablas de dimensiones.

Es la tabla central del modelo dimensional que contienen los datos de estudios que serán el objeto de análisis o también llamados hechos o medidas, además contienen las respectivas relaciones o claves primarias que las unen con las tablas de dimensiones.

Medidas

Datos que se buscan medir o analizar y que generalmente son representados por valores numéricos.

Tabla de dimensiones

Estas tablas desnormalizadas contienen datos cualitativos, que se unen a la tabla de hechos mediante los campos claves. Los datos que contienen estas tablas son atributos o características que sirven para filtrar o agrupar datos en la tabla de hechos además de complementar los campos de medidas para otorgar mayor significado a la información que se está analizando.

Estas tablas pueden tener más de una relación adicional a la tabla de hechos; pueden asociarse de manera jerárquica con otras tablas de dimensiones.

2.4.4 Diferencia entre Datawarehouse y Datamart

Ambos coinciden en definirse como un almacén de datos que los integra y procesa con el fin de obtener información que ayuda a tomar

decisiones estratégicas, pero la diferencia radica en que un Datawarehouse almacena los datos de todas las áreas de una empresa haciendo su estructura más compleja, en cambio; un Datamart se centra en un departamento específico lo que se consideraría una estructura más sencilla. En la metodología Kimball un Datawarehouse es conformado por varios Datamarts.

Los Datawarehouse al almacenar los datos de toda una organización pueden ofrecer información a cualquier tipo de información a los diferentes departamentos en cambio un Datamart, ofrece información específica a un selecto grupo de usuarios.

Al ser más complejo en su estructura un Datawarehouse demanda más tiempo para su análisis, diseño e implementación que un Datamart.

Resumimos las diferencias en la siguiente tabla:

Tabla 2: Diferencias entre Datawarehouse y Datamart

DIFERENCIAS ENTRE DATAWAREHOUSE Y DATAMART		
Aspecto	Datawarehouse	Datamart
Alcance	Corporativo	Departamental
Temática	Todas las área	Una sola área

DIFERENCIAS ENTRE DATAWAREHOUSE Y DATAMART		
Fuente de datos	Muchas	Pocas
Tamaño	Comúnmente Terabytes	Comúnmente Gigabyte
Implementación	Proceso largo	Proceso mucho más corto
Tipo de información	Detallada	Más Resumida
Inversión	Alta	Baja

2.4.5 ETL

Por sus siglas en inglés (Extract, Transform, Load), es el proceso de extracción, transformación y carga de los datos desde las diferentes fuentes hacia el almacén donde los datos pasan a ser homogéneos. Este proceso es de gran importancia ya que permitirá obtener una vista consolidada de la información para el apoyo en la toma de decisiones empresariales satisfaciendo los requisitos de los usuarios y mejorando la productividad del personal, lo que se convierte en beneficios para la empresa.



Figura 2.6: ETL

Fuente: <https://es.talend.com/resources/what-is-etl/>

Extracción

El objetivo del proceso ETL es tener como resultado final datos limpios y accesibles que permitan ser utilizados en el análisis y operaciones de la empresa. En esta primera fase los datos brutos son consultados y extraídos de todas las fuentes disponibles, por ejemplo:

- Bases de datos.
- Archivos de datos.
- Registros de actividades como logs.

Una vez extraídos los datos, se deben analizar e interpretar verificando que los datos cumplen con la estructura esperada para luego convertirlos en un formato ideal para iniciar el proceso de transformación.

Un punto importante a tomar en cuenta es que este proceso cause el menor impacto posible en los sistemas orígenes de extracción, debido a que generalmente se lo realiza sobre los sistemas transaccionales de las empresas; con el fin de no saturar o afectar el

rendimiento normal de estos sistemas, estas tareas se llevan a cabo en horarios y días de menor impacto. [17]

Transformación

En esta fase se busca transformar los datos obtenidos en la anterior etapa, aplicando una serie de reglas o funciones del negocio para luego proceder a la carga; se realiza una limpieza eliminando cualquier tipo de error, inconsistencia o duplicidad en los datos. Este es el proceso que más tiempo puede demorar debido a que hay que asegurarse que los datos estén depurados, para que sean eficientes y útiles para el proceso que se los requiere.

En este proceso además pueden llevarse a cabo diferentes acciones con el fin de prepararlos para realizar la carga como, por ejemplo, realizar cálculos, unir datos de diferentes fuentes, codificar valores o traducir códigos, etc.

Carga

En esta fase se procede a cargar los datos transformados en la fase anterior en el sistema destino.

Existen muchas maneras de realizar este proceso, pero siempre se debe tener en cuenta que se está interactuando con la base de datos de destino y, por eso, durante esta operación se debe pasar por cada uno de los filtros con las respectivas restricciones que se hayan definido. Si todas estas reglas que se tiene como restricciones están correctamente aplicadas, todo el proceso y la veracidad de los datos está garantizado. [18]

Algunos ejemplos definidos como restricciones pueden ser:

- No exista duplicidad de valores.
- Integridad referencial.
- Datos necesarios/obligatorios.
- Rangos establecidos de valores.

2.4.6 Sistemas de Información



Figura 2.7: Tipos de sistemas de información

TPS: Sistema de Procesamiento de Datos

Ayudan a la operación del día a día.

Características:

- Nivel operacional.
- Entrada: transacciones, eventos.
- Procesos: actualizaciones.
- Salida: reportes detallados con totales de control.
- Usuarios: personal y supervisores.

KMS: Sistema de Trabajo del Conocimiento

Los KMS juegan un rol esencial en ayudar a las compañías a compartir y diseminar conocimiento para la toma de decisiones del negocio por parte de los gerentes y los profesionales del negocio.

- Usan las TI para ayudar a compilar, organizar y compartir conocimiento del negocio dentro de una organización.
- Algunas bases de datos hipermedias corporativas han venido a ser las bases de conocimiento para el almacenamiento y la diseminación del conocimiento del negocio.
- Para muchas compañías, los Portales de Conocimiento Empresarial (EKP, por las siglas en inglés de Enterprise Knowledge Portals) son la entrada a las intranets corporativas que sirven como sus KMS.

DSS: Sistema de Apoyo a la Toma de Decisiones

Apoyan la realización del análisis de gestión. Procesamiento analítico en línea - Online Analytical Processing (OLAP).

Características:

- Nivel gerencial.
- Entrada: Alto volumen de datos.
- Salida:
 - OLAP: Gráficos y hojas electrónicas.

- Dataminig: Secuencias, clasificaciones y grupos.
- Usuarios: Profesionales, analistas y especialistas.

MIS: Sistema de Información Gerencial

Decisiones estructuradas en base a información interna y reportes de control. Comparación de datos pasados y presentes.

Características:

- Nivel gerencial.
- Entrada: Volumen alto de datos.
- Proceso: Modelos simples.
- Salidas: Reportes y gráficas de resumen.
- Usuarios: Mandos medios.

ESS: Sistema de Apoyo a Ejecutivos (EIS)

Características:

- Nivel: Alta gerencia.
- Entrada: Datos agregados.
- Internos: Aplicación MIS o ERP y DSS.
- Externos: Internet y proveedores de información.
- Proceso: Interactivo.
- Salida: Alarmas, proyecciones y gráficos.
- Usuarios: Gerentes de máximo nivel.

- Diseñado para cada organización.
- Enlaza al ejecutivo de más alto nivel (CEO) con toda la organización.
- La implantación requiere un equipo de apoyo de alto nivel profesional.

CAPÍTULO 3

DEFINICIÓN DE LA SITUACIÓN ACTUAL Y REQUERIMIENTOS

3.1 Definición de la situación actual

3.1.1 Hardware

En la actualidad la empresa desarrolladora de software dispone de una instancia AWS en arrendamiento con el proveedor Amazon. Las características de esta instancia son las siguientes: m4.xlarge con 4 CPU virtuales Intel Xeon E5-2686 v4 de 2.4 Ghz y 16 GB de memoria, además de un ancho de banda de 750 Mbps. El espacio de disco duro es variable de acuerdo al número de clientes activos. Es en esta instancia que se alojan las bases de datos de los clientes que mantienen alquiler de hosting para el uso del sistema ERP. La conectividad por parte de los clientes se da vía internet, mediante los respectivos servicios y puertos publicados.

Por parte del cliente, los requisitos mínimos de hardware que se exige son:

- Windows 8 o superior.
- Procesador core i3 o superior.
- Memoria Ram 4 GB o superior.

3.1.2 Software

El ERP, el cual usan los usuarios, es un software de tipo cliente – servidor que está construido con tecnología Microsoft.Net y Microsoft SQL y que ofrece los módulos de: Ventas, Compras, Planificación Comercial, Inventario, Bodegas, Cuentas por Cobrar, Cuentas por Pagar, Promociones, Contabilidad, Caja y Bancos, Impuestos y Parámetros Generales.

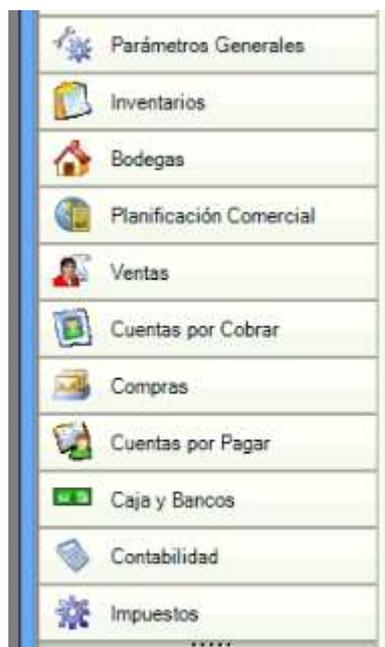


Figura 3.1: Módulos ERP

Los usuarios realizan todas sus operaciones en línea directamente hacia el hosting, es así que obtienen cualquier reporte dentro de cada módulo de manera instantánea.

3.1.3 Reportes disponibles

Dentro del ERP cada módulo contiene una carpeta exclusiva de reportes con información que le corresponde, en este proyecto se analizan únicamente los reportes dentro del módulo de ventas. La empresa desarrolladora de software no personaliza reportes, ya que siguen un formato de estandarización el cual es funcional para todos

sus clientes. Los usuarios tienen disponibles los siguientes reportes en el módulo de ventas:

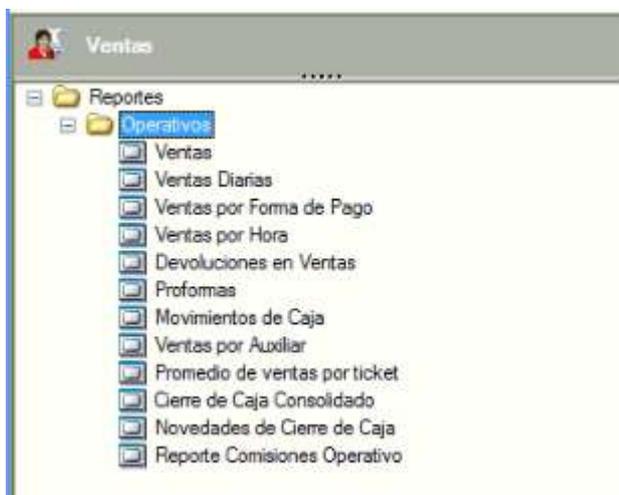


Figura 3.2: Reportes módulo ventas

Estos reportes se visualizan en el siguiente formato, los cuales contienen un promedio de 23 columnas de información:

RESUMEN DE FACTURAS DE VENTAS																							
Criterios:																							
Fecha	No Docum.	Tip	#	Fecha	Cliente	C.E./	Parcial	Dom.	Imp.	Desc.	Subt.	Desc.	Tar.	Tar.	Tar.	Imp-Excto.	Total	Costo	Costo	Util.	% Util.	Estado	Cre En
			Autoriza	Validez	Validez	RELC			MC			MC	0%	15%	30%	adicionales			Base.				
Estado: EMITIDAS																							
2014/02/11	000000000	NCTSA	1247507	18/11/2013	ALFA ROMEO	0000000	22.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	22.6	0.0	0.0	0.0	0.0	0.0	0.0
2014/02/12	000000000	NCTSA	1247508	18/11/2013	ALFA ROMEO	0000000	22.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	22.4	0.0	0.0	0.0	0.0	0.0	0.0
2014/02/17	000000000	NCTSA	1247509	18/11/2013	ALFA ROMEO	0000000	75.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	75.2	0.0	0.0	0.0	0.0	0.0	0.0
2014/02/17	000000000	NCTSA	1247510	18/11/2013	ALFA ROMEO	0000000	21.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	21.8	0.0	0.0	0.0	0.0	0.0	0.0
2014/02/18	000000000	NCTSA	1247511	18/11/2013	ALFA ROMEO	0000000	80.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	80.0	0.0	0.0	0.0	0.0	0.0	0.0
2014/02/22	000000000	NCTSA	1247512	18/11/2013	ALFA ROMEO	0000000	19.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	19.8	0.0	0.0	0.0	0.0	0.0	0.0
2014/02/24	000000000	NCTSA	1247513	18/11/2013	ALFA ROMEO	0000000	31.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	31.4	0.0	0.0	0.0	0.0	0.0	0.0
2014/02/24	000000000	NCTSA	1247514	18/11/2013	ALFA ROMEO	0000000	20.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	20.0	0.0	0.0	0.0	0.0	0.0	0.0
2014/02/27	000000000	NCTSA	1247515	18/11/2013	ALFA ROMEO	0000000	97.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	97.5	0.0	0.0	0.0	0.0	0.0	0.0
2014/02/28	000000000	NCTSA	1247516	18/11/2013	ALFA ROMEO	0000000	22.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	22.0	0.0	0.0	0.0	0.0	0.0	0.0
2014/02/28	000000000	NCTSA	1247517	18/11/2013	ALFA ROMEO	0000000	22.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	22.0	0.0	0.0	0.0	0.0	0.0	0.0
2014/02/28	000000000	NCTSA	1247518	18/11/2013	ALFA ROMEO	0000000	22.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	22.0	0.0	0.0	0.0	0.0	0.0	0.0
2014/02/28	000000000	NCTSA	1247519	18/11/2013	ALFA ROMEO	0000000	62.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	62.0	0.0	0.0	0.0	0.0	0.0	0.0
2014/02/19	000000000	NCTSA	1247520	18/11/2013	ALFA ROMEO	0000000	42.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	42.0	0.0	0.0	0.0	0.0	0.0	0.0
2014/02/24	000000000	NCTSA	1247521	18/11/2013	ALFA ROMEO	0000000	4.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4.8	0.0	0.0	0.0	0.0	0.0	0.0
2014/02/24	000000000	NCTSA	1247522	18/11/2013	ALFA ROMEO	0000000	22.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	22.0	0.0	0.0	0.0	0.0	0.0	0.0
2014/02/24	000000000	NCTSA	1247523	18/11/2013	ALFA ROMEO	0000000	22.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	22.0	0.0	0.0	0.0	0.0	0.0	0.0
2014/02/24	000000000	NCTSA	1247524	18/11/2013	ALFA ROMEO	0000000	22.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	22.0	0.0	0.0	0.0	0.0	0.0	0.0

Figura 3.3: Ejemplo de reporte

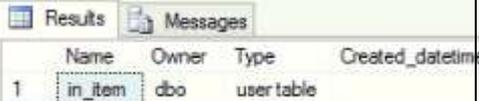
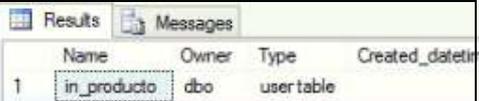
3.1.4 Fuentes de información

Las fuentes de información se obtendrán desde las tablas del ERP donde se producen todas las operaciones transaccionales. Los clientes realizan todos sus movimientos dentro del ERP, por lo que no existirán fuentes adicionales.

Tabla 3: Tablas y campos del ERP

TABLAS Y CAMPOS A USAR DEL ERP	
Tablas ERP	Campos
ve_venta_cab	 <ul style="list-style-type: none"> • id_venta_cab • id_numero_doc • id_bodega • fecha • descripcion • total_parcial_0 • total_parcial_1 • total_descto_0 • total_descto_1 • total_iva_0 • total_iva_1 • total_netto_0 • total_netto_1 • costo_total_0 • costo_total_1 • total_ice

TABLAS Y CAMPOS A USAR DEL ERP	
Tablas ERP	Campos
	<ul style="list-style-type: none"> • fecha_entero • estado_venta • vendedor
ve_venta_det	 <ul style="list-style-type: none"> • id_venta_det • id_venta_cab • id_numero_doc • secuencia_det • id_bodega • id_item • tipo_item • cantidad_unid • cantidad_frac • precio_0 • precio_1 • porc_desccto • descuento_0 • descuento_1 • iva_0 • iva_1 • neto_0 • neto_1 • costo_unitario_0 • costo_unitario_1 • costo_total_0 • costo_total_1

TABLAS Y CAMPOS A USAR DEL ERP	
Tablas ERP	Campos
In_item	 <ul style="list-style-type: none"> • id_item • descripcion • descripcion_larga • id_clasif_1 • id_clasif_2 • fecha_ingreso • precio • id_estado_item • tipo_item • aplica_iva • aplica_ice • porc_utilidad • aplica_irbp • caja_primaria • id_marca_item • precio_fraccion
In_producto	 <ul style="list-style-type: none"> • id_producto • presentacion • medida • concentracion • stock_unidad

TABLAS Y CAMPOS A USAR DEL ERP							
Tablas ERP	Campos						
	<ul style="list-style-type: none"> • stock_fraccion • stk_ing_unid • stk_ing_fracc • stk_egr_unid • stk_egr_fracc • num_fraccion • id_fabricante • costo_prom_0 • costo_prom_1 • vvf • pvf • fra_minima • fra_maxima • cronico • oro 						
Ve_faltante	<div style="border: 1px solid gray; padding: 5px; margin-bottom: 10px;"> Results Messages <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th>Name</th> <th>Owner</th> <th>Type</th> </tr> </thead> <tbody> <tr> <td>ve_faltante</td> <td>dbo</td> <td>usertable</td> </tr> </tbody> </table> </div> <ul style="list-style-type: none"> • id_faltante • id_item • descripcion • unidades • fracciones • motivo • fecha_hora • id_bodega • usuario • estado 	Name	Owner	Type	ve_faltante	dbo	usertable
Name	Owner	Type					
ve_faltante	dbo	usertable					

TABLAS Y CAMPOS A USAR DEL ERP	
Tablas ERP	Campos
	<ul style="list-style-type: none"> • unidades_pedidas • fracciones_pedidas
sg_usuario	 <ul style="list-style-type: none"> • id_usuario • id_empresa • descripcion • login • contraseña • activo
Ve_venta_estado	<ul style="list-style-type: none"> • Estado

3.2 Levantamiento de información de procesos actuales

Dentro de los procesos actuales que realizan los clientes para la generación de información procesada, y que necesitan para soporte en la toma de decisiones, se tienen los dos siguientes:

- En la Figura 3.4 se observa el proceso cotidiano donde el cliente simplemente genera la consulta con los parámetros requeridos en el reporte del ERP y visualiza el resultado en el módulo de reportes.

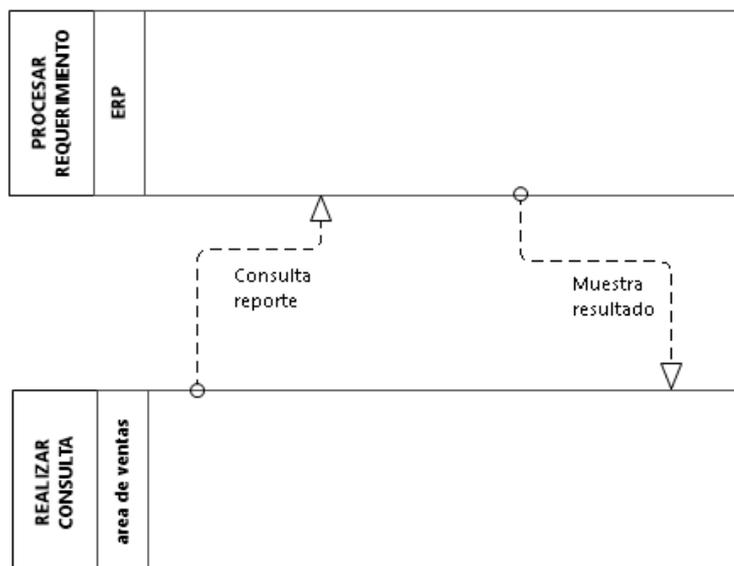


Figura 3.4: Proceso consulta reporte

- En la Figura 3.5 se visualiza el proceso donde el cliente genera la consulta con los parámetros requeridos en el reporte del ERP, descarga los reportes del sistema y los entrega a una tercera persona para que esta los procese y genere los informes estadísticos y cuadros dinámicos según el requerimiento del cliente.

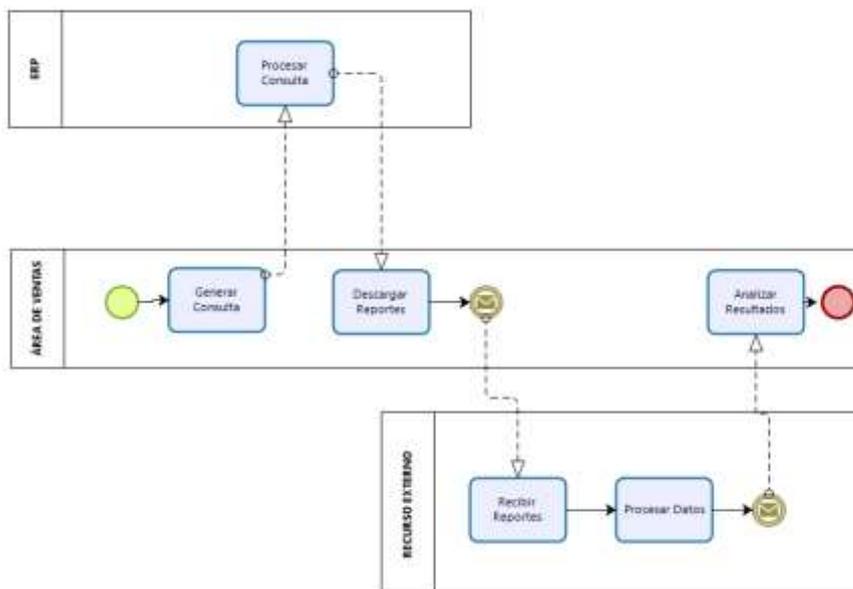


Figura 3.5: Proceso reportes por un recurso externo

3.3 Definición de actores y casos de uso

Se identifican los actores que son afectados, los que intervienen y aquellos que podría interesarles el proceso. Para esta actividad, se toma como referencia un cliente escogido con quien se realizará el proyecto piloto.

Tabla 4: Actores

ACTORES		
NOMBRE	CARGO	EMPRESA
Juan Ronquillo	Coordinador general	Proveedor
Viviana Magallanes	Desarrollador	Proveedor

ACTORES		
NOMBRE	CARGO	EMPRESA
Angel Jiménez	Implementador de soluciones	Proveedor
Oscar Vincos	DBA	Proveedor
Melissa Muñoz	Asistente gerencia ventas	Cliente
Federico Lebet	Gerente	Cliente

Dentro de los actores se ha incluido personas de la empresa que proveen los servicios de ERP además del cliente. El Coordinador General dentro de la empresa desarrolladora de software, para autorizar la construcción de la solución planteada además del análisis del beneficio como valor agregado para la compañía hacia sus clientes y el Jefe del Proyecto. El Implementador quien será el encargado de realizar el levantamiento de información y modelamiento de los paneles de información, el Desarrollador que tendrá a cargo realizar las consultas de extracción desde las fuentes de datos y el DBA, quien está encargado de crear el modelo del Datamart e implementar la solución realizada. De parte de los clientes interviene la Asistente de Gerencia de Ventas, quien es la encargada de gestionar los datos para entregarlos a su jefe inmediato y el Gerente, encargado de toma de decisiones y dirigir el negocio.

Tabla 5: CU-001 Caso de uso

CU-001 CASO DE USO															
CU-001	Obtener panel de información desde el sistema														
Versión	1.0 (03/03/2020)														
Precondición	El usuario ingresó a la aplicación móvil o página web de Microsoft Power BI y se autenticó con su usuario y contraseña.														
Descripción	El sistema deberá comportarse como se describe en el siguiente caso de uso cuando el usuario requiera visualizar los reportes asignados.														
Secuencia Normal	<table border="1"> <thead> <tr> <th>Paso</th> <th>Acción</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>El usuario solicita al sistema iniciar el proceso de visualización de paneles.</td> </tr> <tr> <td>2</td> <td>El sistema muestra los paneles de información que la cuenta del usuario tiene asignados.</td> </tr> <tr> <td>3</td> <td>El usuario selecciona el panel de información requerido.</td> </tr> <tr> <td>4</td> <td>El sistema muestra el panel de información con los valores estándares configurados además de los filtros para cambiar el reporte.</td> </tr> <tr> <td>5</td> <td>El usuario ingresa los nuevos valores en los filtros que contiene el reporte.</td> </tr> <tr> <td>6</td> <td>El sistema actualiza la información del reporte.</td> </tr> </tbody> </table>	Paso	Acción	1	El usuario solicita al sistema iniciar el proceso de visualización de paneles.	2	El sistema muestra los paneles de información que la cuenta del usuario tiene asignados.	3	El usuario selecciona el panel de información requerido.	4	El sistema muestra el panel de información con los valores estándares configurados además de los filtros para cambiar el reporte.	5	El usuario ingresa los nuevos valores en los filtros que contiene el reporte.	6	El sistema actualiza la información del reporte.
	Paso	Acción													
	1	El usuario solicita al sistema iniciar el proceso de visualización de paneles.													
	2	El sistema muestra los paneles de información que la cuenta del usuario tiene asignados.													
	3	El usuario selecciona el panel de información requerido.													
	4	El sistema muestra el panel de información con los valores estándares configurados además de los filtros para cambiar el reporte.													
	5	El usuario ingresa los nuevos valores en los filtros que contiene el reporte.													
6	El sistema actualiza la información del reporte.														
Excepciones	<table border="1"> <thead> <tr> <th>Paso</th> <th>Acción</th> </tr> </thead> <tbody> <tr> <td rowspan="3">5</td> <td>Si el usuario no tuviere la necesidad de visualizar otra información con respecto a la mostrada por defecto.</td> </tr> <tr> <td>E1: EL usuario no manipula los filtros disponibles.</td> </tr> <tr> <td>E2: El sistema no realiza ninguna acción adicional.</td> </tr> </tbody> </table>	Paso	Acción	5	Si el usuario no tuviere la necesidad de visualizar otra información con respecto a la mostrada por defecto.	E1: EL usuario no manipula los filtros disponibles.	E2: El sistema no realiza ninguna acción adicional.								
	Paso	Acción													
	5	Si el usuario no tuviere la necesidad de visualizar otra información con respecto a la mostrada por defecto.													
E1: EL usuario no manipula los filtros disponibles.															
E2: El sistema no realiza ninguna acción adicional.															
Comentarios	El número máximo de paneles de información disponibles para los usuarios es de cuatro y el número máximo de usuarios por cliente será de dos.														

3.4 Definición del proceso mejorado

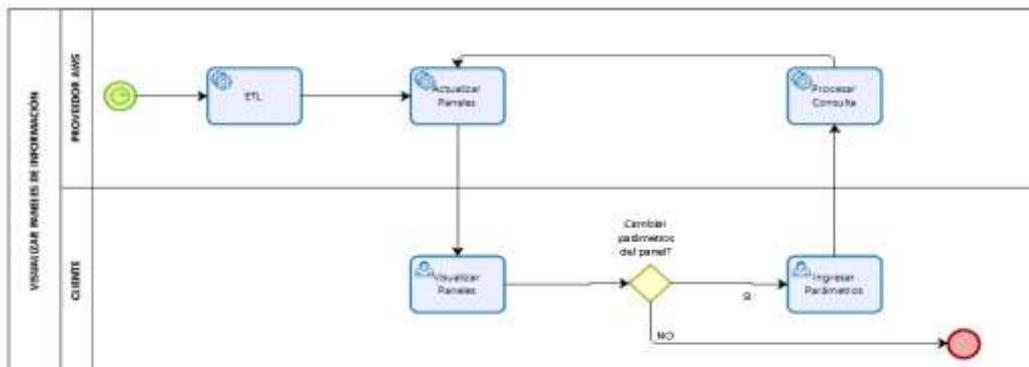


Figura 3.6: Proceso mejorado

En la Figura 3.6 se establece el proceso donde la empresa desarrolladora de software como proveedor del servicio ERP en AWS, y el cliente interactúan de manera directa y el proceso de generación de los paneles de información es automática.

Este proceso contiene las tareas de parte del proveedor: ETL, actualiza paneles y procesa consulta que son de tipo servicio, es decir automáticas; sin intervención de algún usuario para que se cumpla. Estas tareas serán implementadas previamente por el personal establecido como responsable. Las tareas que corresponden al cliente: visualizar paneles e ingresar parámetros serán de tipo usuario, lo cual indica que necesitan de la interacción de una persona para llevarse a cabo.

3.5 Levantamiento de Requerimientos Funcionales

- El sistema deberá poder ser accesible desde cualquier dispositivo y desde cualquier lugar que se tenga acceso a internet.
- Se permitirá el acceso solo a usuarios registrados con sus respectivas credenciales.
- Los usuarios que ingresen al sistema podrán tener disponibles los reportes ya establecidos anteriormente en este documento.
- El proceso de extracción y tratamiento de la información deberá ser automático y sin afectar las operaciones de los usuarios.
- El proceso de actualización de datos debe ser automático.
- El sistema mostrará al usuario los paneles con información estándar y con parámetros ya definidos.

Tabla 6: Requerimiento funcional del sistema

REQUERIMIENTO FUNCIONAL DEL SISTEMA	
Panel de información	Definiciones
Ventas por categorías por local y vendedor	Mostrará por defecto los valores de ventas del día anterior agrupados por cada categoría. Si el cliente tuviera más de un local, mostrará el consolidado de todos los locales.

REQUERIMIENTO FUNCIONAL DEL SISTEMA	
Panel de información	Definiciones
	Filtros: <ul style="list-style-type: none"> • Local (En caso de existir más de uno). • Vendedor. • Categoría. • Fecha.
Oportunidades de ventas perdidas detallado por productos y por local	Mostrará por defecto solo el Top 10 de productos que hayan completado el número de fracciones correspondientes a una caja entera, durante el mes en curso. Filtros: <ul style="list-style-type: none"> • Local. • Fecha.
Curva ABC de productos (Top 10 de productos con mayor y menor rotación)	Se mostrará por defecto dos listados con los 10 productos de mayor y menor ventas de unidades y fracciones durante el mes en curso. Filtros:

REQUERIMIENTO FUNCIONAL DEL SISTEMA	
Panel de información	Definiciones
	<ul style="list-style-type: none"> • Local. • Fecha.
Curva ABC de rentabilidad (Top 10 de productos vendidos con mayor y menor rentabilidad)	<p>Se mostrará por defecto dos listados con los 10 productos de mayor y menor utilidad generada para la empresa durante el mes en curso.</p> <p>Filtros:</p> <ul style="list-style-type: none"> • Local. • Fecha.

El sistema deberá permitir la interacción del usuario con los diferentes filtros que contengan los paneles, los cuales se han establecido anteriormente.

3.6 Alcance del Proyecto

El alcance de este proyecto cubre la optimización del proceso de extracción, tratamiento y presentación de la información que le corresponde al Departamento de Ventas de los clientes que consumen el recurso del servicio del ERP, creando un proceso ETL

y un Datamart con el fin de presentar cuatro paneles de información que son:

- Ventas por categorías por local y vendedor.
- Oportunidades de ventas perdidas detallado por productos y por local.
- Curva ABC de productos (Top 10 de productos con mayor y menor rotación).
- Curva ABC de rentabilidad (Top 10 de productos vendidos con mayor y menor rentabilidad).

Estos paneles se implementarán como proyecto piloto para un cliente del segmento retail que se encuentra en hosting de AWS.

CAPÍTULO 4

ANÁLISIS Y DISEÑO

4.1 Análisis de requerimientos

Según lo expuesto en el Capítulo 3 por el usuario que ha sido seleccionado, el cual cumple con los requisitos planteados para ser parte de la prueba piloto de esta aplicación y en conjunto con el equipo de trabajo de la empresa desarrolladora del ERP, todos los campos necesarios para la construcción de la solución existen dentro de las bases de datos, además de los recursos de hardware y software que se tiene en la actualidad sirven para este propósito. Como política de la empresa desarrolladora del ERP de tener la menor cantidad de personalizaciones exclusivas por cliente dentro del sistema, se identifica que el Datamart y los paneles aquí planteados se pueden estandarizar para todos los clientes del segmento retail.

Los requerimientos de acceso y autenticación son plenamente manejables con las aplicaciones que se utilizarán en este proyecto. Para la estructura del Datamart se utilizará un modelo de tipo estrella que comprende cinco tablas de dimensiones y una tabla de hechos.

4.2 Arquitectura

Para este proyecto se utilizará tecnología Microsoft según se detalla en capítulos anteriores, debido a que la empresa desarrolladora del ERP tiene todas estas herramientas disponibles y en la siguiente figura se visualiza las herramientas que interactúan en cada proceso.

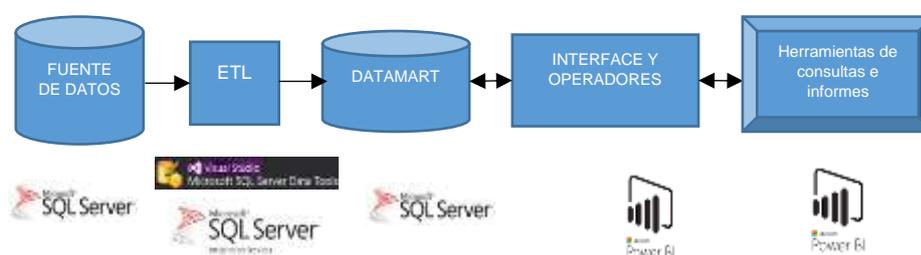


Figura 4.1: Arquitectura

Para la consulta de los datos, se obtendrán estos directamente desde la base de datos del ERP que se encuentra en SQL Server. A nivel del ETL, se gestionará a través de SQL Server Data Tools para Visual Studio y la herramienta Integration Services de SQL Server el proceso de extracción, tratamiento y carga de la información hacia el

Datamart creado en SQL Server que se encontrará en la misma instancia de AWS.

Mediante la aplicación Microsoft Power BI desktop se construyen los paneles, cada uno de sus parámetros y la interface con el Datamart. Finalmente, para la consulta de informes y análisis de datos por parte del cliente se hará uso de la herramienta Microsoft Power BI, la cual tendrá como alcance mostrar los paneles descritos como solución en el Capítulo 1.

4.3 Diseño del Datamart

De acuerdo al alcance y análisis del proyecto, esta aplicación está destinada al área de ventas y a toda persona interesada en el proceso dentro de cada empresa cliente. A continuación, se ilustra el modelo dimensional utilizado para la creación del Datamart.

4.3.1 Medidas

Los objetivos que se intentan medir con esta solución son los siguientes:

- Valores totales en ventas dado por el campo Total_Venta que suma las ventas de un determinado producto, vendedor o sucursal en un lapso determinado de tiempo.
- Productos con mayor rentabilidad en ventas, para lo cual se tiene un campo Total_utilidad.
- Listado de productos con mayor y menor unidades vendidas además de los productos que no se vendieron y sus cantidades durante un rango de tiempo, por lo que se tendrá un campo donde se sumen las cantidades del producto vendido con el nombre de Cant_venta.

4.3.2 Dimensiones

Se definen los nombres de las siguientes dimensiones: Dim_producto, Dim_tiempo, Dim_vendedor, Dim_sucursal, Dim_tipo y la tabla de hechos Fact_ventas.

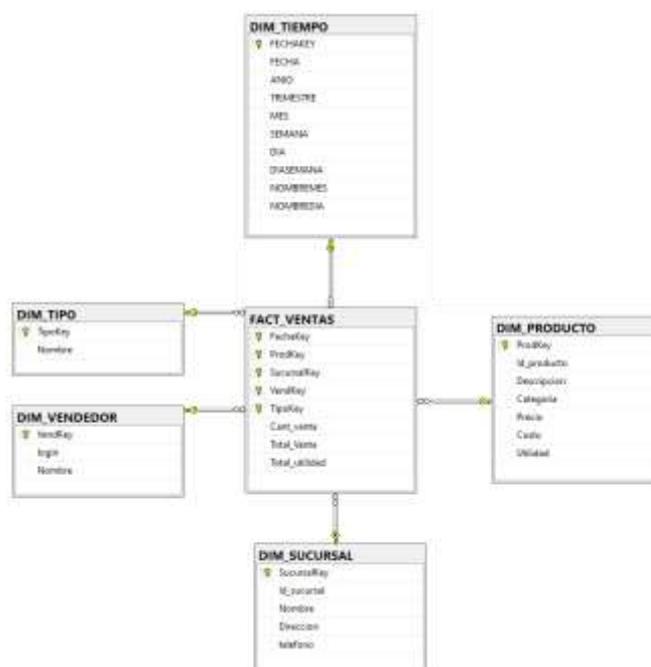


Figura 4.2: Modelo dimensional Datamart Ventas

4.4 Definición de métodos y extracción de información, construcción del ETL

4.4.1 Creación Datamart

Para crear la base de datos se utiliza la siguiente sentencia:

```
USE [master]
GO
CREATE DATABASE [DMVENTAS_COERIEGO]
CONTAINMENT = NONE
ON PRIMARY
( NAME = N'DMVENTAS', FILENAME = N'C:\Program Files\Microsoft SQL
Server\MSSQL14.MSSQLSERVER\MSSQL\DATA\DMVENTAS.mdf', SIZE =
8192KB, MAXSIZE = UNLIMITED, FILEGROWTH = 65536KB )
LOG ON
( NAME = N'DMVENTAS_log', FILENAME = N'C:\Program Files\Microsoft SQL
Server\MSSQL14.MSSQLSERVER\MSSQL\DATA\DMVENTAS_log.ldf', SIZE =
8192KB, MAXSIZE = 2048GB, FILEGROWTH = 65536KB )
GO
ALTER DATABASE [DMVENTAS_COERIEGO] SET COMPATIBILITY_LEVEL =
140
GO
```

```
IF (1 = FULLTEXTSERVICEPROPERTY('IsFullTextInstalled'))
begin
EXEC [DMVENTAS_COERIEGO].[dbo].[sp_fulltext_database] @action = 'enable'
end
GO

ALTER DATABASE [DMVENTAS_COERIEGO] SET ANSI_NULL_DEFAULT
OFF
GO

ALTER DATABASE [DMVENTAS_COERIEGO] SET ANSI_NULLS OFF
GO

ALTER DATABASE [DMVENTAS_COERIEGO] SET ANSI_PADDING OFF
GO

ALTER DATABASE [DMVENTAS_COERIEGO] SET ANSI_WARNINGS OFF
GO

ALTER DATABASE [DMVENTAS_COERIEGO] SET ARITHABORT OFF
GO

ALTER DATABASE [DMVENTAS_COERIEGO] SET AUTO_CLOSE OFF
GO

ALTER DATABASE [DMVENTAS_COERIEGO] SET AUTO_SHRINK OFF
GO

ALTER DATABASE [DMVENTAS_COERIEGO] SET
AUTO_UPDATE_STATISTICS ON
GO

ALTER DATABASE [DMVENTAS_COERIEGO] SET
CURSOR_CLOSE_ON_COMMIT OFF
GO

ALTER DATABASE [DMVENTAS_COERIEGO] SET CURSOR_DEFAULT
GLOBAL
GO

ALTER DATABASE [DMVENTAS_COERIEGO] SET
CONCAT_NULL_YIELDS_NULL OFF
GO

ALTER DATABASE [DMVENTAS_COERIEGO] SET NUMERIC_ROUNDABORT
OFF
GO

ALTER DATABASE [DMVENTAS_COERIEGO] SET QUOTED_IDENTIFIER OFF
GO

ALTER DATABASE [DMVENTAS_COERIEGO] SET RECURSIVE_TRIGGERS
OFF
```

```
GO
ALTER DATABASE [DMVENTAS_COERIEGO] SET DISABLE_BROKER
GO

ALTER DATABASE [DMVENTAS_COERIEGO] SET
AUTO_UPDATE_STATISTICS_ASYNC OFF
GO

ALTER DATABASE [DMVENTAS_COERIEGO] SET
DATE_CORRELATION_OPTIMIZATION OFF
GO

ALTER DATABASE [DMVENTAS_COERIEGO] SET TRUSTWORTHY OFF
GO

ALTER DATABASE [DMVENTAS_COERIEGO] SET
ALLOW_SNAPSHOT_ISOLATION OFF
GO

ALTER DATABASE [DMVENTAS_COERIEGO] SET PARAMETERIZATION
SIMPLE
GO

ALTER DATABASE [DMVENTAS_COERIEGO] SET
READ_COMMITTED_SNAPSHOT OFF
GO

ALTER DATABASE [DMVENTAS_COERIEGO] SET
HONOR_BROKER_PRIORITY OFF
GO

ALTER DATABASE [DMVENTAS_COERIEGO] SET RECOVERY SIMPLE
GO

ALTER DATABASE [DMVENTAS_COERIEGO] SET MULTI_USER
GO
ALTER DATABASE [DMVENTAS_COERIEGO] SET PAGE_VERIFY
CHECKSUM
GO

ALTER DATABASE [DMVENTAS_COERIEGO] SET DB_CHAINING OFF
GO

ALTER DATABASE [DMVENTAS_COERIEGO] SET FILESTREAM(
NON_TRANSACTED_ACCESS = OFF )
GO

ALTER DATABASE [DMVENTAS_COERIEGO] SET
TARGET_RECOVERY_TIME = 60 SECONDS
GO

ALTER DATABASE [DMVENTAS_COERIEGO] SET DELAYED_DURABILITY =
DISABLED
```

```
GO
```

```
ALTER DATABASE [DMVENTAS_COERIEGO] SET QUERY_STORE = OFF  
GO
```

```
ALTER DATABASE [DMVENTAS_COERIEGO] SET READ_WRITE  
GO
```

Se crea inicialmente la estructura de las tablas y campos del

Datamart, para eso utilizaremos la siguiente sentencia:

```
USE [DMVENTAS_COERIEGO]
```

```
GO
```

```
SET ANSI_NULLS ON
```

```
GO
```

```
SET QUOTED_IDENTIFIER ON
```

```
GO
```

```
CREATE TABLE [dbo].[DIM_PRODUCTO](  
    [ProdKey] [int] Identity(1,1) NOT NULL,  
    [Id_producto] [int] NOT NULL,  
    [Descripcion] [varchar](50) NOT NULL,  
    [Categoria] [varchar](50) NOT NULL,  
    [Precio] [varchar](50) NOT NULL,  
    [Costo] [money] NOT NULL,  
    [Utilidad] [decimal](10,2)NOT NULL,
```

```
CONSTRAINT [PK_DIM_PRODUCTO] PRIMARY KEY CLUSTERED
```

```
( [ProdKey] ASC)WITH (PAD_INDEX = OFF, STATISTICS_NORECOMPUTE  
= OFF, IGNORE_DUP_KEY = OFF, ALLOW_ROW_LOCKS = ON,  
ALLOW_PAGE_LOCKS = ON) ON [PRIMARY]
```

```
) ON [PRIMARY]
```

```
GO
```

```
SET ANSI_NULLS ON
```

```
GO
```

```
SET QUOTED_IDENTIFIER ON
GO
CREATE TABLE [dbo].[DIM_SUCURSAL](
    SucursalKey int identity(1,1) NOT NULL,
    Id_sucursal int NOT NULL,
    Nombre varchar(20) NOT NULL,
    Direccion varchar(50) NOT NULL,
    telefono Varchar(20) NOT NULL
CONSTRAINT [PK_DIM_SUCURSAL] PRIMARY KEY CLUSTERED
(
    [SucursalKey] ASC
)WITH (PAD_INDEX = OFF, STATISTICS_NORECOMPUTE = OFF,
IGNORE_DUP_KEY = OFF, ALLOW_ROW_LOCKS = ON,
ALLOW_PAGE_LOCKS = ON) ON [PRIMARY]
) ON [PRIMARY]
GO
```

```
SET ANSI_NULLS ON
GO
```

```
SET QUOTED_IDENTIFIER ON
GO
CREATE TABLE [dbo].[DIM_TIPO](
    TipoKey int identity(1,1) NOT NULL,
    Nombre varchar(20) NOT NULL
CONSTRAINT [PK_DIM_TIPO] PRIMARY KEY CLUSTERED
(
    [TipoKey] ASC
)WITH (PAD_INDEX = OFF, STATISTICS_NORECOMPUTE = OFF,
IGNORE_DUP_KEY = OFF, ALLOW_ROW_LOCKS = ON,
ALLOW_PAGE_LOCKS = ON) ON [PRIMARY]
) ON [PRIMARY]
GO
```

```
SET ANSI_NULLS ON
```

```
GO
```

```
SET QUOTED_IDENTIFIER ON
```

```
GO
```

```
CREATE TABLE [dbo].[DIM_VENDEDOR](  
    VendKey int identity(1,1) NOT NULL,  
    login varchar (20) NOT NULL,  
    Nombre varchar(50) NOT NULL,  
    CONSTRAINT [PK_DIM_VENDEDOR] PRIMARY KEY CLUSTERED  
  
    (  
        [VendKey] ASC  
    )WITH (PAD_INDEX = OFF, STATISTICS_NORECOMPUTE = OFF,  
    IGNORE_DUP_KEY = OFF, ALLOW_ROW_LOCKS = ON,  
    ALLOW_PAGE_LOCKS = ON) ON [PRIMARY]  
    ) ON [PRIMARY]  
GO
```

```
SET ANSI_NULLS ON
```

```
GO
```

```
SET QUOTED_IDENTIFIER ON
```

```
GO
```

```
CREATE TABLE DIM_TIEMPO
```

```
(  
    FECHAKEY INT NOT NULL,  
    FECHA DATE NOT NULL,  
    ANIO SMALLINT NOT NULL,  
    TRIMESTRE SMALLINT NOT NULL,  
    MES SMALLINT NOT NULL,  
    SEMANA SMALLINT NOT NULL,  
    DIA SMALLINT NOT NULL,
```

```
DIASEMANA SMALLINT NOT NULL,  
NOMBREMES CHAR(15) NOT NULL,  
NOMBREDIA CHAR(10) NOT NULL  
CONSTRAINT PK_DIM_TIEMPO PRIMARY KEY CLUSTERED  
(  
    FECHAKEY ASC  
)  
WITH (PAD_INDEX = OFF, STATISTICS_NORECOMPUTE = OFF,  
    IGNORE_DUP_KEY = OFF, ALLOW_ROW_LOCKS = ON,  
    ALLOW_PAGE_LOCKS = ON) ON [PRIMARY]  
) ON [PRIMARY]  
GO  
  
SET ANSI_NULLS ON  
GO  
SET QUOTED_IDENTIFIER ON  
GO  
  
CREATE TABLE [dbo].[FACT_VENTAS](  
    FechaKey int NOT NULL,  
    ProdKey int NOT NULL,  
    SucursalKey int NOT NULL,  
    VendKey int NOT NULL,  
    TipoKey int NOT NULL,  
    Cant_venta int NOT NULL,  
    Total_Venta money NOT NULL,  
    Total_utilidad decimal(10,2) NOT NULL  
    CONSTRAINT [PK_FACT_VENTAS] PRIMARY KEY CLUSTERED  
(  
    [FechaKey] ASC,  
    ProdKey ASC,
```

```
SucursalKey ASC,  
VendKey ASC,  
Tipokey ASC  
)WITH (PAD_INDEX = OFF, STATISTICS_NORECOMPUTE = OFF,  
IGNORE_DUP_KEY = OFF, ALLOW_ROW_LOCKS = ON,  
ALLOW_PAGE_LOCKS = ON) ON [PRIMARY]  
) ON [PRIMARY]  
GO  
  
BEGIN TRANSACTION  
SET QUOTED_IDENTIFIER ON  
SET ARITHABORT ON  
SET NUMERIC_ROUNDABORT OFF  
SET CONCAT_NULL_YIELDS_NULL ON  
SET ANSI_NULLS ON  
SET ANSI_PADDING ON  
SET ANSI_WARNINGS ON  
COMMIT  
  
BEGIN TRANSACTION  
GO  
ALTER TABLE dbo.DIM_VENDEDOR SET (LOCK_ESCALATION = TABLE)  
GO  
COMMIT  
  
BEGIN TRANSACTION  
GO  
ALTER TABLE dbo.DIM_TIEMPO SET (LOCK_ESCALATION = TABLE)  
GO  
COMMIT  
  
BEGIN TRANSACTION  
GO  
ALTER TABLE dbo.DIM_SUCURSAL SET (LOCK_ESCALATION = TABLE)  
GO
```

```
COMMIT
BEGIN TRANSACTION
GO
ALTER TABLE dbo.DIM_PRODUCTO SET (LOCK_ESCALATION = TABLE)
GO
COMMIT
BEGIN TRANSACTION
GO
ALTER TABLE dbo.DIM_TIPO SET (LOCK_ESCALATION = TABLE)
GO
COMMIT
BEGIN TRANSACTION
GO
ALTER TABLE dbo.FACT_VENTAS ADD CONSTRAINT
    FK_FACT_VENTAS_DIM_PRODUCTO FOREIGN KEY
    (
        ProdKey
    ) REFERENCES dbo.DIM_PRODUCTO
    (
        ProdKey
    ) ON UPDATE NO ACTION
    ON DELETE NO ACTION

GO
ALTER TABLE dbo.FACT_VENTAS ADD CONSTRAINT
    FK_FACT_VENTAS_DIM_SUCURSAL FOREIGN KEY
    (
        SucursalKey
    ) REFERENCES dbo.DIM_SUCURSAL
    (
```

```
SucursalKey
) ON UPDATE NO ACTION
  ON DELETE NO ACTION
GO
ALTER TABLE dbo.FACT_VENTAS ADD CONSTRAINT
  FK_FACT_VENTAS_DIM_TIPO FOREIGN KEY
  (
    TipoKey
  ) REFERENCES dbo.DIM_TIPO
  (
    TipoKey
  ) ON UPDATE NO ACTION
  ON DELETE NO ACTION

GO
ALTER TABLE dbo.FACT_VENTAS ADD CONSTRAINT
  FK_FACT_VENTAS_DIM_FECHA FOREIGN KEY
  (
    FechaKey
  ) REFERENCES dbo.DIM_TIEMPO
  (
    FechaKey
  ) ON UPDATE NO ACTION
  ON DELETE NO ACTION

GO
ALTER TABLE dbo.FACT_VENTAS ADD CONSTRAINT
  FK_FACT_VENTAS_DIM_VENDEDOR FOREIGN KEY
  (
    VendKey
```

```
) REFERENCES dbo.DIM_VENDEDOR
(
VendKey
) ON UPDATE NO ACTION
ON DELETE NO ACTION

GO

ALTER TABLE dbo.FACT_VENTAS SET (LOCK_ESCALATION = TABLE)
GO
COMMIT
```

4.4.2 Creación y preparación del proyecto SQL Integration Services

Para este proceso según lo expuesto en la Sección 4.2, se utilizará la herramienta Integration Services de SQL que se habilita desde el paquete de SQL Server Data Tools. Desde el menú archivo se creará un nuevo proyecto de tipo Integration Services como lo muestran las Figuras 4.3 y 4.4.

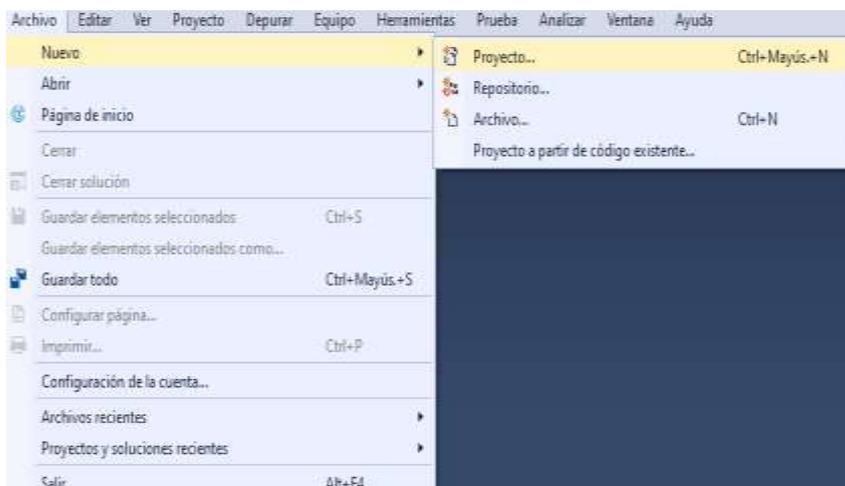


Figura 4.3: Crear proyecto *Integration Services* (1)

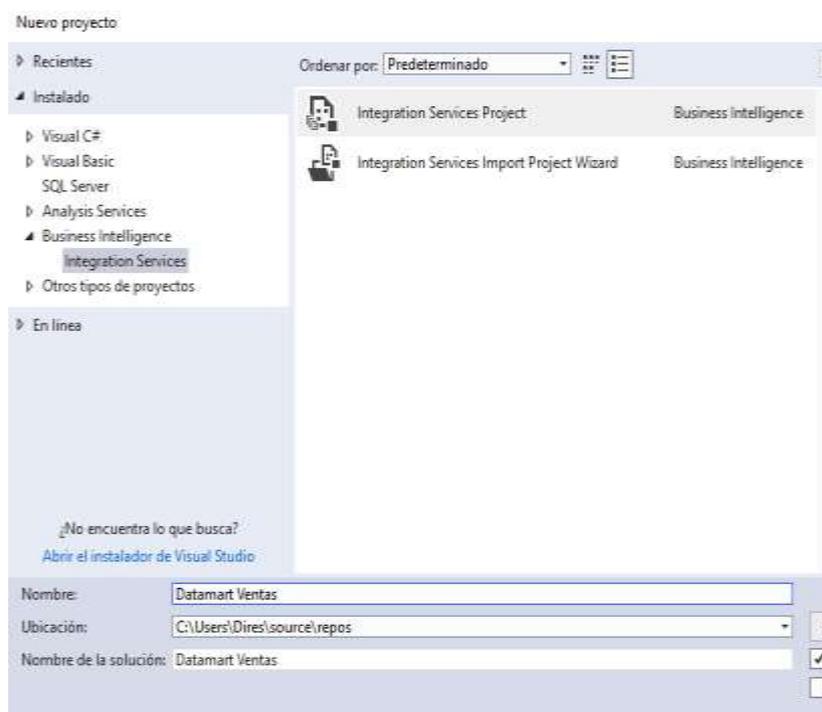


Figura 4.4: Crear proyecto *Integration Services* (2)

Se asigna nombre a los paquetes de SQL Server Integration Services, que para este proyecto serán: Dimensiones.dtsx y Fact_Ventas.dtsx.

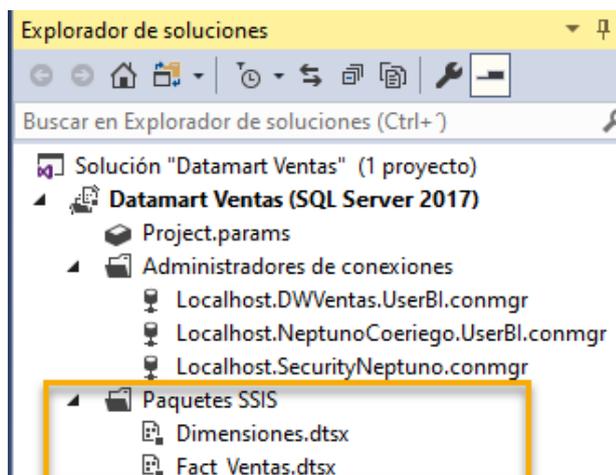


Figura 4.5: Paquetes SSIS

Dentro del paquete Dimensiones.dtsx se contienen las tareas para la carga de datos en las dimensiones tiempo, producto, vendedor, sucursal y tipo. El paquete Fact_ventas.dtsx contiene la tarea de carga de datos en la tabla de hechos.

Se definen las conexiones de origen y destino, desde el administrador de conexiones bajo los siguientes parámetros:

- Tipo de proveedor .Net\Sql DataProvider.
- Las bases de datos origen y destino están alojadas en el hosting de AWS, por lo que se debe colocar la dirección de dicho servidor.
- Se define un usuario exclusivo en el contenedor de las bases de datos SQL Server para este proceso con permisos limitados de solo lectura. Este usuario es UserBI.

- Base de datos origen con información de usuarios: NeptunoSecurity.

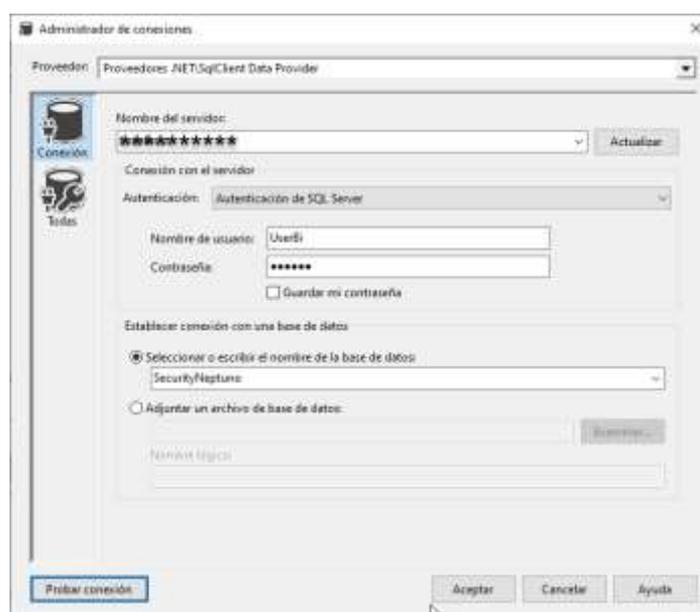


Figura 4.6: ETL Definición Origen de datos DB Security

- Base de datos de negocios origen: NeptunoCoeriego.

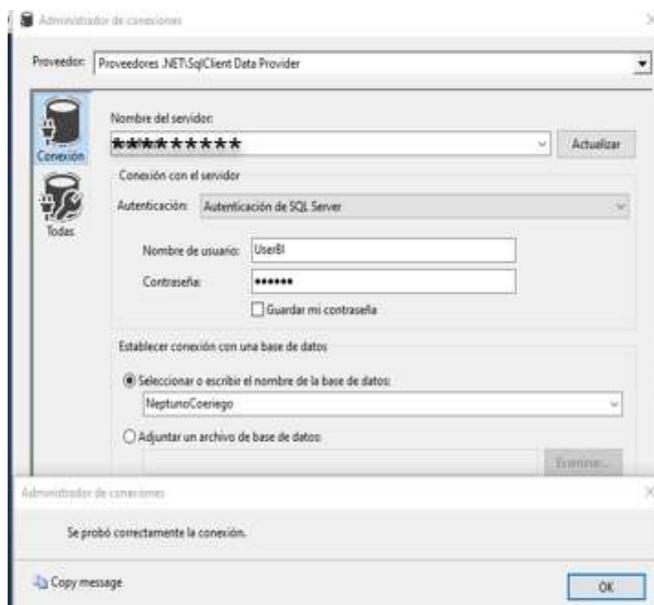


Figura 4.7: ETL Definición origen de datos

- Base de datos destino Datamart: DWVentas.

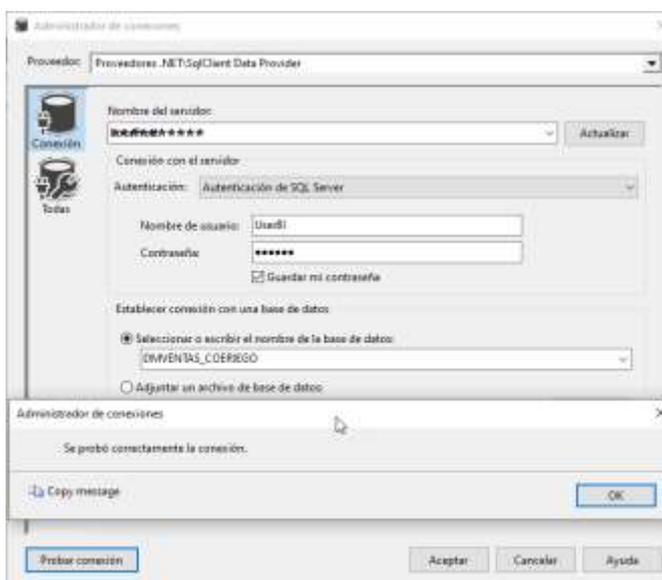


Figura 4.8: ETL Definición destino de datos

Dentro de un mismo flujo de control se ubican las tareas para las cuatro dimensiones antes mencionadas. Se crea una tarea de flujo de datos para los procesos que alimentan cada una de las dimensiones.

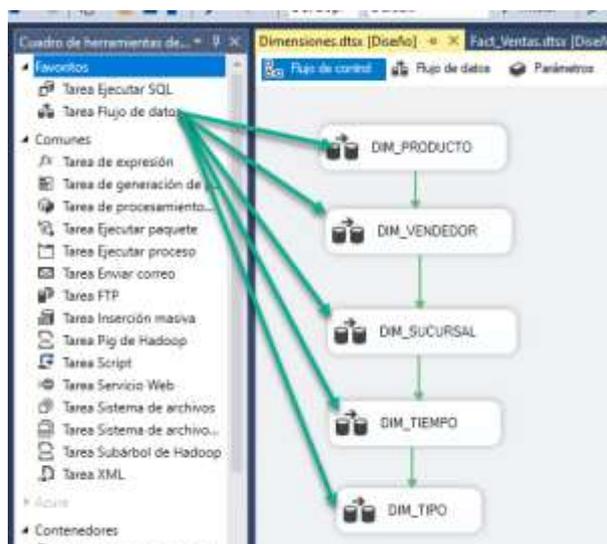


Figura 4.9: Tareas de flujo de datos de dimensiones

4.4.3 Flujo de datos base para el proceso ETL de las tablas de dimensión

Para la construcción del proceso ETL del Datamart y sus dimensiones producto, vendedor, tiempo, sucursal y tipo según muestra la Figura 4.10, se genera un estándar para el flujo de datos el cual permitirá extraer los datos desde el origen para luego hacer una comparación con los datos que ya existen en el Datamart y al final proceder a cargar la información de manera incremental. Para

cada una de las cinco dimensiones definidas, se realizarán las siguientes tareas:

- 1.- Consulta de datos desde origen.
- 2.- Ordenar datos.
- 3.- Combinación de mezcla.
- 4.- División condicional.
- 5.- Carga de datos.

A esto se le sumará alguna tarea adicional propia de cada dimensión.

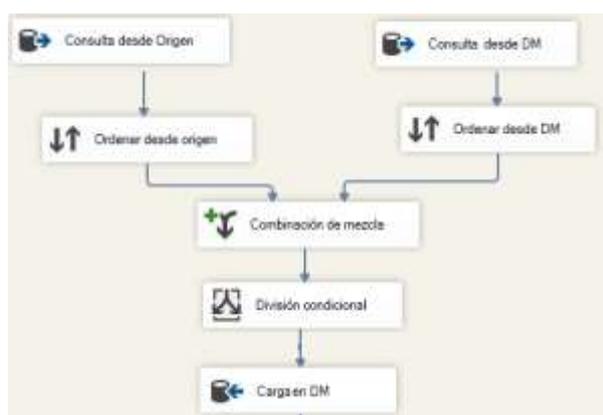


Figura 4.10: Estructura básica del flujo de datos

Leer productos desde origen

Se incluye la consulta que servirá para la extracción de los datos desde la base origen utilizando la conexión hacia dicha base.

Consulta desde DM

Consta de la consulta de los datos existentes dentro de la tabla de dimensión.

Ordenar

Ambas consultas se ordenan por el campo Id de cada dimensión.

Combinación de mezcla

Se realiza la combinación de las dos entradas resultantes de las consultas anteriores con el fin de obtener un listado único de ambas tablas.

División condicional

Al tener ambas tablas combinadas el resultado es un listado de todos los elementos juntos, con la diferencia que los registros que ya existen tienen asignados un campo Id y los que no, tendrán ese campo con un valor nulo. Al ser evaluados con la expresión `isnull(Key "nombreid")` se separan los registros nuevos de los existentes, produciendo así una carga incremental de la información.

Carga en DM

Se cargan los datos a la tabla en la base destino correspondiente.

4.4.4 Creación del flujo de datos para proceso ETL de la dimensión Dim_Producto

A continuación, se describe la configuración detallada de cada una de las tareas antes mencionadas además de la adición de las tareas "Cambiar a Mayúsculas" y "Delimita nombre de Producto" como lo muestra la Figura 4.11.



Figura 4.11: Flujo de datos Dim_Producto

Leer productos desde origen:

Para poder leer los datos desde origen seleccionando la cadena de conexión hacia la base de datos de negocios, se lo realizará mediante el siguiente query:

```
select a.id_item as Id_producto, a.descripcion_larga as Nombre_Item, precio, b.pvf
as costo, c.descripcion as Categoria, (a.precio - b.pvf) as Utilidad from in_item a
inner join in_producto b on a.id_item = b.id_producto
inner join in_nodo_clasif_1 c on a.id_clasif_1 = c.id_nodo_clasif_1 where
a.id_clasif_1 is not null
```

Leer productos desde DM

Para establecer comunicación desde el Datamart, se usa la cadena de conexión creada y para obtener los datos que existen en la actualidad en la tabla dim_producto se utilizará el siguiente query:

```
Select id_producto from dim_producto
```

Combinación de mezcla

Se combinan todos los campos de ambas entradas.

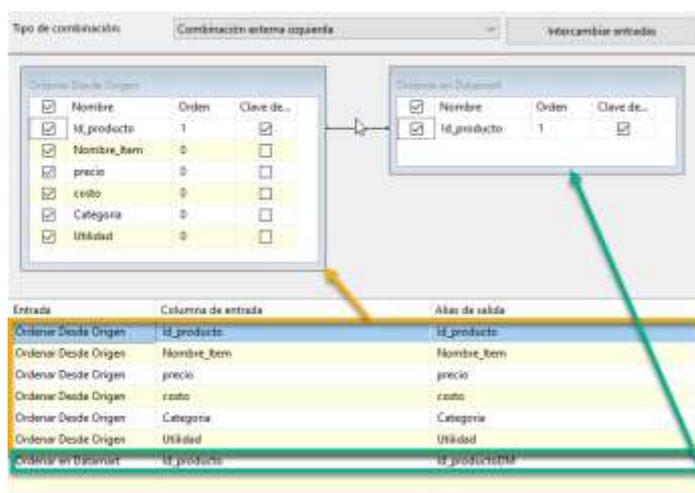


Figura 4.12: Combinación mezcla Dim_Producto

Al id_producto que viene del Datamart se le cambia de nombre para diferenciarlo del id_producto de la base origen.

División condicional

Parámetro condicional para la separación: ISNULL(Id_productoDM).

Cambiar a mayúsculas

Se ha agregado una transformación de mapa de caracteres para estandarizar que todos los nombres de los productos y categorías estén en mayúsculas, así poder tener los datos uniformes.

Delimitar nombre de producto

Al ejecutar la extracción hay datos que tienen nombres excesivamente largos, para no tener novedades al momento de realizar la carga y sobre todo que el panel en BI se vea uniforme, se delimita el campo nombre del producto de los datos de origen de esa manera se asegura que el tamaño del campo ya está validado en el proceso de transformación.

Carga productos DM

Proceso de carga de datos al Datamart, ya con todos los campos validados. La relación de los campos es automática ya que tienen los mismos nombres de campos, al tener los procesos de combinación y mezcla en este paso, se cargarán los datos que no existen con el fin de no borrar nuevamente la tabla y volver a cargar toda la información.

4.4.5 Creación del flujo de datos para proceso ETL de la dimensión Dim_Vendedor

Tal como se explica en la Sección 4.4.3, el proceso en el flujo de datos consta de las cinco tareas básicas establecidas. Para este proceso no hay tareas adicionales, por lo que se detalla el contenido para el proceso incremental de vendedores, según se visualiza en la Figura 4.13.

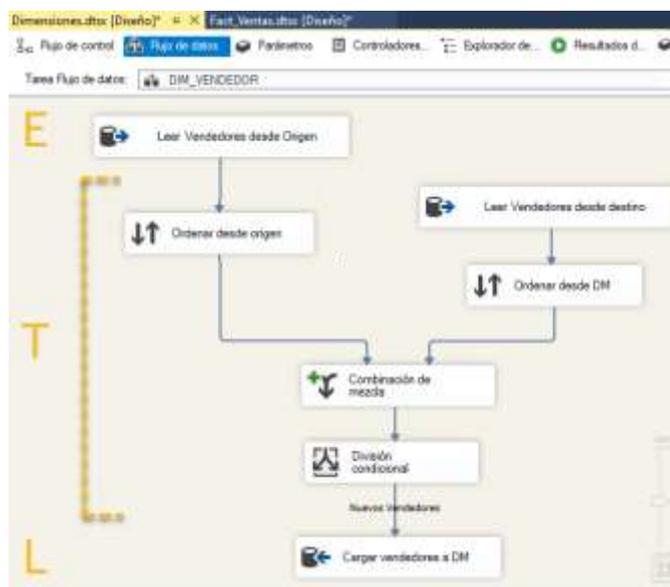


Figura 4.13: Flujo de datos Dim_vendedor

Leer vendedores desde origen

Para poder leer los datos desde origen seleccionando la cadena de conexión hacia la base de datos de negocios, se lo realizará mediante el siguiente query:

```
Select login, descripcion as Nombres from sg_usuario
```

Leer vendedores desde destino

Se establece comunicación desde el Datamart usando la cadena de conexión creada para ello y para obtener los datos que existen en la actualidad en la tabla dim_vendedor, se utilizará la siguiente consulta:

```
Select login from dim_vendedor
```

Ordenar

Ordena los datos para realizar una combinación por el campo login.

Combinación de mezcla

Se combinan todos los campos de ambas entradas.

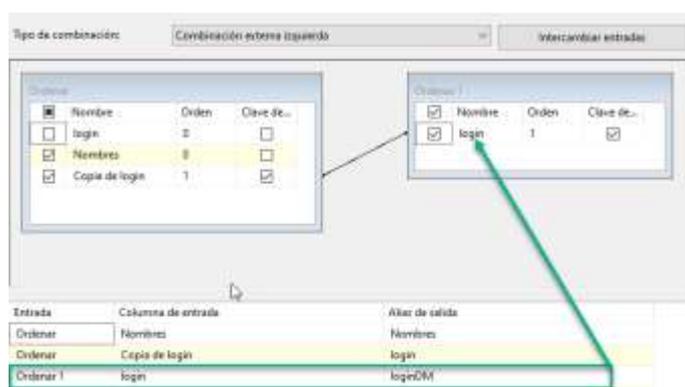


Figura 4.14: Combinación de mezcla Dim_vendedor

Al campo loginDM proveniente del Datamart se le cambia de nombre para diferenciarlo del campo de la base origen.

División condicional

Parámetro condicional para la separación: ISNULL(loginDM).

Cargar vendedores a DM

Llena la tabla dim_vendedor con los datos resultantes de los procesos de tratamiento.

4.4.6 Creación del flujo de datos para proceso ETL de la dimensión Dim_Sucursal

Para el llenado de la dimensión sucursal se agrega dos tareas adicionales a las cinco básicas descritas que son “Mayúsculas” y “Delimitar campos”. A continuación, se describe el contenido de cada tarea según la Figura 4.15.



Figura 4.15: Flujo de datos Dim_sucursal

Leer sucursales desde origen

La extracción de los datos desde la base de negocios origen se realiza mediante la siguiente consulta:

```
Select id_bodega, Direccion, substring( Nombre,1,20)as nombre, Substring (
telefono,1,20) as telefono from in_bodega
```

```
where id_bodega in (select distinct id_bodega from ve_venta_cab)
```

Leer sucursales desde DM

La extracción de los datos existentes en el Datamart se lo realiza mediante la siguiente consulta:

```
Select id_sucursal from dim_sucursal
```

Combinación de mezcla

Combinación de ambas tablas.

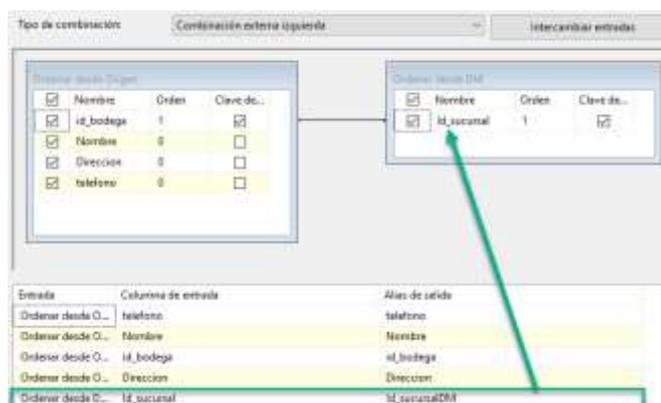


Figura 4.16: Combinación de mezcla Dim_sucursal

Se cambia el nombre del campo Id_sucursal que viene de la tabla del Datamart para poderlo diferenciar del Id_sucursal de origen.

División condicional

Parámetro de condición: ISNULL(Id_sucursalDM).

Mayúsculas

Para mantener la uniformidad de los datos, todos los nombres de las sucursales quedan en mayúscula.

Delimitar campos

Se delimita el largo del campo Nombre para evitar tener datos que puedan afectar la carga por ser demasiado largos.

Cargar sucursales a DM

Carga datos resultantes del proceso de tratamiento en la tabla de Datamart.

4.4.7 Creación del flujo de datos para proceso ETL de la dimensión Dim_Tiempo

Para procesar los datos de la dimensión tiempo se toma en cuenta para el registro inicial la fecha más antigua que se encuentre en la tabla de ventas de la base de negocios. Se tiene como única tarea adicional delimitar campos como se muestran en la Figura 4.17.

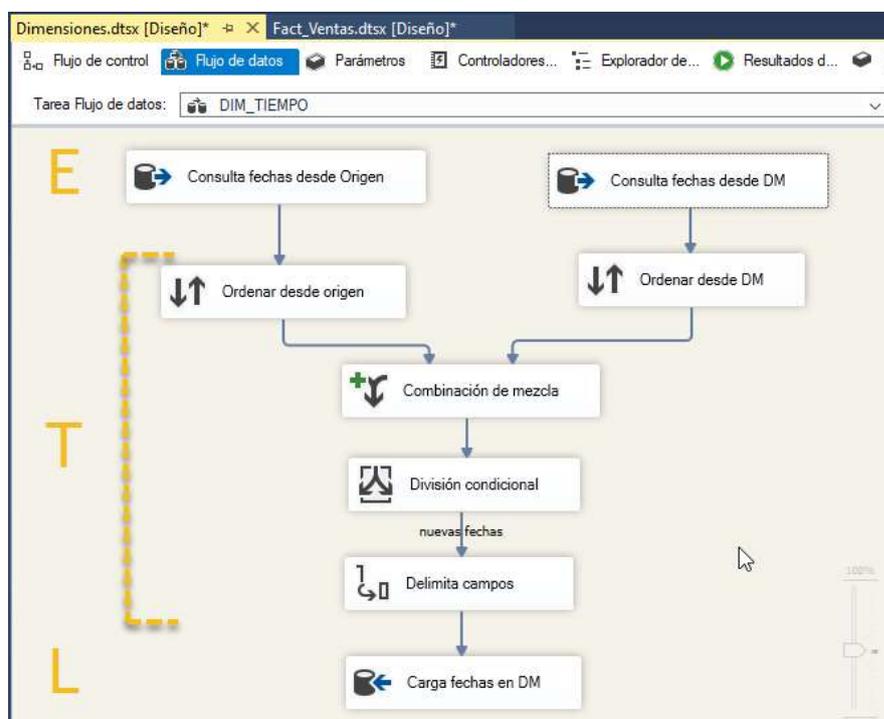


Figura 4.17: Flujo de datos Dim_Tiempo

Consulta fecha desde origen

Para realizar el proceso de extracción de datos desde la base de datos origen se utiliza la siguiente consulta:

```

select fecha_entero, CONVERT(varchar,fecha,23) as fecha, year(fecha) as anio,
month(fecha) as mes, DAY(fecha) as dia, DATEPART(qq,(fecha)) as trimestre,
DATENAME(mm, (fecha)) as NombreMes,datepart (dw,(fecha)) as DiaSemana,
DATENAME(dw,(fecha)) as NombreDia, DATEPART(ww, (fecha))as semana from
ve_venta_cab group by fecha_entero, CONVERT(varchar,fecha,23), year(fecha),
month(fecha), DAY(fecha), DATEPART(qq,(fecha)), DATENAME(mm,
(fecha)),datepart (dw,(fecha)), DATENAME(Dw,(fecha)), datepart (ww, (fecha))

order by year(fecha), month(fecha), DAY(fecha)

```

Consulta fechas desde DM

Para extraer los datos desde la tabla del Datamart, se lo realizará con la siguiente consulta:

```
select fechakey from dim_tiempo
```

Ordenar

Ordenar por campo fecha_entero desde del origen y campo fechakey del Datamart.

Combinación de mezcla

Combinación de ambas tablas.

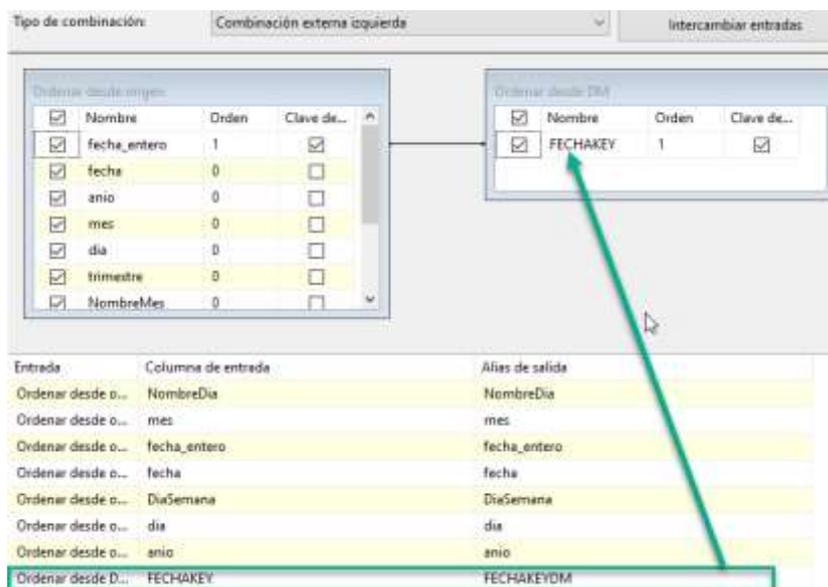


Figura 4.18: Combinación de mezcla dim_tiempo

En el campo keyfecha se coloca un identificador que se pueda diferenciar de los demás campos.

División condicional

Parámetro de condición: ISNULL(FECHAKEYDM).

Delimita campos

En la transformación tipo conversión de datos se realiza las siguientes transformaciones:

Tabla 7: Transformación de campos en Dim_tiempo

TRANSFORMACIÓN DE CAMPOS DIM_TIEMPO		
Entrada	Alias de Salida	Tipo de datos
anio	Copia de anio	Entero de dos bytes con signo [DT_I2]
dia	Copia de dia	Entero de dos bytes con signo [DT_I2]
fecha	Copia de fecha	Fecha [DATE]
mes	Copia de mes	Entero de dos bytes con signo [DT_I2]
NombreDia	Copia de NombreDia	Cadena Unicode 10
NombreMes	Copia de NombreMes	Cadena Unicode 15
semana	Copia de semana	Entero de dos bytes con signo [DT_I2]
trimestre	Copia de trimestre	Entero de dos bytes con signo [DT_I2]

Cargar fechas en DM

Gracias a los procesos anteriores, se identifican los registros nuevos y se cargan al Datamart.

4.4.8 Creación del flujo de datos para proceso ETL de la dimensión Dim_Tipo

En la dimensión tipo están los datos sobre el tipo de transacción, es decir venta o no vendido que significa que es una venta perdida por falta de stock. Para esta dimensión se tiene la siguiente estructura según lo muestra la Figura 4.19:

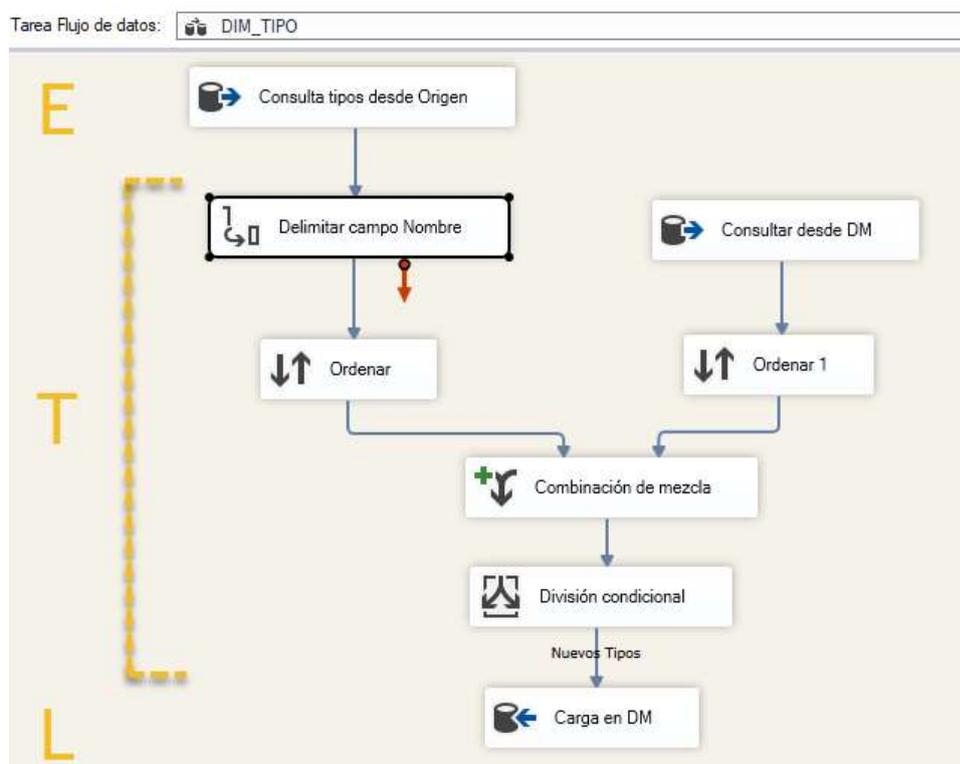


Figura 4.19: Flujo de datos Dim_tipo

Consulta tipos desde origen

Extracción de datos desde origen:

```
select * from ve_tipo_venta
```

Consulta desde DM

Consulta de datos existentes en el Datamart y disponibles para comparación:

```
select * from dim_Tipo
```

Combinación de mezcla

Combinación del contenido de ambos orígenes.

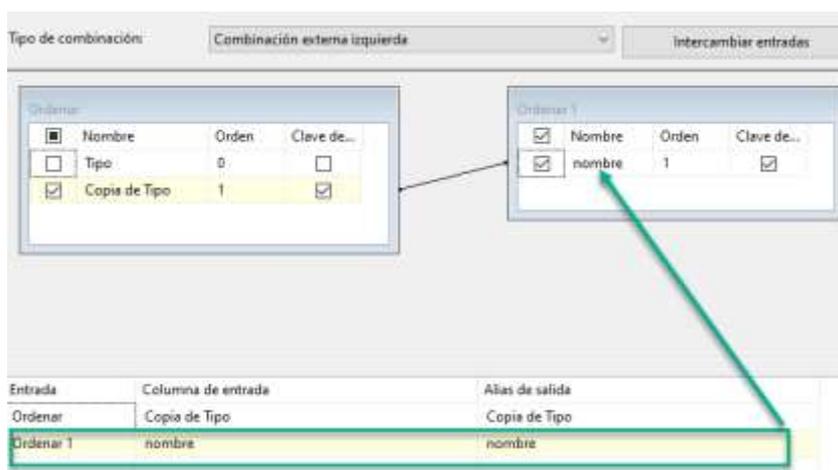


Figura 4.20: Combinación de mezcla Dim_tipo

Se diferencia el campo del Datamart con el que se realizará la validación de tipos nuevos.

División condicional

Parámetro de validación para separar tipos nuevos: Isnull (nombre).

Carga en DM

La carga se realiza de un único campo que es el nombre del tipo de transacción de venta.

4.4.9 Creación del flujo de datos para proceso ETL de la tabla de hechos Fact_ventas

Se crea un paquete SSIS para alojar el proceso ETL de la tabla de hechos llamada Fact_ventas además de una tarea de tipo Ejecutar SQL, la misma que se encarga de borrar los datos en la tabla del Datamart para volver a llenarla, ya que, a diferencia de las dimensiones, esta carga no se realiza de manera incremental. Este proceso además realiza la unión de dos tipos de transacciones que son las ventas emitidas y las ventas perdidas con su respectivo identificador. Dentro del flujo de datos contiene las respectivas relaciones con las tablas de dimensiones mediante el respectivo campo Key de Dim_producto, Dim_vendedor, Dim_tiempo, Dim_sucursal y Dim_tipo, para lo cual se utiliza la herramienta de búsqueda para realizar dicha actividad.

En la Figura 4.21 el flujo de control muestra la tarea inicial que borra los datos en la tabla del Datamart y a continuación ejecuta el flujo de datos para volver a poblar la tabla.

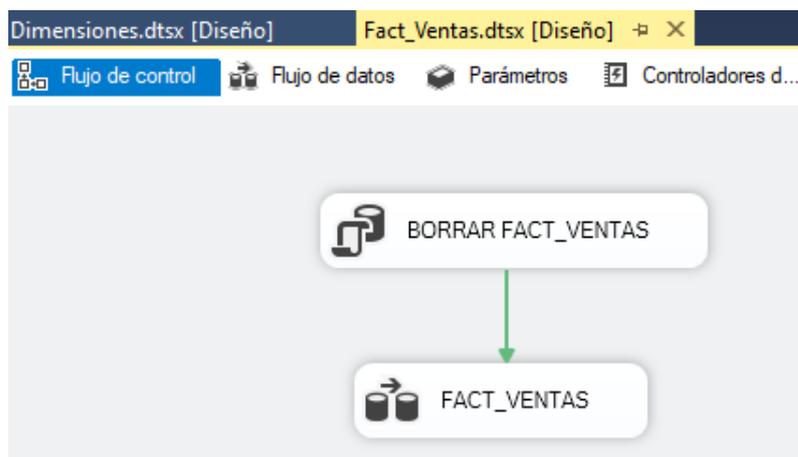


Figura 4.21: Flujo de datos para Fact_ventas

Los parámetros configurados dentro de la tarea de tipo ejecutar SQL BORRAR FACT_VENTAS son los siguientes:

<ul style="list-style-type: none"> Conjunto de resultados <ul style="list-style-type: none"> ResultSet: Ninguno General <ul style="list-style-type: none"> Name: BORRAR FACT_VENTAS Description: Tarea Ejecutar SQL Instrucción SQL <ul style="list-style-type: none"> ConnectionType: ADO.NET Connection: localhost.DMVENTAS_COERIEGO.userBI SQLSourceType: Entrada directa SQLStatement: truncate table Fact_ventas IsQueryStoredProcedure: False BypassPrepare: True Opciones <ul style="list-style-type: none"> TimeOut: 0 CodePage: 1252 TypeConversionMode: Permitido
<p>Name Especifica el nombre de la tarea.</p>

Figura 4.22: Parámetros en Borrar Fact_ventas

A continuación, se describe el proceso dentro del flujo de datos que alimentará la tabla fact_ventas.

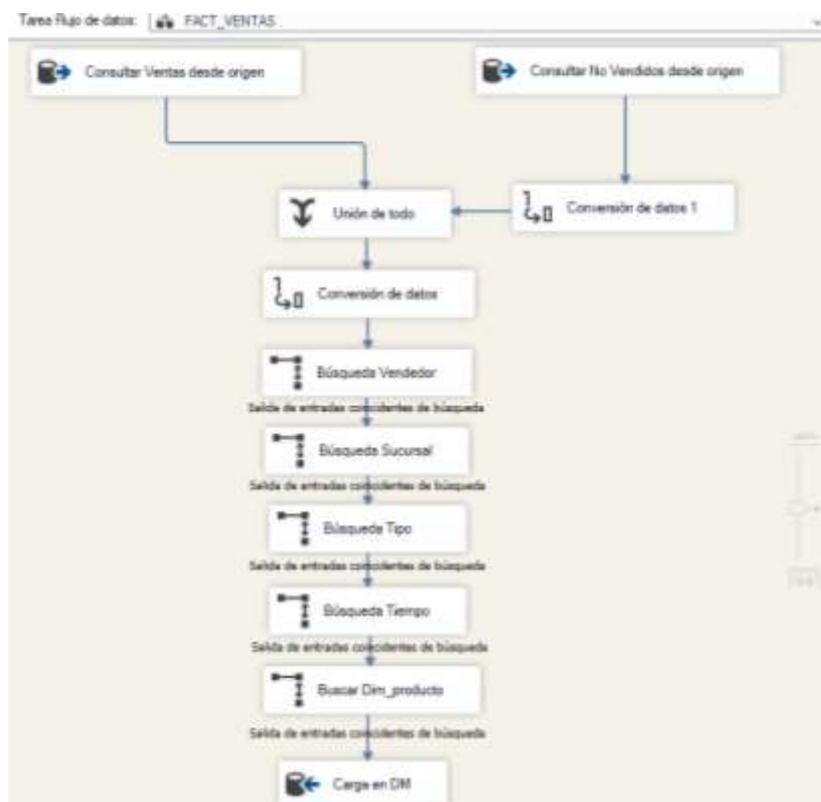


Figura 4.23: Flujo de datos para Fact_ventas

Este flujo de datos extrae la información desde dos diferentes tablas que contienen la misma estructura, las cuales corresponden a las ventas emitidas y ventas perdidas que luego son unidas para formar una sola.

Consultar ventas desde origen

La extracción de las tablas de ventas emitidas de la base de negocios origen, se lo realiza mediante la siguiente consulta:

```

select b.id_bodega as Sucursal, b.vendedor,
a.id_item,b.fecha_entero,
sum(((a.cantidad_unid*c.num_fraccion)+a.cantidad_frac)) as

```

```

Cant_Unidades, round (sum(a.neto_1),2) as neto_1, round
(sum((a.neto_1-a.costo_total_1)),2) as Utilidad, '1' as Estado
from ve_venta_det a inner join ve_venta_cab b on
a.id_venta_cab=b.id_venta_cab

inner join in_producto c on a.id_item = c.id_producto

group by fecha_entero, b.id_bodega, b.vendedor, a.id_item

```

Consulta no vendidos desde origen

En esta tarea se extraen los datos de los no vendidos con el siguiente query:

```

Select top 20 id_bodega as sucursal, a.usuario as vendedor,
id_item, convert (varchar,fecha_hora,112) as fecha_entero,
sum((a.unidades*b.num_fraccion)+ a.fracciones)as cant_venta,
convert(money,0) as neto_1, convert(money,0) as Utilidad, '2'as
Estado from ve_faltante a

inner join in_producto b on a.id_item = b.id_producto where
a.unidades > b.num_fraccion

group by convert (varchar,fecha_hora,112),id_bodega, a.usuario,
id_item

```

Para mantener el mismo formato de los campos y luego se pueda realizar la unión de las dos tablas, se utiliza la herramienta de conversión de datos de los campos que no coincidan en su tipo o tamaño.

Conversión de datos

El campo fecha que se obtiene de la tabla faltantes referente a los no vendidos no coincide en el formato, por lo que se cambia a entero.

Unión de todo

Se unen las dos tablas.

Búsquedas

Mediante esta herramienta se realiza el enlace de los respectivos campos claves de las dimensiones entre la tabla de ventas desde origen y las tablas de dimensiones del Datamart.

Las configuraciones de estas búsquedas se definen de la siguiente manera:

Tabla 8: Búsquedas para Fact_ventas

BÚSQUEDAS PARA FACT_VENTAS		
Búsqueda	Columna de Entrada - desde Origen	Columna disponible en Datamart
Dim_Tiempo	Fecha_entero	Fechakey
Dim_Sucural	Sucursal	Sucursalkey
Dim_vendedor	vendedor	Vendedorkey
Dim_producto	Id_item	Productokey
Dim_tipo	Estado	TipoKey

Cargar ventas en DM

Se cargan los datos en la tabla del Datamart Fact_ventas.

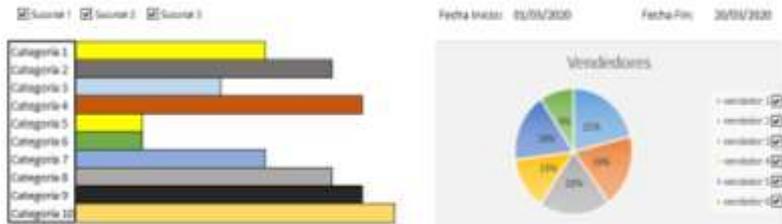
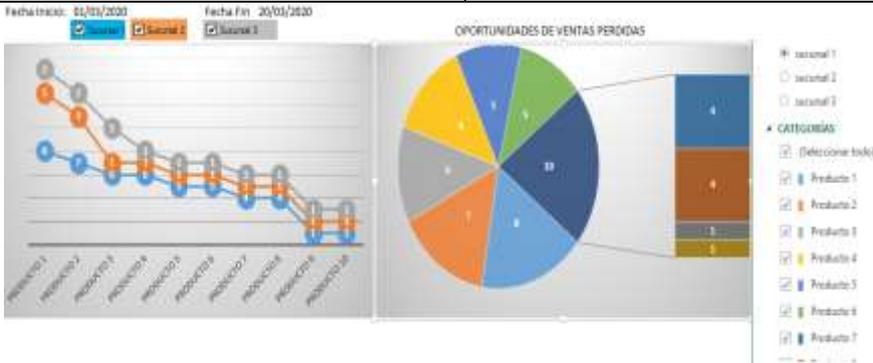
4.5 Diseño y creación de los paneles de información

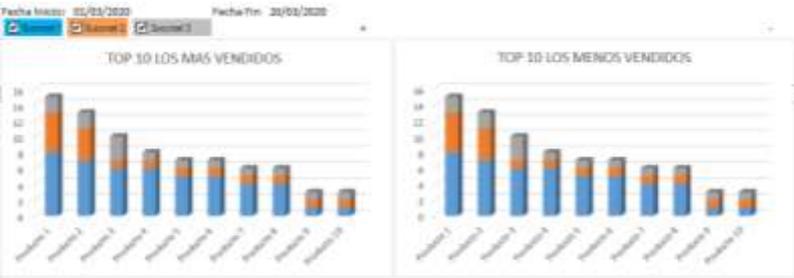
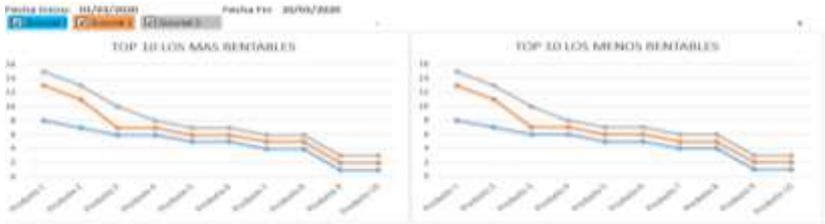
4.5.1 Diseño de paneles

Para el diseño de estos paneles se ha tomado en cuenta los requerimientos de los usuarios, definición en conjunto con el equipo de trabajo y usuario piloto, llevando estos diseños hacia la estandarización para que sean de utilidad en cada área de ventas de todos los clientes que usan el ERP.

Se crean estos paneles, bajo el criterio del usuario, necesitarán de uno y hasta dos clics, es decir; los paneles al iniciar la aplicación mostrarán el reporte con información estándar que contará con los datos del mes en curso y cualquier variación sobre el requerimiento de visualización del usuario, se podrá realizar con el menor número de clics posible.

Tabla 9: Diseño de Paneles

DISEÑO DE PANELES	
Panel	Descripción
Ventas por categorías, por local y vendedor	<p>Jerarquías por Dimensión:</p> <ul style="list-style-type: none"> • Sucursal • Vendedor • Fecha
	
Oportunidades de ventas perdidas detalladas por productos y por local	<p>Jerarquías por Dimensión:</p> <ul style="list-style-type: none"> • Sucursal • Fecha
	
Curva ABC de productos (Top 10 de productos con mayor y menor rotación)	<p>Jerarquías por Dimensión:</p> <ul style="list-style-type: none"> • Sucursal • Tipo Rotación: Alta/Baja • Fecha

DISEÑO DE PANELES	
Panel	Descripción
	
<p>Curva ABC de rentabilidad (Top 10 de productos vendidos con mayor y menor rentabilidad)</p>	<p>Jerarquías por Dimensión:</p> <ul style="list-style-type: none"> • Sucursal • Tipo Rentabilidad: Alta o Baja • Fecha
	

4.5.2 Creación de Paneles

Para la creación de los paneles se utilizará la herramienta Power BI de Microsoft en su versión de escritorio, donde se crea la cadena de conexión hacia el Datamart, para tener acceso a los datos ya cargados durante el proceso ETL.

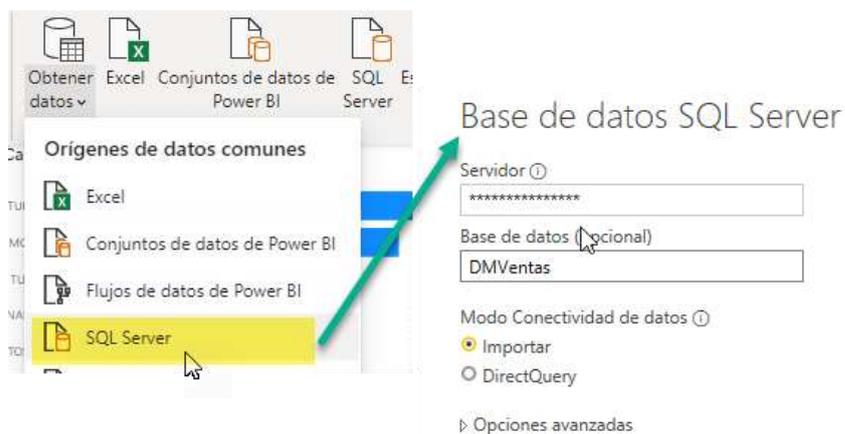


Figura 4.24: Conexión Power BI Datamart

Una vez realizada la conexión, se tienen disponibles las tablas que contienen el Datamart según lo muestra la Figura 4.25.



Figura 4.25: Campos en Power BI

De esta manera se agregan los campos a utilizar en cada uno de los paneles definidos para este proyecto en la Sección 4.5.1.

Panel de ventas por categoría, local y vendedor

Para este panel se crea una nueva medida llamada **Suma_Total** bajo la siguiente formula: $\text{Round}(\text{SUM}(\text{FACT_VENTAS}[\text{Total_Venta}]);2)$, la cual es la sumatoria de los valores netos de cada venta. Cada gráfico utiliza los siguientes campos de las tablas de dimensiones:

Tabla 10: Campos en BI - Panel ventas por categorías, local y vendedor

PANEL VENTAS POR CATEGORÍAS, LOCAL Y VENDEDOR	
Tabla	Campo
Gráfico de barras agrupadas	
Dim_producto	Categoría
Fact_ventas	Total_venta
Gráfico circular	
Dim_vendedor	Nombre
Fact_ventas	Suma_Total (Nueva medida)
Gráfico de segmentación de datos	
Fact_ventas	Fechakey
Dim_sucursal	Nombre
Gráfico de tabla	
Fact_ventas	Suma_Total (Nueva medida)
Fact_ventas	Total_utilidad

Al contener la tabla de hechos Fact_ventas los registros tanto de ventas emitidas como ventas perdidas, es necesario incluir un filtro con el tipo de transacción 1, que corresponde a ventas emitidas, según lo muestra en la Figura 4.26.

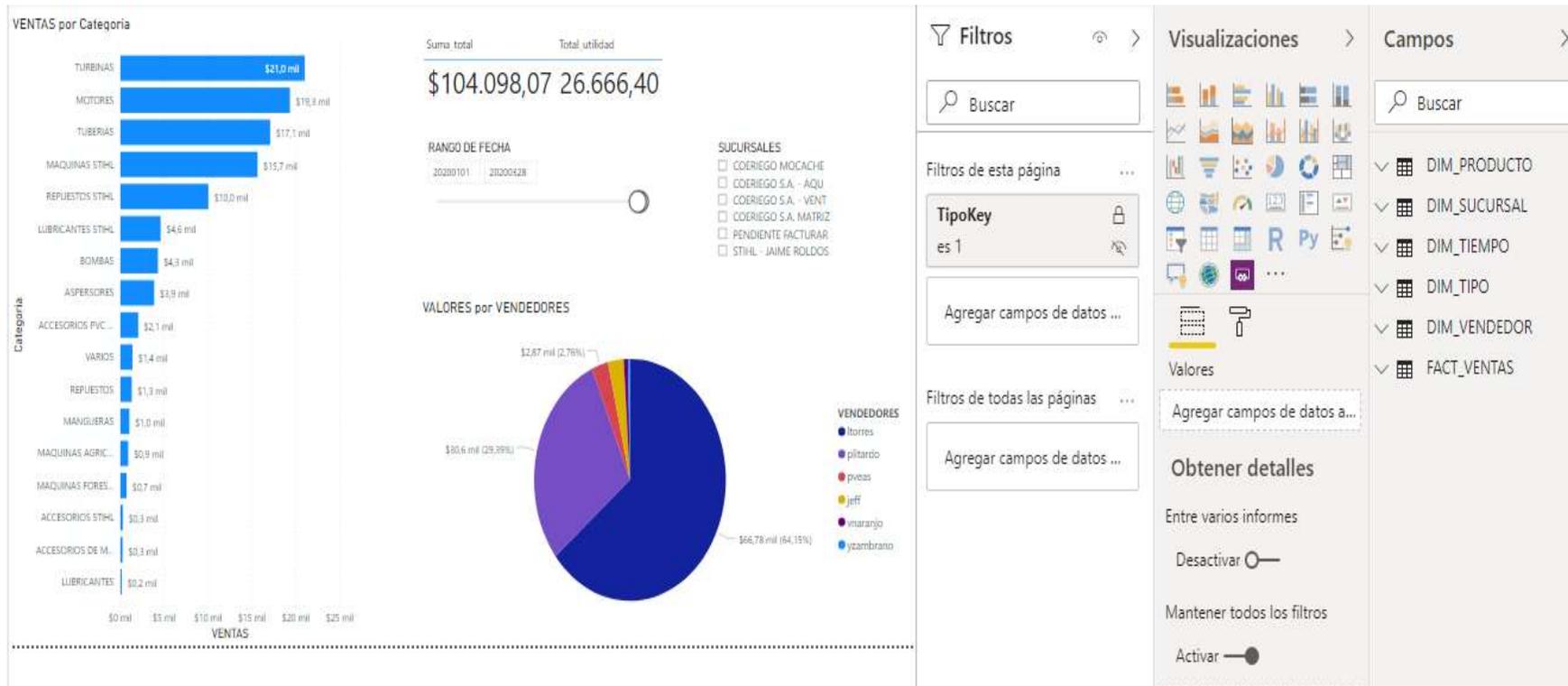


Figura 4.26: Panel de ventas por categoría, local y vendedor

Panel de oportunidades de ventas perdidas

Este reporte esta detallado por producto y por local. Utiliza un filtro por tipo de transacción 2, que corresponde a las ventas no realizadas, además se utiliza un filtro que permite mostrar un número limitado de resultados, ya que el exceso de productos listados hace que el reporte se distorsione y no se pueda visualizar, es por esto que la cantidad a mostrar serán los diez productos con mayor cantidad de unidades no vendidas. Los campos y tablas utilizadas en el panel son los siguientes:

Tabla 11: Campos en BI Panel Ventas perdidas por producto y sucursal

PANEL VENTAS PERDIDAS POR PRODUCTO Y SUCURSAL	
Tabla	Campo
Gráfico de líneas	
Dim_producto	Descripción
Dim_sucursal	Nombre
Fact_ventas	Cant_venta
Gráfico de anillos	
Dim_sucursal	Nombre
Fact_ventas	Cant_venta
Gráfico de segmentación de datos	
Fact_ventas	Fechakey
Dim_sucursal	Nombre

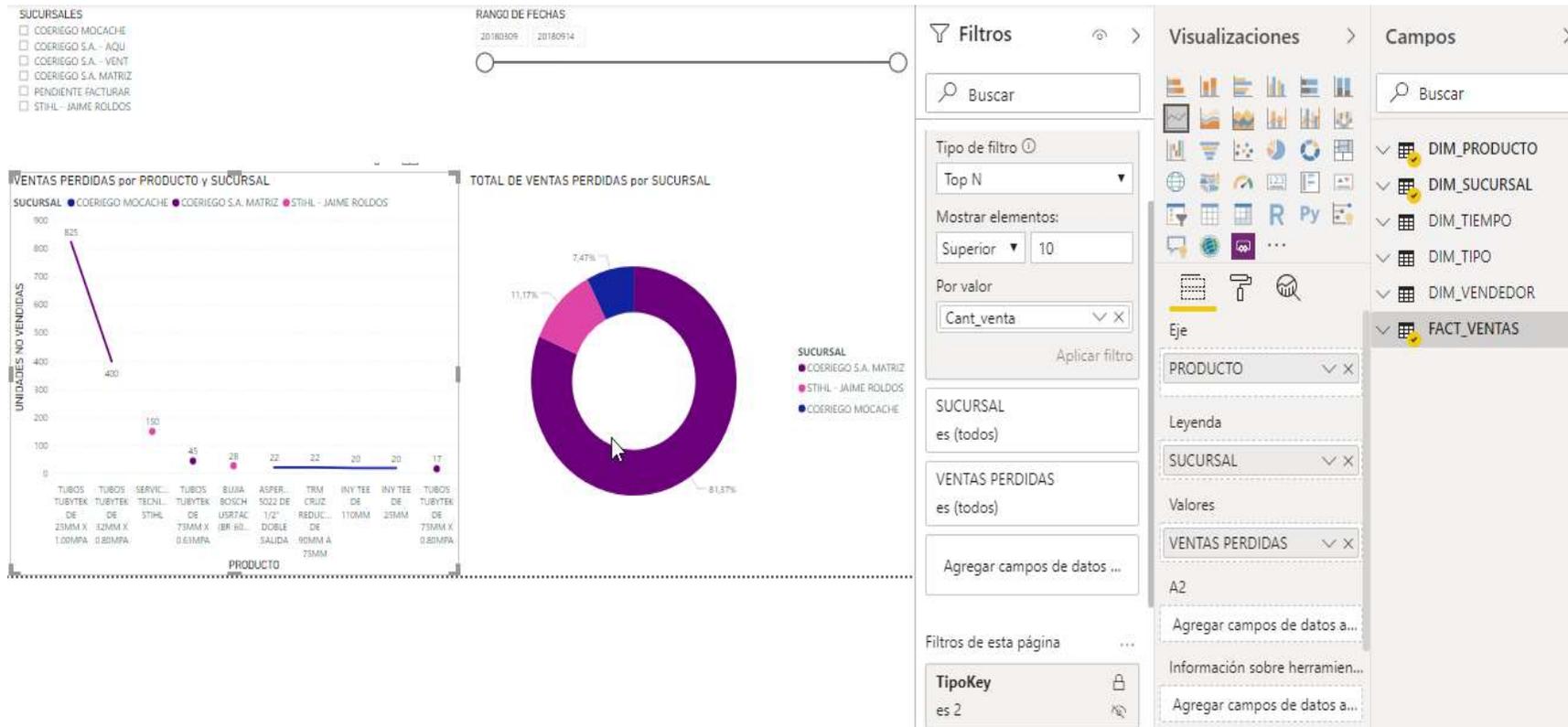


Figura 4.27: Panel ventas perdidas por producto y sucursal

Panel Curva ABC de productos

Reporte detallado por producto y por sucursales de los productos de mayor y menor rotación en ventas. Este reporte utiliza el filtro de transacción de tipo 1 que corresponde a las ventas emitidas, además de utilizar un filtro de tipo top N para mostrar la cantidad de datos requeridos para el panel; uno de los gráficos se parametriza con el tipo de elementos top superiores para los productos de mayor rotación y el otro con el tipo de elementos top inferiores para los de menor rotación. Las tablas y campos utilizados son los siguientes:

Tabla 12: Campos en BI Panel Curva ABC de productos

PANEL CURVA ABC DE PRODUCTOS	
Tabla	Campo
Gráfico de columnas apiladas	
Dim_producto	Descripción
Dim_sucursal	Nombre
Fact_ventas	Cant_venta
Gráfico de anillos	
Dim_producto	Descripción
Dim_sucursal	Nombre
Fact_ventas	Cant_venta
Gráfico de segmentación de datos	
Fact_ventas	Fechakey
Dim_sucursal	Nombre

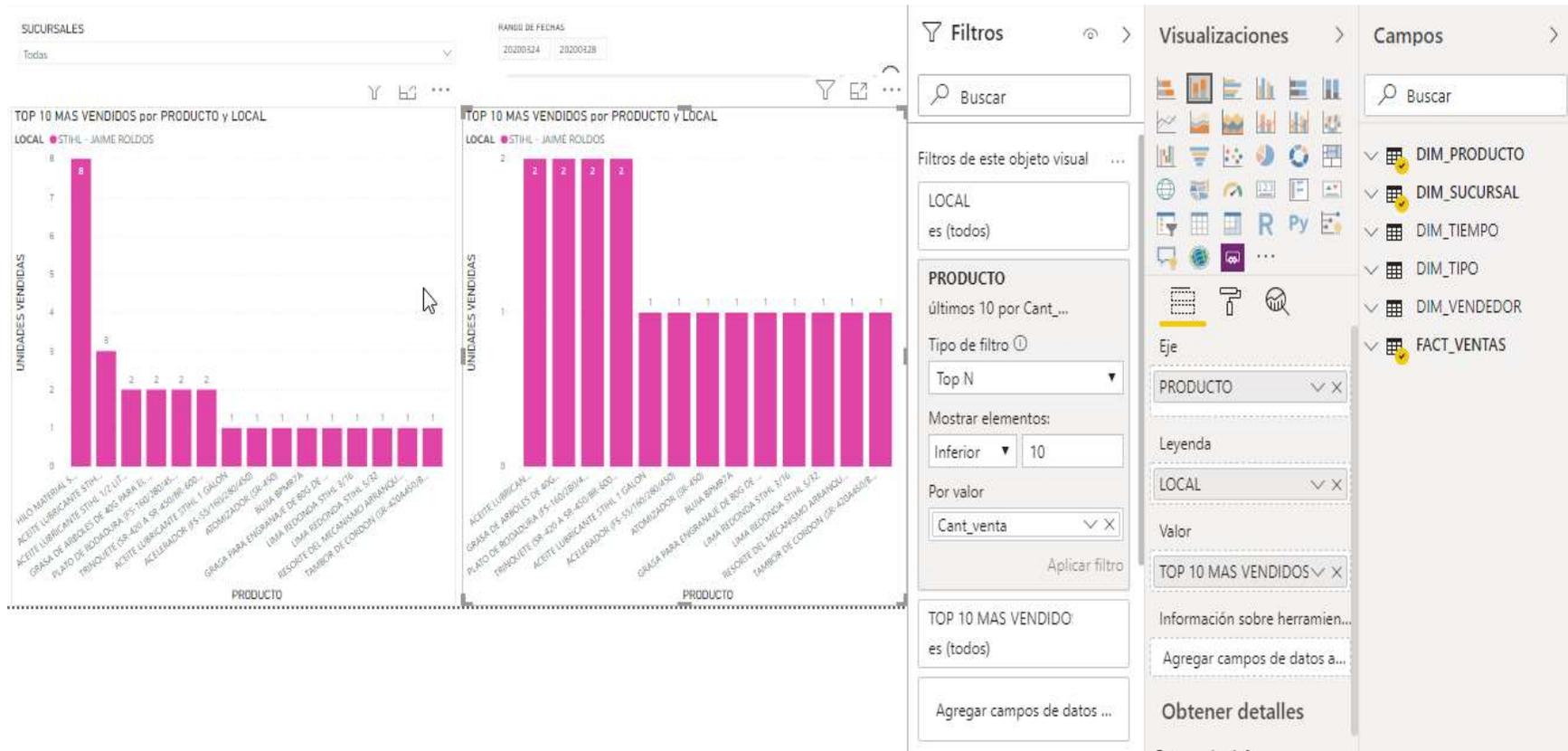


Figura 4.28: Panel ABC Productos

Panel ABC Rentabilidad

Reporte detallado por producto y por sucursales de los productos de mayor y menor rentabilidad en ventas. Este reporte utiliza el filtro de transacción de tipo 1 que corresponde a las ventas emitidas, además de utilizar un filtro de tipo top N para mostrar la cantidad de datos requeridos para el panel; uno de los gráficos se parametriza con el tipo de elementos top superiores para los productos de mayor rotación y el otro con el tipo de elementos top inferiores para los de menor rotación. Las tablas y campos utilizados son los siguientes:

Tabla 13: Campos en BI Panel Curva ABC de rentabilidad

PANEL CURVA ABC DE RENTABILIDAD	
Tabla	Campo
Gráfico de líneas	
Dim_producto	Descripción
Fact_ventas	Cant_venta
Gráfico de columnas apiladas	
Dim_producto	Descripción
Fact_ventas	Cant_venta
Gráfico de segmentación de datos	
Fact_ventas	Fechakey
Dim_sucursal	Nombre

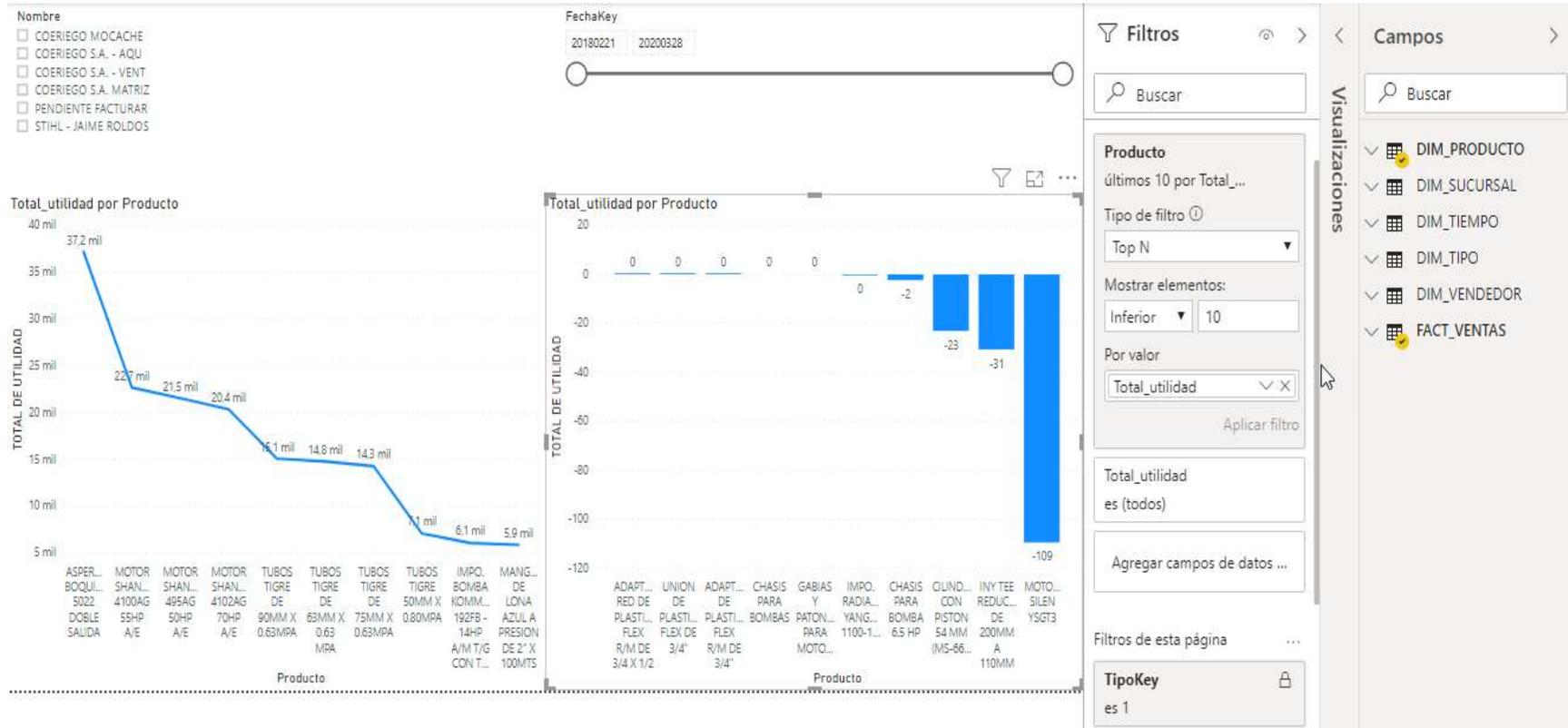


Figura 4.29: Panel ABC rentabilidad

CAPÍTULO 5

IMPLEMENTACIÓN Y ANÁLISIS DE RESULTADOS

5.1 Plan de pruebas

Las pruebas necesarias consistieron en utilizar un ambiente de control de calidad que se ubica en un servidor local en las instalaciones de la empresa desarrolladora de software, que contiene los respaldos de las bases de datos transaccionales de los clientes que mantienen arrendamiento del ERP en hosting. Los actores responsables para realizar dichas pruebas son los siguientes:

Tabla 14: Actores responsables en pruebas

ACTORES RESPONSABLES EN PRUEBAS	
ACTOR	ROL
Viviana Magallanes	Desarrollador
Oscar Vincés	Administrador de base de datos

ACTORES RESPONSABLES EN PRUEBAS	
ACTOR	ROL
Melissa Muñoz	Cliente

El periodo que se destina para realizar estas pruebas es de dos días, donde se presentarán dos variantes en la actualización de los datos:

- Día 1 - Actualización 00:00.
- Día 2 - Actualización 00:00 y 12:00.

El objetivo de probar con esas dos variantes es que el usuario evalúe la funcionalidad de contar con información del día en curso con una actualización a las 12 del mediodía y otra solo hasta el día anterior. El administrador de base de datos comprobará el comportamiento del sistema al actualizar en horas hábiles donde los usuarios trabajan normalmente.

Categorías de resultados:

Exitoso: Cumple en su totalidad con los resultados esperados.

Error: No cumple con los resultados esperados.

Tolerable: Cumple con los resultados, pero no con el desempeño esperado.

Criterios a evaluar:

Criterios con categoría exitoso y error:

- Éxito de las tareas de ETL de dimensiones.
- Consistencia en los datos cargados en las dimensiones.
- Éxito en las tareas de ETL de hechos.
- Consistencia en los datos cargados en hechos.

Criterios con categoría exitoso, error y tolerable:

- Visualización de datos en paneles del usuario.
- Consumo de recursos del servidor al momento de ejecutar las tareas del nuevo proceso.
- Tiempo de ejecución de las tareas del ETL.

5.1.1 Pase para implementación en ambiente de pruebas**Objetivo**

Implementación de proceso ETL y paneles de información en ambiente de QA.

Archivos de implementación (DBA)

Para la creación del Datamart y su estructura, se utilizarán los siguientes archivos que se envían al administrador de base de datos:

Crear_dm.sql

Crear_estructura.sql

Tabla 15: Archivos implementación para pruebas

ARCHIVOS DE IMPLEMENTACIÓN PARA PRUEBAS			
Servidor	Instancia	Base de datos	Archivo
Srvqdesarrollo	SQL2017	Master	Crear_dm.sql
Srvqdesarrollo	SQL2017	Master	Crear_estructura.sql

Paquetes de SSIS (Desarrollo)

Para la implementación del paquete de SSIS en el ambiente de control de calidad disponible para las pruebas, se ejecuta la acción implementar desde la aplicación visual estudio donde se encuentran las tareas que se han creado para el proceso ETL.

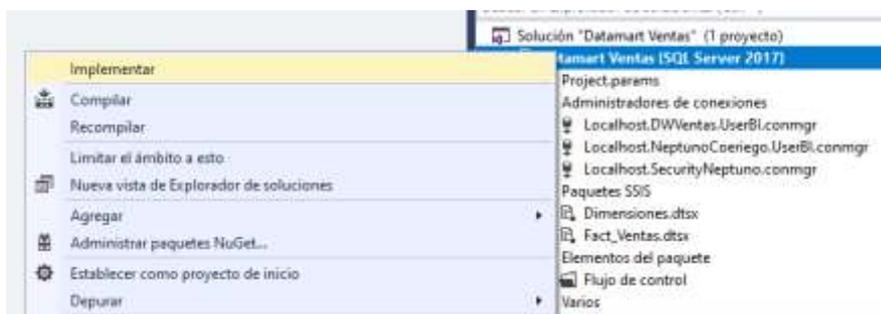


Figura 5.1: Implementación paquete SSIS en pruebas

Utilizando los datos establecidos para formar la cadena de conexión.

Tabla 16: Cadena de conexión servidor de pruebas

CADENA DE CONEXIÓN SERVIDOR DE PRUEBAS			
Servidor	Instancia	Usuario	Contraseña
Srvqadesarrollo	SQL2017	UserBI	*****

Definición de tarea programada

Para la ejecución de este proceso se establece una programación automática dentro de dos Jobs en el SQL Server, que para este caso será una configuración para el primer día con una actualización fuera de horarios laborales y la segunda con una actualización adicional en horas laborales según se lo indica en la sección 5.1.

Tabla 17: Job DM Ventas SSIS pruebas

JOB DM VENTAS SSIS PRUEBAS	
Generales	
Nombre del Job	JobDMVentasCoeriego1
Pasos	
Nombre	JobDMVentasCoeriego1
Tipo	Sql Server Integration Services Packages
Servidor	Srvqadesarrollo/Sql2017
Autenticación	UserBI
Paquete	Paquete SSIS creado en paso anterior
Programación	
Nombre	JobDMVentasCoeriego1
Tipo de tarea	Recurrente
Fecha y hora	06/05/2020 00:00

Tabla 18: Job DM ventas SSIS pruebas 2

JOB DM VENTAS PRUEBAS 2	
Generales	
Nombre del Job	JobDMVentasCoeriego2
Pasos	
Nombre	JobDMVentasCoeriego2
Tipo	Sql Server Integration Services Packages
Servidor	Srvqadesarrollo/Sql2017
Autenticación	UserBI
Paquete	Paquete SSIS creado en paso anterior
Programación	
Nombre	JobDMVentasCoeriego2
Tipo de tarea	Una sola vez
Fecha y hora	07/05/2020 12:00

5.2 Resultados de las pruebas

A continuación, se detallan los resultados obtenidos de las pruebas realizadas.

Tabla 19: Resultados de pruebas

RESULTADOS DE LAS PRUEBAS		
Tiempo de ejecución de las tareas del ETL	Exitoso	1.- Carga inicial: 7 segundos. 2.- Primera Actualización: 6 segundos. 3.- Segunda Actualización: 6 segundos.
Éxito de las tareas de ETL de dimensiones	Exitoso	
Consistencia en los datos cargados en las dimensiones	Exitoso	
Éxito en las tareas de ETL de hechos	Exitoso	
Consistencia en los datos cargados en hechos	Exitoso	
Visualización de datos en paneles del usuario	Exitoso	
Consumo de recursos del servidor al momento de ejecutar las tareas del nuevo proceso	Exitoso	No se visualiza consumo de recurso anormal.

Las pruebas resultaron satisfactorias en todas las fases y sus respectivos actores pudieron ejecutar las tareas sin novedades.

La carga de datos desde las bases transaccionales hacia el Datamart fluye sin ningún problema por lo que el Administrador de base de datos no observa ningún comportamiento inusual en el servidor, ni en el ERP durante de la ejecución de los procesos. Las tareas de visualización en los paneles para el usuario fueron exitosas, el desarrollador verificó la actualización de datos y la publicación de estos en la aplicación del usuario.

Por parte del usuario final, este ingresó a la aplicación verificando los paneles activos y la correcta interacción con las opciones disponibles. Para la prueba de revisar la utilidad de tener una actualización a mitad del día en curso, se verifica que no es necesario debido a que las cajas están en apertura por lo que los valores pueden variar, lo que puede provocar una percepción errónea. En conjunto con el equipo de trabajo, se realizó la validación de la consistencia de los datos mostrados en la aplicación Microsoft Power BI con los datos de la base transaccional original.

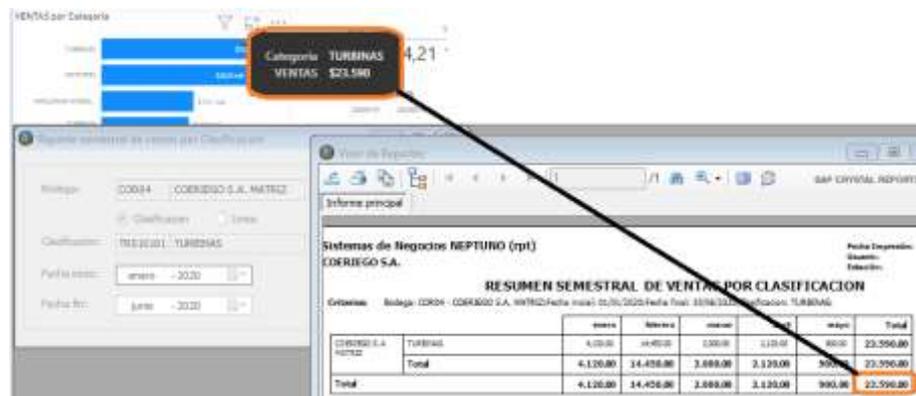


Figura 5.3: Evidencia de cuadro de datos clasificación

Se realiza una nueva validación sobre el producto más rentable dentro de toda la compañía durante el periodo 2020.



Figura 5.4: Evidencia cuadro de datos rentabilidad

El panel que corresponde a la información del producto más vendido por local, indica el siguiente resultado:

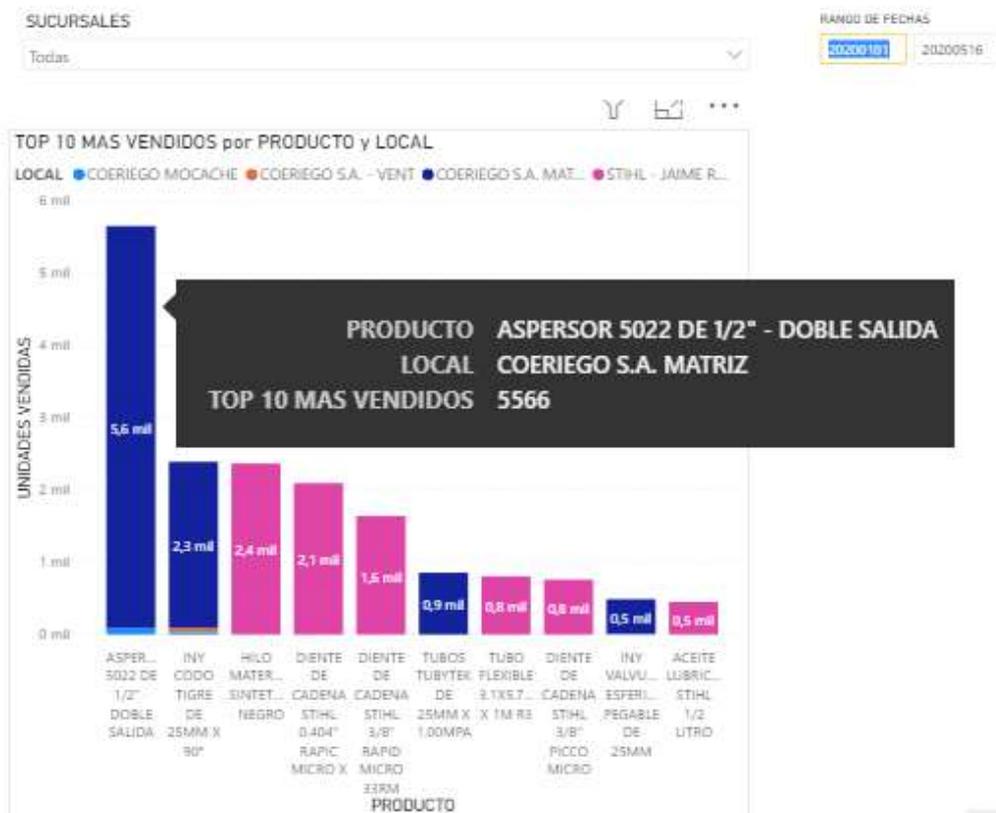


Figura 5.5: Evidencia cuadro de datos, más vendidos

Se valida en los reportes del ERP (ver figura a continuación), que el resultado es el mismo.

Sistemas de Negocios NEPTUNO (rpt)
COERIEGO S.A.

CURVA ABC DE UNIDADES VENDIDAS

Criterios: Bodega: 1; Fecha inicial: 01/01/2020; Fecha final: 18/05/2020;

Id Item	Descripción	Presentación	Frac.	Fabri.	Cant. Unit	Valoriz.	%Valor	V. Ac
COERIEGO S.A. MATRIZ								
319	FUMIGADORA DE ESPALDA YARMA DE 16 L FUMIGADORA DE ESPAL		1	TRACT	130	11.849,78	14,67	11,8
3219	TURBINA A CAUDAL DE 25HBC	NINGUNO	1	ALESS	1	9.200,00	11,76	21,3
3274	TURBINA A CAUDAL DE 14HBC	NINGUNO	1	ALESS	3	3.670,69	7,26	27,2
1872	MOTOR SHANGHAI 4332AG 70HP A/E	4102AG	1	TRACT	1	4.404,29	5,53	31,6
3218	TUBO NOVAPORT PLUS DE 975MM X 6M	NINGUNO	1	ABL.	7	4.042,00	5,03	35,7
1462	MOTOR SHANGHAI 493AG A/E DIESEL - 1495		1	ALESS	1	3.392,86	4,20	39,1
1806	ASPERSOR 5022 DE 1/2" - DOBLE SALIDA DOBLE SALIDA		1	ALESS	5.836	3.330,36	4,12	42,4

Figura 5.6: Comparativo más vendidos

5.3 Plan de implementación

Objetivo

Implementación de proceso ETL y paneles de información en ambiente de producción en el servidor de AWS.

Archivos de implementación (DBA)

Para la creación del Datamart y su estructura, se utilizarán los siguientes archivos que son enviados al administrador de base de datos:

Crear_dm.sql

Crear_estructura.sql

Tabla 20: Archivos de implementación para producción

ARCHIVOS DE IMPLEMENTACIÓN PARA PRODUCCIÓN			
Servidor	Instancia	Base de datos	Archivo
Srvdirgyeapli01	SQL2017	Master	Crear_dm.sql
Srvdirgyeapli01	SQL2017	Master	Crear_estructura.sql

Paquetes de SSIS (Desarrollo)

Para la implementación del paquete de SSIS en el ambiente de producción, se ejecuta la acción implementar desde la aplicación visual estudio donde se encuentran las tareas que se han creado para el proceso ETL.

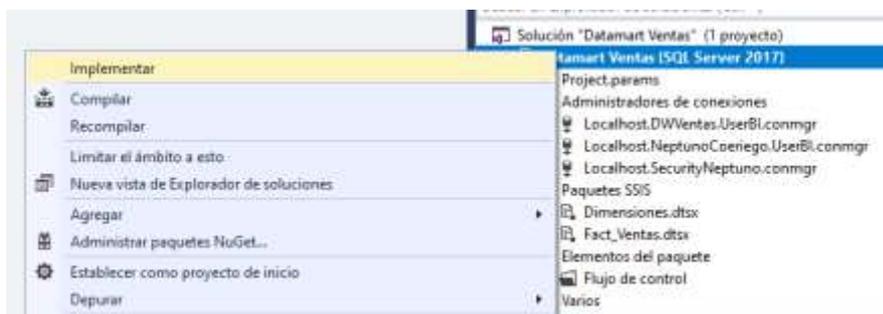


Figura 5.7: Implementación paquete SSIS en producción

Utilizando los datos establecidos para formar la cadena de conexión.

Tabla 21: Cadena de conexión servidor de producción

CADENA DE CONEXIÓN SERVIDOR DE PRODCUCCIÓN			
Servidor	Instancia	Usuario	Contraseña
Srvdirgyeapli01	SQL2017	UserBI	*****

Definición de tarea programada

Para la ejecución del paquete de SSIS, se establece una programación automática dentro del SQL Server.

Tabla 22: Job DM Ventas SSIS producción

JOB DM_VENTAS SSIS PRODUCCIÓN	
Generales	
Nombre del Job	JobDMVentasCoeriego
Pasos	
Nombre	JobDMVentasCoeriego
Tipo	Sql Server Integration Services Packages
Servidor	Srvdirgyeapli01/Sql2017
Autenticación	UserBI
Paquete	Ventas_Coeriego
Programación	
Nombre	JobDMVentasCoeriego

JOB DM_VENTAS SSIS PRODUCCIÓN	
Tipo de tarea	Recurrente
Fecha y hora	00:00

Publicación de los paneles

Mediante la herramienta con la que se construyó los paneles (Microsoft Power BI), se publican los paneles con el usuario creado específicamente para cada usuario, lo que se realiza a través de correos que tienen licenciamiento de dicha herramienta.

5.4 Análisis de resultados

Beneficios

Los beneficios que brinda la implementación de esta solución están orientados hacia dos frentes, al cliente final y a la empresa desarrolladora de software dueña del ERP. Con respecto al beneficio para la empresa desarrolladora de software, se brinda una solución que permita aprovechar los recursos disponibles dentro de la empresa que no se han explotado en su totalidad, teniendo como resultado una herramienta que se puede ofrecer como un servicio agregado a los ya existentes, lo que hace más atractiva la cartera de

productos que ofrece. Es importante tener en cuenta que también significa una disminución en el uso de los recursos de la empresa al generar reportes personalizados y luego incluirlos en una actualización como se hace normalmente. Además del crecimiento profesional del personal de la empresa, ya que desarrolla nuevas habilidades y experiencia en estas herramientas que son indispensables en el mercado actual.

Al cliente final le representa optimización de tiempo, al requerir informes con datos consolidados orientados a la toma de decisiones de manera autorizada, con la confianza de contar con datos consistentes y ya no tener que buscar terceros para realizar estos trabajos. Adicionalmente el beneficio de poder acceder a la información desde cualquier dispositivo, desde cualquier lugar y a cualquier hora, centralizando diversos servicios con un mismo proveedor.

Riesgos

Este proyecto describe los beneficios de utilizar las herramientas de inteligencia de negocios y el servicio que se puede ofrecer a empresas de cualquier tamaño, pero la implementación y el uso de

este en un escenario empresarial real es determinado por las personas quienes dirigen la empresa. El uso de todas las herramientas está bajo licenciamiento y dado que la empresa mantiene un estatus de partner gold con la marca, tiene beneficios de paquetes de licencias gratuitas de todos sus productos por lo que algún cambio en este estado afectaría el uso de dichas licencias.

Al tratarse de un servicio agregado que se propone brindar a los usuarios que consumen el servicio de hosting en AWS, pudieran pedir generar informes de otro tipo consumiendo más recursos de trabajo de lo previsto, por lo que se debe analizar la inclusión de este servicio bajo condiciones que incluyan los costos de algún desarrollo adicional.

CONCLUSIONES Y RECOMENDACIONES

Conclusiones

1. Todos los sistemas utilizados por las empresas para el registro de transacciones, contienen información importante que puede ser aprovechada mediante los procesos adecuados para procesarla y convertirla en información útil.
2. Las soluciones de Inteligencia de Negocios (BI) proveen de información organizada y consolidada que puede ser aplicada a una o a todas las áreas de una empresa como apoyo en la toma de decisiones.
3. Automatizar el proceso de extracción y tratamiento de la información ayuda en la disminución del uso de recursos de tiempo y esfuerzo al personal de TI, además de conseguir eliminar los errores que pudieren ser cometidos por las personas al realizar estos procesos de forma manual.
4. La subutilización de recursos limita el crecimiento de las empresas ya que existen herramientas tecnológicas para mejorar los procesos de la organización que no se utilizan, ya sea por falta de conocimiento o visión de quienes la dirigen.
5. Existen muchas herramientas para la implementación de tecnologías de almacenamiento de datos e inteligencia de negocios, ya sea software propietario o software libre, lo importante es hacer uso de ellas para obtener el mayor beneficio posible.

Recomendaciones

1. Promocionar e incentivar el uso de la inteligencia de negocios en las empresas clientes como apoyo en la toma de decisiones, utilizando la tecnología como aliado estratégico.
2. Socializar entre los clientes de la empresa dueña del ERP estos modelos de paneles de información básicos y estándares como servicio agregado, dejando abierta la opción de personalización y creación de nuevos paneles bajo un costo establecido.
3. Al ser este un proyecto piloto se debe realizar en el futuro, en caso de ser implementado, un análisis de acuerdo al número de usuarios recurrentes usando esta aplicación para reestructurar la ejecución de estos procesos y no afectar el rendimiento del ERP.

BIBLIOGRAFÍA

- [1] powerdata, «powerdata,» [En línea]. Available: <https://www.powerdata.es/data-warehouse>. [Último acceso: 08 Febrero 2020].
- [2] sinnexus, «sinnexus,» [En línea]. Available: https://www.sinnexus.com/business_intelligence/piramide_negocio.aspx. [Último acceso: 09 02 2020].
- [3] N. Diazgranados, «estrategiaynegocios.net,» 07 Julio 2016. [En línea]. Available: <https://www.estrategiaynegocios.net/opinion/977752-345/diferencia-entre-dato-informaci%C3%B3n-y-conocimiento>.
- [4] J. Castro, «/blog.corponet.com.,» 12 Agosto 2015. [En línea]. Available: <https://blog.corponet.com.mx/que-es-la-inteligencia-de-negocios>.
- [5] «aws.amazon.com,» [En línea]. Available: <https://aws.amazon.com/es/about-aws/>. [Último acceso: 20 Febrero 2020].
- [6] «docs.microsoft.com,» 07 Junio 2018. [En línea]. Available: <https://docs.microsoft.com/en-us/sql/integration-services/sql-server-integration-services?view=sql-server-ver15>.
- [7] «microsoft.com,» 15 Agosto 2019. [En línea]. Available: <https://docs.microsoft.com/es-es/sql/ssdt/download-sql-server-data-tools-ssdt?view=sql-server-ver15>.
- [8] «powerbi.microsoft.com,» [En línea]. Available: <https://powerbi.microsoft.com/es-es/what-is-power-bi/>. [Último acceso: 20 Febrero 2020].
- [9] D. Prudenciano, «tecnomark,» [En línea]. Available: <http://tecnomark.es/los-almacenes-de-datos-data-warehouse/>. [Último acceso: 15 Febrero 2020].
- [10] W. H. Inmon, «Building the Data warehouse,» *Indianapolis : Wiley Publishing, Inc., 2005*.
- [11] M. Á. R. Rojas, «<http://scielo.br/>,» [En línea]. Available: <http://www.scielo.br/pdf/ci/v34n2/28555.pdf>. [Último acceso: 15 02 2020].

- [12] E. Editorial, «reportedigital.com,» reportedigital, 31 Julio 2018. [En línea]. Available: <https://reportedigital.com/cloud/beneficios-data-warehouse-toma-decisiones-organizacionales/>.
- [13] sinnexus.com, «sinnexus.com,» [En línea]. Available: https://www.sinnexus.com/business_intelligence/datamart.aspx. [Último acceso: 12 Febrero 2020].
- [14] «esan.edu.pe,» 07 Marzo 2019. [En línea]. Available: <https://www.esan.edu.pe/apuntes-empresariales/2019/03/que-es-data-mart-y-por-que-es-importante-implementarlo-en-tu-empresa/>.
- [15] S. Pearlman, «es.talend.com,» 19 Febrero 2019. [En línea]. Available: <https://es.talend.com/resources/what-is-data-mart/>.
- [16] J. García, «mundodb.es,» 09 Noviembre 2013. [En línea]. Available: <http://mundodb.es/disenio-data-warehouse-hechos-y-dimensiones-modelo-estrella-vs-copo-de-nieve>.
- [17] S. Pearlman, «es.talend.com,» 19 Agosto 2019. [En línea]. Available: <https://es.talend.com/resources/what-is-etl/>.
- [18] «blog.powerdata.es,» 13 Junio 2013. [En línea]. Available: <https://blog.powerdata.es/el-valor-de-la-gestion-de-datos/bid/288859/procesos-etl-extracci-n-transformaci-n-carga>.