

ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL



Facultad de Ingeniería en Electricidad y Computación

**“IMPLEMENTACIÓN DE UN CUBO DE INFORMACIÓN PARA LA
SEGMENTACIÓN DE CLIENTES DE UNA EMPRESA DE
PRESTACIONES DE SERVICIOS FINANCIEROS”**

TRABAJO DE TITULACIÓN

Previo a la obtención del Título de:

MAGÍSTER EN SISTEMAS DE INFORMACIÓN GERENCIAL

Presentado por

EVELYN TATIANA PONGUILLO RODRÍGUEZ

Guayaquil – Ecuador

2018

AGRADECIMIENTO

A Dios todopoderoso, quien me ha llenado de bendiciones y forjado mi camino, dirigiéndolo por el sendero correcto, el que en todo momento está a mi lado apoyándome, agradecida por su infinita misericordia y amor reflejado en mi vida, permitiéndome alcanzar nuevas metas.

Agradezco a mis padres Esther y Alberto porque ellos son mi motivación y razón para continuar día a día luchando para alcanzar mis anhelos, a ellos les debo todos mis logros por su paciencia y apoyo incondicional, todo lo que soy es gracias a ellos.

A mis hermanos y amigos que siempre tuvieron unas palabras de aliento para no desmayar, y que me apoyaron en todo lo que estuvo a su alcance.

DEDICATORIA

Dedico este proyecto a Dios porque me brindó la fuerza necesaria para continuar en todo momento, guiando cada paso para llegar a la meta.

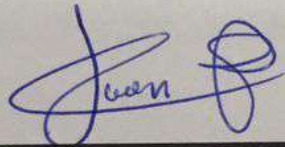
Dedico este proyecto a Dios, a mi familia y amigos porque han sido parte fundamental en mi vida, brindándome todo su apoyo y voz de aliento para continuar y no desmayar, ayudándome a superar día a día.

TRIBUNAL DE SUSTENTACIÓN



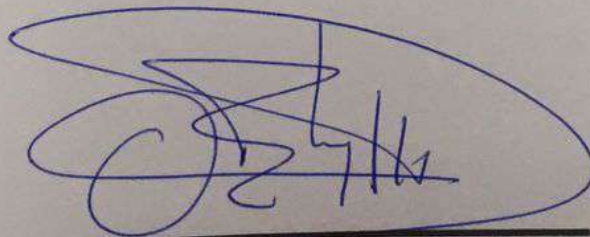
Mgs. Lenin Freire C.

DIRECTOR MSIG



Mgs. Juan Carlos García P.

DIRECTOR DEL PROYECTO DE
GRADUACIÓN



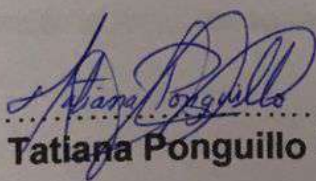
Mgs. Omar Maldonado D.

MIEMBRO DEL TRIBUNAL

DECLARACIÓN EXPRESA

“La responsabilidad del contenido de este Trabajo de titulación, me corresponde exclusivamente; y el patrimonio intelectual de la misma a la ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL”.

(Reglamento de Graduación de ESPOL)



.....
Evelyn Tatiana Ponguillo Rodríguez

RESUMEN

La empresa administradora de servicios financieros requiere informes para la toma de decisiones y mejorar la captación de clientes, para lo cual necesita inmediatez en la respuestas de los servicios e información solicitada, provocando atrasos en otras tareas para solventar la necesidad del área, por lo que el departamento de Gestión de Clientes necesita la implementación de un Datawarehouse (DW) para la segmentación de clientes, mediante un cubo de información o Datamart (DM), el cual es el objetivo principal del proyecto, lo que permitirá facilitar el análisis y toma de decisiones mediante herramientas de Inteligencia de Negocio (BI).

El departamento de gestión del cliente, para cumplir con su objetivo solicita al departamento de sistemas la información necesaria para realizar sus reportes, sin embargo, surgen problemas al momento de realizar dicho requerimiento, como por ejemplo, el informe puede demorar dependiendo de la complejidad hasta 3 días, debido al poco conocimiento del personal sobre la lógica de negocios, por esta misma razón existen problemas de duplicidad de datos e inconsistencia de información, entre otros. Lo que actualmente dificulta la toma de decisiones.

La ejecución del proyecto, inicia con una parte investigativa sobre conceptos de Inteligencia de negocios y segmentación de clientes, con la finalidad de aportar a la solución el mejor esquema de resolución, mediante el diseño del cubo de información y depuración de datos, de esta manera la información estaría disponible en cualquier momento, impactando positivamente los tiempos de respuestas permitiendo la visualización a través de los archivos de Excel para la toma de decisiones oportuna.

El presente trabajo se considera como una etapa inicial para un futuro desarrollo de un área de Inteligencia de Negocios en la empresa, de tal modo que pueda escalar a soluciones de tableros de control y gestión de indicadores.

ÍNDICE GENERAL

AGRADECIMIENTO.....	I
DEDICATORIA.....	II
TRIBUNAL DE SUSTENTACIÓN.....	III
DECLARACIÓN EXPRESA.....	IV
RESUMEN.....	V
ÍNDICE GENERAL.....	VII
ABREVIATURAS Y SIMBOLOGÍA.....	X
ÍNDICE DE FIGURAS.....	XI
ÍNDICE DE TABLAS.....	XIV
1 CAPÍTULO 1.....	1
GENERALIDADES.....	1
1.1 Antecedentes.....	1
1.2 Descripción del Problema.....	2
1.3 Solución Propuesta.....	4
1.4 Objetivo General.....	9
1.5 Objetivos Específicos.....	9
1.6 Metodología.....	10
2 CAPÍTULO 2.....	11

MARCO TEÓRICO	11
2.1 Conceptos básicos de Inteligencia de Negocios.....	11
2.1.1. Conceptos de Componentes de BI	12
2.1.2. Datawarehouse	13
2.1.3. Bodega de datos y Datamart.....	15
2.1.4. Procesos de extracción, transformación y carga de datos.....	17
2.1.5. Cubos de Información	18
2.2 Esquema de Almacén de Datos	19
2.3 Conceptos Básicos de la Metodología a desarrollar	22
2.4 Segmentación de Clientes.....	23
3 CAPÍTULO 3	26
LEVANTAMIENTO Y ANÁLISIS DE LA INFORMACIÓN.....	26
3.1 Recopilación de datos de la situación actual	26
3.2 Documentación del Requerimiento.....	31
3.3 Diagnóstico de Recursos Técnicos.....	44
4 CAPÍTULO 4	46
DISEÑO DE LA SOLUCIÓN.....	46
4.1 Diseño y definición de Modelo de Datos.....	46
4.2 Diseño de procesos de extracción, transformación y carga (ETL)	50
4.3 Diseño de Cubo OLAP	53

4.4	Diseño de reportes	54
5	CAPÍTULO 5	68
	DESARROLLO E IMPLEMENTACIÓN DE LA SOLUCIÓN.....	68
5.1	Plan de Desarrollo e Implementación.....	68
5.1.1.	Proceso de Extracción para Segmentación de Clientes	69
5.1.2.	Carga del proceso de segmentación de clientes	74
5.1.3.	Cubo de Información de Segmentación de Clientes	77
5.2	Plan de Pruebas.....	88
6	CAPÍTULO 6.....	93
	ANÁLISIS DE RESULTADOS.....	93
6.1	Información Integrada y Actualizada	93
6.2	Revisión de resultados obtenidos.....	94
6.3	Mejor uso de Recursos	98
	CONCLUSIONES Y RECOMENDACIONES	101
	BIBLIOGRAFÍA.....	103
	ANEXO - CRONOGRAMA.....	105
	ANEXO - ENCUESTA	107

ABREVIATURAS Y SIMBOLOGÍA

BD:	Base de datos
BI:	Business Intelligence
DM:	Datamart
DS:	Data Source
DWH:	Datawarehouse
ETL:	Extracción, transformación y carga.
HW:	Hardware
OLAP:	On-Line Analytical Processing
OLTP:	On-Line Transaction Processing
SSAS:	SQL Server Analysis Services
SSIS:	SQL Server Integration Services
SW:	Software

ÍNDICE DE FIGURAS

Figura 1.1: Diagrama de Business Intelligence	5
Figura 2.1: Solución BI	13
Figura 2.2: Base de datos operativas vs Base de datos Datawarehouse.....	15
Figura 2.3: Marco general para los procesos de ETL.....	17
Figura 2.4: Diagrama de esquema Snowflake	20
Figura 2.5: Diagrama de un esquema estrella	21
Figura 2.6: Metodología Kimball - Arquitectura Bus del DW.....	22
Figura 2.7: Metodología Inmon DW Corporativo	23
Figura 3.1: Estructura Organizacional.....	26
Figura 3.2: Esquema de la Solución Planteada	31
Figura 3.3: Caso de Uso FrontEnd	42
Figura 3.4: Caso de Uso BackEnd.....	43
Figura 4.1: Esquema de Fuentes de Datos.....	47
Figura 4.2: Tablas temporales	51
Figura 4.3: Arquitectura del DW.....	53
Figura 4.4: Diseño Estrella.....	54
Figura 4.5: Reporte Total de Clientes	56
Figura 4.6: Reporte de Clientes por Ubicación Geográfica	57
Figura 4.7: Reporte de Clientes por Datos Generales.....	58
Figura 4.8: Reporte de Clientes por Datos Laborales	58
Figura 4.9: Reporte de Clientes por Datos Empresariales	59
Figura 4.10: Reporte de Clientes por fecha de ingreso	60
Figura 4.11: Reporte de Clientes activos por datos generales	60

Figura 4.12: Reporte de Clientes activos por Datos Laborales	61
Figura 4.13: Reporte de Clientes activos por Datos Empresariales	62
Figura 4.14: Reporte de Clientes activos por fecha de Ingreso.....	62
Figura 4.15: Clientes Actualizados, con firma y regalo por Datos Generales	63
Figura 4.16: Clientes Actualizados, con firma y regalo por Datos Laborales.....	64
Figura 4.17: Clientes Actualizados, con firma y regalo por datos Empresariales	65
Figura 4.18: Clientes Actualizados por fecha de ingreso	65
Figura 4.19: Clientes con regalo por fechas de Ingreso	66
Figura 4.20: Clientes con firma por fechas de ingreso	66
Figura 5.1: Dimensiones Generales.....	72
Figura 5.2: ETL-Dimensión Calendario (Tiempo).....	73
Figura 5.3: ETL-Hechos.....	74
Figura 5.4: Carga de datos a base consolidada.....	75
Figura 5.5: Carga principal del modelo Estrella.....	76
Figura 5.6: Carga del paquete de clientes	76
Figura 5.7: Implementación Esquema Estrella (Staging Database)	77
Figura 5.8: Desarrollo del cubo de información.....	78
Figura 5.9: Procesamiento del cubo de información	78
Figura 5.10: Activación de Power Pivot en Excel.....	80
Figura 5.11: Opciones de Excel.....	81
Figura 5.12: Activación Power Pivot y Power View	81
Figura 5.13: Vista de Power Pivot en Excel	82
Figura 5.14: Conexión de datos externos	82
Figura 5.15: Conexión Analysis Services.....	83

Figura 5.16: Asistente de Importación de tablas	83
Figura 5.17: Especificación de consulta MDX	84
Figura 5.18: Pantalla de Diseño de Consultas	84
Figura 5.19: Importación de datos	85
Figura 5.20: Datos cargados en Power Pivot	85
Figura 5.21: Pantalla de Diseño Power View	86
Figura 5.22: Reporte de Clientes por Ubicación Geográfica	86
Figura 5.23: Reporte de Clientes por Datos Generales.....	87
Figura 5.24: Reporte de Clientes por Datos Laborales	88

ÍNDICE DE TABLAS

Tabla 1: Roles involucrados en el proceso	29
Tabla 2: Análisis FODA de Gestión de Clientes.....	29
Tabla 3: Listado de dimensiones	34
Tabla 4: Listado de medidas.....	36
Tabla 5: Diseño de Cubo	36
Tabla 6: Diseño de Fórmulas.....	37
Tabla 7: Listado de Reportes.....	39
Tabla 8: Recursos de Hardware y Software.....	45
Tabla 9: Herramientas BI a utilizar para el DW	50
Tabla 10: Fuente de datos para ETL Dimensiones	52
Tabla 11: Fuentes de datos para ETL Medidas	53
Tabla 12: Definición de Reportes de Gestión de Clientes	55
Tabla 13: Actividades de Implementación de Bodega de datos	68
Tabla 14: Pasos de Ejecución	79
Tabla 15: Resultados de Pruebas Unitarias.....	89
Tabla 16: Resultados de Pruebas con Usuarios	91
Tabla 17: Matriz de resultados Obtenidos (Problemas del Giro del Negocio)	94
Tabla 18: Matriz de resultados Obtenidos (Problemas Tecnológicos).....	96
Tabla 19: Matriz de Resultados Obtenidos (Nuevos Problemas Encontrados)	97
Tabla 20: Alcances Propuestos y tiempos optimizados	99
Tabla 21: Criterios para la Toma de Decisiones	100

INTRODUCCIÓN

Debido a los acelerados cambios que la empresa experimenta, y a la competencia que existe en el sector financiero que cada día es mayor, la organización tiene la necesidad constante de brindar servicios acordes a la exigencia del cliente, y que a su vez ayuden al incremento y fidelización de los mismos.

Todos los movimientos operacionales de la empresa, generan datos que terminan siendo el producto secundario de todas las transacciones que se realizan. Estos se guardan en la base de datos (BD) del sistema, dichos datos son el soporte de los procesos del negocio, pero actualmente no generan valor agregado, debido a que no existen reportes que ayuden al análisis de la información que permita elaborar las estrategias en tiempo razonables.

Considerando la evolución tecnológica y la tendencia de obtener información de forma oportuna para la toma de decisiones, surge la necesidad de

implementar un procedimiento que le permita realizar análisis de información, explorando los datos de la empresa para lograr estrategias en el momento preciso, además de permitir monitorear de forma efectiva sus operaciones con el objeto de conseguir una ventaja competitiva al momento de captar nuevos clientes y retenerlos.

El área de sistemas es quien ha venido solventando dicha necesidad de información, lo cual ha provocado una carga operativa para el departamento, además de que los tiempos de respuesta no son óptimos para su respectivo análisis, y en ocasiones la información suele ser errónea por la falta de conocimiento del negocio del personal involucrado en el desarrollo del requerimiento.

El presente proyecto tiene como solución, la utilización de una herramienta de inteligencia de negocios que a través del proceso de extracción, transformación y carga (ETL), permiten crear un Datamart (DM) que formará parte del Datawarehouse (DW) o almacén de datos que proveerá de información necesaria para el análisis. Originando así la oportunidad de tener

los mecanismos para reaccionar rápidamente a la necesidad de los clientes, mediante la transformación de la data transaccional en información analítica que a su vez genera conocimiento para la empresa.

CAPÍTULO 1

GENERALIDADES

1.1 Antecedentes

La empresa administradora de servicios financieros con más de 10 años en el mercado a nivel nacional, cuenta con un sistema de prestaciones sociales y financieras. Su principal objetivo es ofrecer un excelente servicio para mejorar la satisfacción de sus clientes, optimizando procesos y renovando las estrategias de captación de clientes y retención. Sin embargo, no posee una herramienta que facilite la toma de estas decisiones, por lo que se apoyan con el departamento de sistema, solicitando información de segmentación de clientes.

El departamento de Gestión de Clientes realiza sus informes a través de archivos de Excel solicitados al área de Sistemas o matrices en hojas de cálculo conectadas directamente a la base de datos transaccional, por lo que al momento

de actualizarlas en el día puede tomarse varias horas, ya que utilizan las tablas transaccionales y este proceso se realiza todos los días, adicional estos archivos no son seguros ya que contienen las claves de producción.

Los reportes solicitados son basados en las necesidades del área, de esta manera realizan los análisis necesarios para la captación o retención de clientes, definiendo nuevas estrategias, identificando lugares donde deben visitar y en qué sector existe menos clientes, donde y porque podríamos tener mejores resultados.

La mejor manera de moldear los reportes solicitados por el área de gestión de clientes y de mostrar la información global para que puedan observar las diferentes tendencias, serían los cubos de información, que muestra la información desde otras perspectivas, según lo requiera el usuario para poder obtener datos útiles que le ayuden a la toma de decisiones.

1.2 Descripción del Problema

En la actualidad la empresa no cuenta con una herramienta de inteligencia de negocios que provea información a tiempo sobre la segmentación de clientes, permitiendo así poder contar con los mecanismos necesarios para reaccionar rápidamente ante los cambios del mercado, agregando valor a la empresa.

Específicamente el área responsable de la gestión del cliente percibe, que los clientes han dejado de confiar en la institución, debido a la caída de

aportaciones que ha ido creciendo, y el nivel de cobros se ha visto afectado gravemente, este inconveniente ocasiona que surja la necesidad del estudio de segmentación de clientes para verificar en qué segmento es importante captar nuevos clientes o realizar estrategias de retención.

En el caso de necesitar la información de segmentación de clientes, se necesita realizar un requerimiento al área de Sistemas, para que asignen personal que ejecute dicha solicitud, provocando retrasos en los cronogramas que se encuentran en ejecución. Esto provoca que el personal debe de abandonar un proyecto vigente para proceder con lo solicitado, generando costos administrativos adicionales por mala utilización de recursos y posterior contratación de personal para terminar a tiempo los proyectos paralizados, además puede ocurrir que no se atienda el requerimiento de forma inmediata por falta de personal y se entregue el informe atrasado, por lo que el tiempo de espera puede ser muy largo.

Otro de los factores que afecta dicho requerimiento es el escaso conocimiento del recurso asignado, debido al alto nivel de rotación del personal, lo que a su vez evidencia una excesiva exposición a errores debido a la manipulación de la información por personal sin experiencia en el negocio.

En el campo tecnológico, se evidencia inconsistencias debido a que no existe integridad de data, algunos campos se grabaron vacíos, adicional no existe seguridad de información ya que se puede filtrar información valiosa y confidencial a personal no adecuado.

1.3 Solución Propuesta

Tener una herramienta de inteligencia de negocios que permita convertir la información de la base de datos en conocimiento, ayudando a mejorar la captación y retención de clientes para fortalecer el crecimiento de la empresa a través de una mejor gestión de los servicios brindados por la institución, permitiendo que la Alta Dirección tenga información pertinente y oportuna para la toma de decisiones mediante el cubo de información de segmentación de clientes, que ayude al control de los objetivos del negocio.

La solución propuesta apoya a la jefatura de Servicio al Cliente, mediante el desarrollo e implementación de cubos de información para presentar informes, lo que a su vez permite agilizar su elaboración y asegurar su validez.

Dicha implementación constará de procesos fuera de línea que se ejecutaran automáticamente en horarios no laborables, permitiendo contar con información en tiempo real del día anterior.

Para lograr dicha meta, primero extrae información de los orígenes de la base de datos SQL Server transaccional. Luego la transforma mediante el proceso de SQL Server Integration Services de extracción, transformación y carga a un Datamart para optimizar las consultas. Y finalmente la cargan al origen de datos analíticos donde se tiene la información resumida y agrupada. El origen de datos analíticos se escala a la plataforma analítica mediante el proceso de SQL Server Analysis Services, que es donde se vinculan las herramientas de consumo de información. Dicha plataforma

analítica permitirá que se realicen las consultas de forma instantánea, generando el dinamismo y rapidez que el negocio necesita. Se utilizará para la conexión con el Cubo de Información la herramienta de Office Excel, que es la más conocida y dominada por el usuario de la empresa.

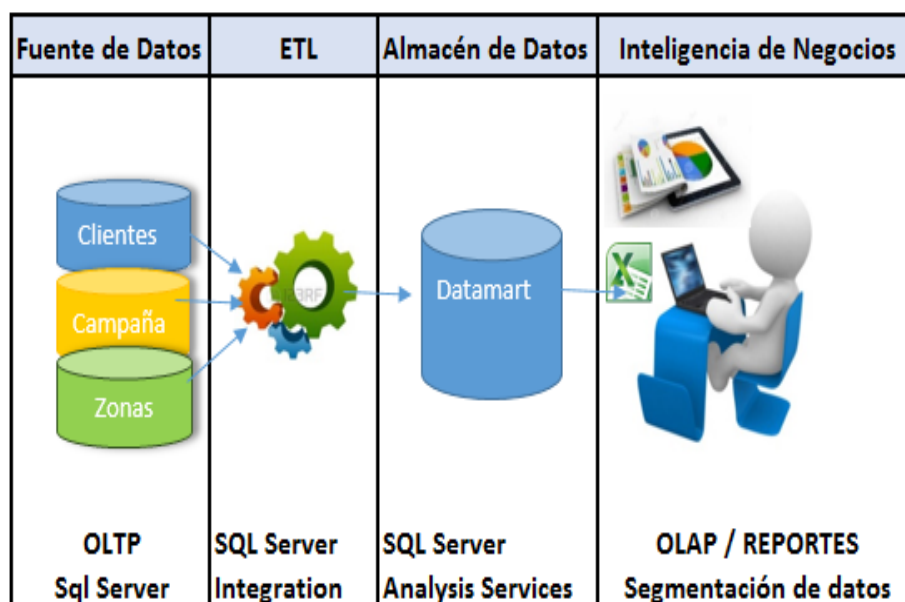


Figura 2.1: Diagrama de Business Intelligence

Se proporcionará un modelo analítico con todas las variables necesarias para establecer la captación y retención mediante la segmentación de clientes/potenciales para crear producto o estrategias acordes a las diferentes realidades y necesidades.

Las diferentes medidas son:

- Número de Clientes: Cantidad total de clientes que posee la empresa y con los que se puede trabajar en las diferentes estrategias.
- Número de Clientes por ubicación geográfica: Cantidad total de clientes según el lugar. Se puede obtener las cantidades de clientes por provincias, cantones y parroquias, para poder determinar las estrategias adecuadas según el lugar.
- Número de Clientes según datos generales, laborales y empresariales. Determinar la cantidad clientes que existe según el género, edad, estado civil, nivel de ingreso, entre otros, para poder personalizar las estrategias según el segmento.
- Número de clientes según fecha de ingreso: Se puede verificar el crecimiento de clientes en meses, o años, y la antigüedad de los mismos.
- Número de clientes activos: Cantidad de clientes activos sobre el total de clientes registrados se encuentran con servicios actualmente, los clientes activos son aquellos que mantienen movimientos en los últimos 6 meses en nuestra empresa.
- Número de clientes activos por ubicación geográfica: Sirve para verificar el nivel de crecimiento de nuevos clientes en cada provincia.
- Número de clientes activos por datos generales, laborales y empresariales. Tasa de crecimiento de clientes dependiendo del nivel de ingreso, carga laboral, genero entre otros.
- Número de clientes activos por fecha: Tasa de crecimiento de cliente en meses específicos. En qué momento se logró una mayor captación de clientes.

- Número de clientes que actualizaron datos: Este indicador nos da resultados del porcentaje de datos actualizados, según las diferentes segmentaciones de clientes (datos generales, laborales, empresariales y fechas).
- Número de clientes que han recibido premios o regalos: Indica la cantidad de clientes que recibieron obsequios por actualización de datos, y en qué sectores hemos trabajado y en cuales nos falta trabajar para incrementar la cantidad de cliente de la empresa.
- Número de clientes que han firmado cartas de descuento bancario para el pago de nuestro servicio: Este indicador sirve para verificar el porcentaje de clientes activos que han firmado carta de descuento y que se han verificado, según la ubicación geográfica, datos generales o antigüedad.

Beneficios de la Solución

La aplicación de la solución propuesta nos brinda los beneficios mencionados a continuación:

- Los procesos de toma de decisión serán mejorados debido a la disponibilidad de información a tiempo.
- Descubrir oportunidades de negocio mediante la medición de comportamiento de los clientes que permitan generar nuevos servicios y mejorar estrategias según la necesidad del segmento.
- Las decisiones se harán de manera más rápida basada en información precisa, garantizando aciertos al mitigar la incertidumbre.

- Los procesos de la empresa serán optimizados ya que se elimina el tiempo que se perdía esperando por información que podía ser incorrecta e incompleta.
- Confiabilidad debido a la gestión de seguridad de información confidencial disponible para el personal adecuado.
- Mejor utilización de recurso humano, debido a que se elimina el uso de personal para entrega de dichos informes, y por lo tanto, se ocupan en atender las tareas prioritarias.
- Eliminación de datos inconsistentes y repetidos en los informes.

A continuación, se detalla el alcance del proyecto:

- Implementación de un Datamart del área de gestión de clientes que permitirá cubrir las necesidades del departamento, definición de generación de informes bases que se presentaran en Excel y permitirán analizar e interpretar la información.
- Se analizará y diseñará las nuevas dimensiones y medidas para la segmentación de clientes.
- Creación de estructuras del Datamart
- Creación de paquetes de Extracción, transformación y carga.
- Se creará el cubo de información que permitirá tener los datos disponibles en todo momento, eliminando la generación de los siguientes reportes manuales:
 - Reportes de total de clientes por provincias, cantones y parroquias.

- Reportes de clientes por género, edad y estado civil.
 - Reportes de clientes que han recibido premios o regalos en campañas por provincias.
 - Reportes de clientes que han firmado cartas de descuento bancarios por provincias.
 - Reportes de clientes que han actualizado datos por provincias.
- Se establecerá los roles y privilegios para los diferentes usuarios.
 - Se analizará los beneficios de la solución

1.4 Objetivo General

Implementar un cubo de información para la segmentación de clientes de una empresa de prestaciones de servicios financieros, luego de su respectivo desarrollo.

1.5 Objetivos Específicos

- Levantar información para el desarrollo e implementación del cubo de información para la segmentación de clientes.
- Conocer los elementos más importantes para diseñar el cubo de información según las necesidades del negocio.
- Implementar el cubo de información para la segmentación de clientes.
- Analizar los resultados de la solución diseñada, contrastando los beneficios obtenidos.

1.6 Metodología

Para implementar el proyecto de Inteligencia de Negocios se va a considerar la metodología de Kimball, primero se procederá a llenar el área temporal con los datos necesarios, para luego mediante procesos de extracción, transformación y carga se proceda a crear el Datamart para la segmentación de clientes que servirá para presentar los datos mediante el cubo de información, que ayudará a la toma de decisiones de manera oportuna.

CAPÍTULO 2

MARCO TEÓRICO

2.1 Conceptos básicos de Inteligencia de Negocios

Hoy en día es evidente la necesidad de tener mejores métodos para extraer y transformar los datos en información válida que genere valor a la empresa. La Inteligencia de Negocio o Business Intelligence es una respuesta a dicha necesidad, y a su vez es una evolución de los sistemas de soporte de decisiones. La inteligencia de Negocios es un conjunto de metodologías, aplicaciones y prácticas orientadas a la organización y administración de información que ayuda a tomar las mejores decisiones a los usuarios de una empresa [1]. Entre las tecnologías de BI tenemos: Reportes, procesamiento analítico en línea (OLAP), análisis visual, análisis predictivo, reglas de negocio e integración de datos.

La implementación de BI en una organización permite tener soporte de las decisiones que se toman, tanto para el usuario interno como externo, ayudando a tener gestión del personal y ventaja sobre sus competidores. Pero no siempre se puede lograr todos los beneficios de BI, ya que se puede cometer errores en la definición del requerimiento de la empresa por no saber determinar la magnitud del problema a solucionar, provocando en ocasiones el fracaso de los proyectos.[2]

Podemos definir a BI como el que nos provee información de control del proceso de negocio y disponer de información orientada a la semántica del usuario facilitando de esta manera la toma de decisiones.[3]

2.1.1. Conceptos de Componentes de BI

Los componentes BI están compuestos por las siguientes capas:

- Fuentes de datos: Son las diversas fuentes donde guardan la información del negocio como: Sistemas OLTP, archivos de Excel, archivos planos, entre otros.
- Datawarehouse y/o Datamart: Son bases de datos multidimensionales que se alimentan mediante procesos de extracción, transformación y carga.
- Sistemas OLAP y de minería de datos: Es una nueva fuente de información procesada que potencia los sistemas de BI.

- Herramientas analíticas y de presentación: Permiten una mejor visualización de los datos para el usuario, accediendo así a la información del negocio para realizar el respectivo análisis.[4]

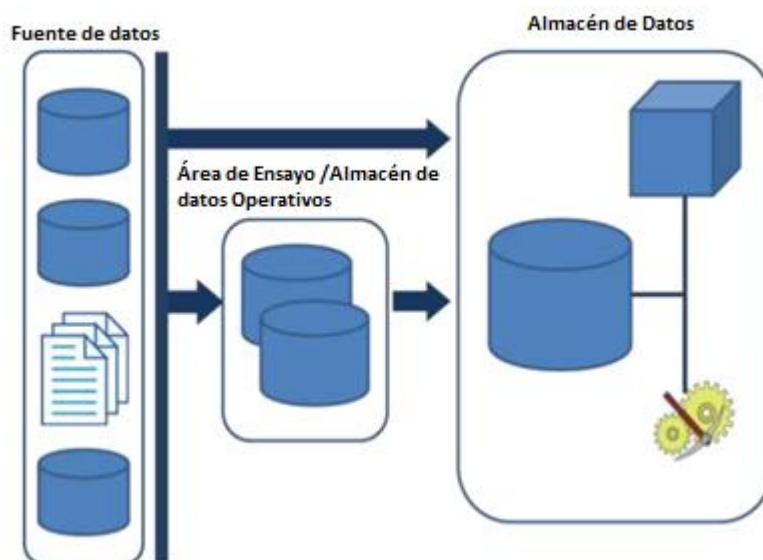


Figura 4.1: Solución BI [4]

2.1.2. Datawarehouse

Es una base de datos que está compuesta de diferentes fuentes con información fiable, depurada y homogénea, que permite realizar un análisis de diversas perspectivas, mejorando la gestión, con tiempos óptimos de respuesta, apoyo y control de la toma de decisiones.

Los beneficios de un Datawarehouse son:

Reducción de coste: Al tener una única base de datos para consultar simplifica la tarea de los usuarios, así como la del personal del departamento de sistemas, provocando beneficios tangibles e intangibles

como: reducción del tiempo invertido para obtener la información, permite ver si existen ventajas competitivas y de esta manera poder reaccionar a tiempo ante una variable crítica, entre otros.

Mejorar el modelo de comportamiento de la empresa: Acceso fácil y flexible a la información de las diferentes áreas de la empresa, obteniendo información cercana a la realidad.

Promover procesos de cambio: Análisis con frecuencia, menos limitado a los directivos, el cual permite obtener conocimiento de los procesos de la empresa, proporcionando la detección de disfunciones y aplicando medidas para el cambio.

Para Bill Inmon un Datawarehouse se caracteriza por ser:

Temático: La organización de datos por temas facilita el entendimiento de los usuarios, por ejemplo, todos los datos de un cliente se agrupan en una sola tabla, los de productos en otra tabla y así sucesivamente.

Integrado: Los datos están integrados de manera que se eliminan las inconsistencias existentes entre los diferentes sistemas operacionales, entre las inconsistencias más comunes tenemos: formatos de fechas, unidades de medidas, nomenclaturas, varias tablas con la misma información.

Histórico: Son datos que cambian en el tiempo, por lo que se quedan reflejados para posteriormente ser visualizados cuando son consultados,

impidiendo la alteración de la realidad que existía al momento de estar almacenados.

No volátil: La información debe ser solo de lectura en una Datawarehouse, una vez ingresada no se debería modificar, ni eliminar.[5]

A continuación se muestran las principales diferencias entre una base de datos operativa y una base de datos Datawarehouse:

ASPECTO	Base de datos OPERATIVA	Base de datos del DATA WAREHOUSE
Objetivo	Operaciones del día a día	Toma de decisiones
Granularidad	Datos detallados	Datos depurados y sumariados
Orientación	Por aplicación o transacciones	Temáticos, categorías amplias
Nivel de integración	Variable, no crítica para OLTP	Racionalización de datos
Temporalidad	30-40 días (1 año máximo)	1-5 años on-line (como norma)
Validez de datos	Válidos para "ahora"	Perspectiva histórica
Usuarios	Administrativos/transacciones	Gestores/análisis
Volatilidad	Actualizable	Sólo accesibles una vez cargados
Procesamiento	Repetitivo y conocido	Puntual y no conocido
Manejo	Un registro cada vez/pocos datos	Manejo masivo de datos
Disponibilidad	Alta	Media
Normalización	Cuarta forma normal	Redundancia
Tecnología	Transaccional	Acceso y gestión

Figura 4.2: Base de datos operativas vs Base de datos Datawarehouse [5]

2.1.3. Bodega de datos y Datamart

Un Datamart es una base que contiene datos departamentales, una versión exclusiva del almacén de datos (Datawarehouse), es un subconjunto de datos de un área del negocio en concreto. Es un sistema de consultas multidimensionales de la información, que se ejecutan a la base de datos a través de herramientas OLAP.

Para Lane (1999), un Datamart es un Datawarehouse de forma sencilla que está enfocado en una área funcional. Debido a que se centra en un área, se constituyen de pocas fuentes de datos que los Datawarehouse.

Se lo define como un conjunto de datos, cuyo objetivo es responder a un análisis o necesidad, los datos están estructurado como copo de nieve o un modelo estrella al igual que un Datawarehouse.[6]

Por su clasificación los Datamart pueden ser dependiente e independiente.

Datamarts dependientes: es un subconjunto o base de datos que se crea en el Almacén de datos.

Datamarts independiente: es un Almacén de datos pequeño para un departamento o negocio estratégico.

Un datamart dependiente utiliza un modelo de datos consistente con calidad, el inconveniente en este modelo es que se debe construir primero el Almacén de datos y luego el datamart. La ventaja es la visualización de los mismos datos a los que acceden todos los usuarios de Almacén de datos.

Los datamart independiente son construidos más rápido y a menor costo que un Almacén de datos.[7]

2.1.4. Procesos de extracción, transformación y carga de datos

Los datos después de ser extraído de diferentes fuentes como archivos planos, fuentes de registro web, sistemas heredados, hojas de cálculo, conjuntos de datos XML, tablas relacionales y fuentes no relacionales, luego se propagan al área de almacenamientos de datos DSA donde se transforman y se limpian antes de cargarse en el depósito de datos.[8]

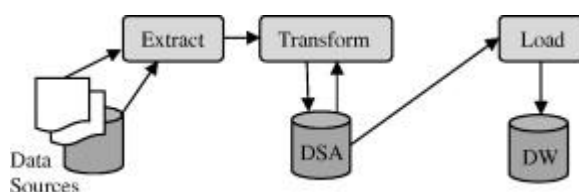


Figura 4.3: Marco general para los procesos de ETL[8]

Es un proceso que organiza y aporta los métodos y herramientas necesarias para mover el flujo de datos de las diferentes fuentes de una organización a un almacén de datos.

Extracción: Extrae los datos desde los sistemas origen y los convierte a un formato preparado para continuar con el proceso de transformación.

Transformación: Utiliza funciones o criterios de negocio de los datos que fueron extraídos para cambiarlos en datos que luego van hacer cargados.

Carga: Proceso de carga de acumulación simple, es aquella que realiza el resumen de todo el periodo de tiempo que existe en una transacción. El proceso Rolling mantiene diferentes niveles de granularidad. Interactúa

con la base destino, garantizando la calidad de datos mediante restricciones y disparadores definidos.

Entre los procesos de ETL que puede realizarse tenemos: migración de datos de diferentes aplicativos o cambios de versión, consolidar datos de sistemas con gran volumen de información y sincronizar base de datos operativos.[9]

2.1.5. Cubos de Información

Cubos OLAP (On-Line Analytical Processing) que significa procesamiento analítico en línea, o cubos de información, se parecen a los cubos de rompecabezas en los juego, en el que se mueve cada celda para armar los colores, en el datawarehouse el objetivo es organizar los datos por tablas o relaciones, estos cubos tienen un sin número de dimensiones a diferencia del juego que solo tiene 3 dimensiones.[10]

Son estructuras multidimensionales que tienen datos resumidos de base de datos transaccionales y diferentes fuentes, es una solución muy utilizada en el campo de Inteligencia de Negocios para acelerar las consultas de un gran volumen de datos para marketing, informes de negocios de ventas, informes gerenciales, minería de datos y áreas similares.

Las base de datos relacionales son muy complejas y relativamente lentas comparadas con la base multidimensional que son muy veloces en sus respuestas, es un mejor modelo de búsqueda para la parte administrativa.

Un cubo tendrá datos de las variables que desea analizar, permitiendo tener una vista lógica de los datos provistos por los sistemas transaccionales hacia el almacén de datos. Esta vista estará disponible dependiendo de las dimensiones y también podrá tener datos calculados. El análisis de los datos se realizará según las dimensiones del cubo, por lo que tendremos un análisis multidimensional.

Una de las opciones para consultar el cubo de información es a través de tablas dinámicas en una hoja de cálculo o programas personalizados, que les permitirá manipular las diferentes perspectivas del cubo con mucha facilidad y rapidez, realizando los filtros o cruce de información que se requiera. [11]

2.2 Esquema de Almacén de Datos

Un almacén de datos, es una colección de dimensiones y variables de diferentes fuentes, en donde las dimensiones son elementos que forman parte del análisis y variables son los valores que se analizan. Así tenemos que un DW procesa datos que se convertirán en conocimiento para la toma de decisiones del usuario.[11]

Dimensiones: Representa una entidad del negocio, que ayuda a colocar en contexto a los datos numéricos de la tabla de hechos, es clave para hacer que el

almacén de datos sea utilizable y comprensible. Satisface la necesidad analítica del usuario, mostrando información descriptiva, fácil de utilizar para los usuarios. Las tablas de dimensiones son los puntos de entrada en la tabla de hechos. [12]

Variables: Representa datos numéricos en el contexto de las dimensiones, en el esquema estrella es la tabla central. Constituye los valores que van hacer analizados, como por ejemplo unidades vendidas o el número de empleados.[13]

Esquema Copo de Nieve

Este esquema es más complejo que el esquema estrella, la principal diferencia es que existen dimensiones que no están relacionadas directamente con la tabla de hecho, sino que se relacionan por medio de otras tablas dimensiones.

Este esquema está formada por una tabla de hecho central que tiene las medidas y una o varias tablas adicionales con un nivel mayor de normalización.[4]

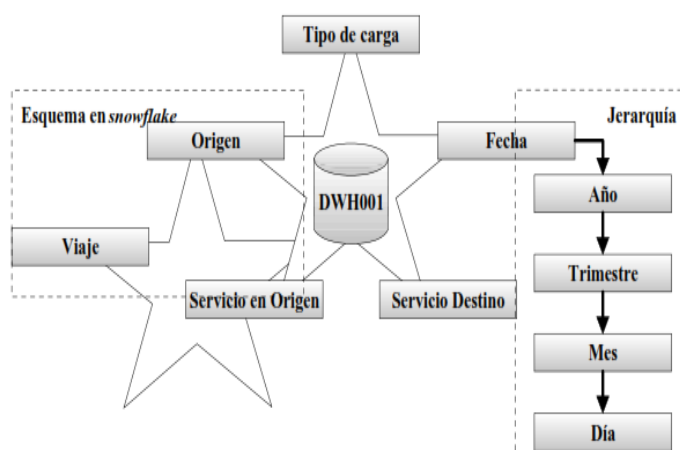


Figura 4.4: Diagrama de esquema Snowflake[4]

Esquema en Estrella

El esquema estrella se caracteriza por formarse de una tabla central denominada tabla de hecho que contiene las métricas y medidas del negocio y está relacionada con varias tablas dimensiones que describen la información existente en un proceso específico del negocio, creando de esta manera una estrella, motivo por el cual se le da el nombre a este esquema que facilita el acceso a datos y el análisis, la tabla de hecho es la única que tiene múltiples uniones (joins) que la conectan con las otras tablas.

Las puntas de la estrella son las tablas de dimensiones, las cuales proporcionan contexto a los datos numéricos.[13]

A diferencia del esquema de copo de nieve este modelo ocupa más espacio en disco, pero resulta más fácil de entender por el usuario y tiene un mejor rendimiento a la hora de ser consultado los datos.[4]

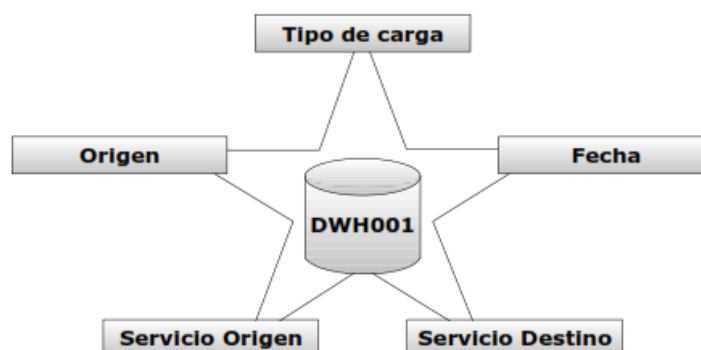


Figura 4.5: Diagrama de un esquema estrella[4]

2.3 Conceptos Básicos de la Metodología a desarrollar

Para el desarrollo de un almacén de datos tenemos las siguientes metodologías a implementar como: Bottom up y Top down.

Metodología Bottom-up

Metodología de Ralph Kimball también denominada Bottom-up (ascendente), pues el Datawarehouse Corporativo, no es otra cosa que la unión de los diferentes Datamarts que están estructurado a través de una arquitectura de bus. Dicha características hacen que el proceso sea más flexible y sencillo de implementar, pues se puede iniciar construyendo un primer elemento como Datamart del sistema de análisis, y después ir agregando otros elementos que compartan las dimensiones definidas anteriormente o incluir nuevas. En esta metodología los procesos de extracción, transformación y carga extraen la información del sistema operacional para ser procesados en el “area stage”, para luego realizar el llenado de los Datamart de manera individual, respetando la estandarización de las dimensiones (dimensiones conformadas).

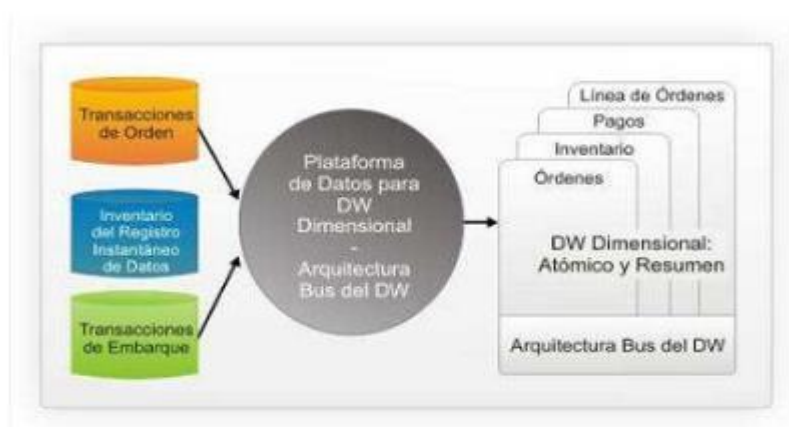


Figura 4.6: Metodología Kimball - Arquitectura Bus del DW[14]

Metodología Top-down

La metodología Top-down “descendente” también es llamado metodología Inmon, donde los datos son extraídos de los diferentes sistemas operacionales por medio de los procesos ETL y cargados al área stage, donde se consolidarán y validarán en el DW, además tenemos los metadatos donde se documentan de forma clara y precisa el contenido de DW. Después de realizar dicho proceso los Datamart departamentales son refrescados para obtener información de este.

Tener un enfoque global, es más difícil realizar un proyecto sencillo ya que se intenta abordar el “todo” para poder ir al “detalle”. [14]

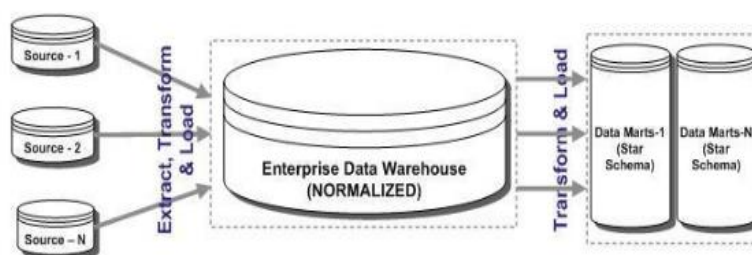


Figura 4.7: Metodología Inmon DW Corporativo[14]

2.4 Segmentación de Clientes

La segmentación de clientes se orienta a clasificar a los clientes ya existentes de un negocio, es una manera analítica de agrupar a los segmentos de acuerdo a características y comportamientos similares. La segmentación tiene un efecto multiplicador ya que el impacto se traduce en incremento de las ventas, mayor satisfacción del cliente y mayor rentabilidad. Así mismo si se habla de estrategias de ventas, es importante determinar las características de sus clientes, para

entender el costo de servir en cada uno de los canales de distribución, permitiendo precisar estrategias claras y rentables para la empresa.

La segmentación ayuda a las empresas a organizar y ajustar servicios y recursos según los grupos o clientes concretos, en otras palabras recibe el servicio de acuerdo al segmento.

Ventajas de una segmentación de clientes

- Orientación al cliente.
- Mejor utilización en el uso de recursos, que se desembocan en la eficiencia de los procesos operativos de la empresa.
- Atención ajustada a las necesidades del cliente.
- Entender mejor las necesidades y características del cliente.
- Maximiza las ventas al tener focalización gracias a una mejor comunicación de mercado.
- Identifica la ubicación de los clientes potenciales en los distintos segmentos.
- Mejor rentabilidad y mayor satisfacción de los clientes.

Segmentación de Mercado

La segmentación de mercado trata de dividir el total de mercado en pequeños grupos o segmentos que sean homogéneos. El objetivo es segmentar apropiadamente el mercado para conocer al cliente y mejorar la exactitud de marketing en la empresa.

El segmento de mercado es un conjunto homogéneo y grande de clientes que se pueden parecer dentro del mercado, por actitudes de compra, hábitos de compra

similares, por ubicación geográfica, y tienen reacciones similares ante la mezcla de marketing.

Tipos de segmentación de mercado

Segmentación Geográfica: Es el fraccionamiento de mercado en base a su provincia, clima, ubicación como país, urbana o rural, etc.

Segmentación Demográfica: Está vinculada a la demanda y es posible de medir. Entre las particularidades demográficas tenemos: el género, el ingreso, estado civil, etnia, la edad, etc.

Segmentación Psicográfica: Inspecciona todo lo relacionado con la conducta, sentimientos de las personas y sus pensamientos, así como las dimensiones de personalidad, de estilo de vida y de valores.

Segmentación por Comportamiento: Detalla la conducta que tiene correspondencia con el producto, maneja las variables de beneficios que el cliente quiere de un producto y la tasa del producto consumido.

Segmentación Socioeconómica: Tiene que ver con los estudios, personalidad, profesión, clase social, valores que defiende, etc.[15]

CAPÍTULO 3

LEVANTAMIENTO Y ANÁLISIS DE LA INFORMACIÓN

3.1 Recopilación de datos de la situación actual

La empresa cuenta con varias áreas de negocio, enfocados en el sector de servicios financieros, a continuación se presentara un macro diagrama de las áreas involucradas en el proyecto.



Figura 6.1: Estructura Organizacional

En la Figura 3.1, se ilustra el organigrama del área responsable de la gestión del cliente de la empresa administradora de servicios financieros, comprendida por el Departamento de Gestión de clientes, el cual reporta al Director Administrativo. El jefe del departamento de Gestión, tiene a su cargo dos supervisores que se encargan de la gestión con los clientes. De acuerdo con el alcance propuesto, se pretende cubrir las necesidades de información de dicha área.

El Departamento de Sistemas es el encargado de cumplir con la entrega de información que solicite el Departamento de Gestión de Clientes.

Entre los principales inconvenientes, que actualmente se posee con respecto a la entrega de información para la toma de decisiones, tenemos:

- La información es entregada en un archivo de Excel para que pueda ser consultada por el usuario, la entrega se realiza a través de una copia local del archivo al equipo del usuario que lo solicita.
- Los archivos contienen toda la información solicitada de las bases de datos transaccionales de la empresa.
- El tema de seguridad de estos archivos es muy bajo y genera carga de trabajo cada vez que solicitan información actualizada.
- Generación de informes gerenciales nuevos que el sistema transaccional no posee, por lo que se genera carga manual inesperada al Departamento de Sistemas, que dependiendo del tipo de informe puede demorar hasta 3 días.

- Escaso conocimiento por parte de las personas que laboran en la empresa por el alto índice de rotación del personal, por lo que los reportes pueden contener errores al momento de ser entregados.

Adicional a los problemas anteriores se suman los siguientes:

- No se obtiene información 100% confiable, debido a la falta de integridad de información, debido a datos vacíos o nulos en campos importantes del cliente, como género, estado civil, entre otros datos que no fueron validados al momento de su ingreso.
- Se obtiene información duplicada, lo cual no permite tener información correcta apegada a la realidad.
- El área de sistemas tiene exceso de actividades, sin contar con los incidentes que surgen en el día a día.
- Las solicitudes de nuevos informes por el área de gestión de clientes deben de entregarse en muy poco tiempo, esto causa que las actividades planificadas en el Departamento de Sistemas se acumulen y afecten al cronograma de proyectos que están en ejecución.

Otro problema muy común lo encontramos en el detalle de la solicitud de los requerimientos, las cuales no están bien definidas, provocando reprocesos en las tareas, lo que provoca que el personal utilice más tiempo de lo necesario para la finalización del requerimiento. Adicional de que no se priorizan las tareas correctamente, y consecuentemente se producen encolamientos de estos informes que son urgentes para la operación de la empresa.

El patrocinador de este proyecto sería el área de gestión al cliente ya que son los beneficiarios directo con la implementación del cubo de información para la segmentación de clientes.

A continuación detallaremos los diferentes roles y funciones involucrados en este proceso:

Tabla 1: Roles involucrados en el proceso

Involucrados	Departamento de Sistemas	Departamento de Gestión de Clientes	Directivos Administrativos
Roles	Analizar	Informar	Controlar y Decidir
Funciones	Análisis de Información	Presentar Información	Control Institucional
	Investigación de Problemas	Responsables de Informes	Establecer Estrategia
	Establecer Propuestas		Toma de Decisiones

Tabla 2: Análisis FODA de Gestión de Clientes

Fortalezas	Oportunidades
Captación de clientes.	Ampliación de cartera de créditos a través de la tecnología.
Oportunidad de entrega de créditos a clientes.	Incremento de las aportaciones que realizan los clientes.
Líder de mercado.	Nuevos Convenios por débitos bancarios.
Excelente atención al cliente.	
Personal capacitado para trabajar.	

Genera flujo de trabajo para otras áreas.	
Debilidades	Amenazas
Falta de promoción de los servicios que brinda la empresa.	Desenfoco de la visión principal del área.
Saturación de mercado.	Cambios del sistema de crédito o regularizaciones estatales.
Exceso de trabajo.	Nuevos competidores en el mercado financiero.
Clientes con mayor exigencia.	Fuga de recursos especializados por mejores ofertas laborales.
Centralización de conocimientos en recursos.	

Con el levantamiento de información realizado y el análisis FODA, se demuestra que la empresa tiene fortalezas y grandes oportunidades que minimizan las debilidades y amenazas, sin embargo existen puntos críticos que se deben trabajar como la Generación de flujo de trabajo para otras áreas y Personal capacitado para trabajar, en los cuales se centran los problemas detallados a continuación:

- Se cuenta con solo dos especialistas que se encargan de solventar todos los requerimientos del área, en otras palabras no todo el personal del área esta especializado para ejecutar los requerimientos existentes, por lo que no existe tiempo ni una persona encargada de verificar la data existente, eliminando así la duplicidad y velar por la integridad de datos.
- El área de desarrollo, recibe muchas solicitudes de requerimientos de reportes que no pueden obtener mediante el sistema, lo cual hace que se desenfocan de su visión principal, provocando sobrecarga de trabajo, y atrasos en la entrega de los diferentes proyectos.

La solución que se plantea se enfoca en mejorar los puntos antes mencionados, eliminando problemas tecnológicos del negocio, eliminando duplicidad de datos e integridad de data, entre otros, contribuyendo a la entrega de información útil para la toma de decisiones, beneficiando al negocio y disminuyendo la carga operativa con los procesos automatizados. A continuación esquema de solución planteada:

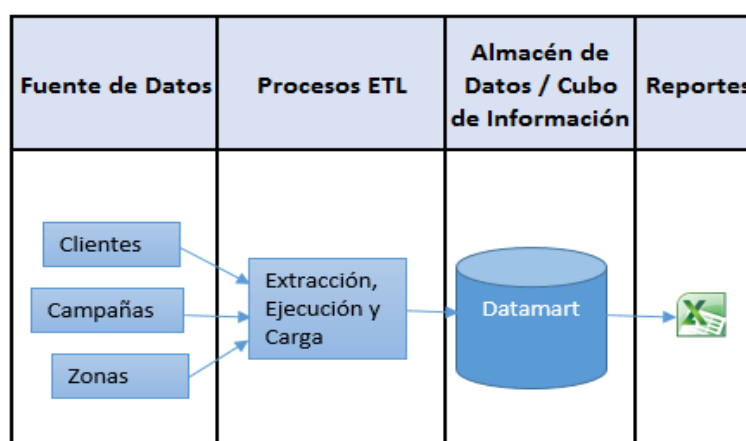


Figura 6.2: Esquema de la Solución Planteada

3.2 Documentación del Requerimiento

El área de Gestión de cliente requiere que se resuelva el principal inconveniente de obtener información necesaria para su área de manera rápida y dinámica, como apoyo a sus procesos para presentar la información para la toma de decisiones, los mismo que en la actualidad se presentan muy pocas veces y de manera estática, lo que conlleva a realizar tareas adicionales como ajustes manuales a dichos reportes para obtener información de valor.

De esta manera, se evidencia la necesidad de proveer una herramienta que brinde informes de forma rápida y que permitan tener un proceso dinámico para el análisis de información, de tal manera que permita obtener una visión general del área de negocio.

Con respecto a la fuente de información que se necesita para satisfacer la necesidad de negocio del área de gestión de clientes, el área de Sistema determinó que la información se encuentra en las base de datos del ERP, dentro de su estructura de BD, pero no cuenta con una herramienta que tenga el efecto dinámico para sus reportes, optimizando el tiempo de respuesta para poder analizar la información.

Dimensiones y medidas

Como parte del levantamiento de información con las áreas de Gestión de clientes, se establecieron las dimensiones y medidas e informes que se implementarán.

El área de Gestión de clientes ha definido diferentes vistas que ayudarán a reflejar los resultados esperados que brindarán el soporte requerido al usuario final, para el análisis de segmentación de clientes y toma de decisiones importantes. A continuación se detalla el análisis correspondiente:

- Número de clientes: Cantidad total de clientes que posee la empresa y con los que se puede trabajar en las diferentes campañas.

- Número de clientes por ubicación geográfica: Cantidad total de clientes según el lugar. Se puede obtener las cantidades de clientes por provincias, cantones y parroquias, para poder determinar las campañas adecuadas según el lugar.
- Número de clientes según datos generales, laborales y empresariales: Determinar la cantidad clientes que existe según el género, edad, estado civil, nivel de ingreso, entre otros, para poder personalizar las campañas según el segmento.
- Número de clientes según fecha de ingreso: Se puede verificar el crecimiento de clientes en meses, o años, y la antigüedad de los mismos.
- Número de clientes activos: Cantidad de clientes activos sobre el total de clientes registrados se encuentran con servicios actualmente, los clientes activos son aquellos que mantienen movimientos en los últimos 6 meses en nuestra empresa.
- Número de clientes activos por ubicación geográfica: Sirve para verificar el nivel de crecimiento de nuevos clientes en cada provincia.
- Número de clientes activos por datos generales, laborales y empresariales: Tasa de crecimiento de clientes dependiendo del nivel de ingreso, carga laboral, género entre otros.
- Número de clientes activos por fecha: Tasa de crecimiento de cliente en meses específicos. En qué momento se logró una mayor captación de clientes.
- Número de clientes que actualizaron datos: Este indicador nos da resultados del porcentaje de datos actualizado en la última campaña según las

diferentes segmentaciones de clientes (datos generales, laborales, empresariales y fechas).

- Número de clientes que han recibido premios o regalos en campañas: Indica la cantidad de clientes que recibieron obsequios en la última campaña de actualización de datos, y en qué sectores hemos trabajado y en cuales falta trabajar para incrementar la cantidad de cliente de la empresa.
- Número de clientes que han firmado cartas de descuento bancario para el pago del servicio: Este indicador sirve para verificar el porcentaje de clientes activos que han firmado la carta de descuento y que se han verificado, según la ubicación geográfica, datos generales o antigüedad.

La Tabla 3 detalla el listado de las dimensiones definidas para este proyecto.

Tabla 3: Listado de dimensiones

Nombre de la Dimensión	Nombre del Nivel	Descripción del Nivel (ejemplo)	Descripción de la Dimensión
Ubicación División Política Administrativa	Provincia	Guayas	Ayuda a segmentar los datos en jerárquica por provincias, cantones y parroquias.
	Cantón	Portoviejo	
	Parroquia	Ximena	
Ubicación nuevo modelo de gestión	Zona	Zona 1	Permite segmentar la información del cubo jerárquicamente por provincias, zonas y cantones.
	Distrito	03D01	
	Circuito	01D01C01	
Calendario	Año	2012	Permite segmentar la información jerárquicamente por año, mes y día.
	Mes	1	
	Día	15	
Relación Laboral	Tipo de Relación Laboral	Nombramiento (N)	Permite segmentar la información del cubo por relación laboral del cliente con su institución.

Nombre de la Dimensión	Nombre del Nivel	Descripción del Nivel (ejemplo)	Descripción de la Dimensión
Rango Nivel de Ingreso	Nivel de Ingreso	Menos de \$500	Permite segmentar la información del cubo por los rangos de nivel de ingresos del cliente.
Jornada Laboral	Jornada Laboral	Matutina	Permite segmentar la información del cubo por la jornada laboral de sus clientes.
Rango Años Servicios	Años de Servicios	Menos de 3 años	Permite segmentar el cubo por los rangos de años de servicio de sus clientes.
Estado Civil	Estado Civil	Soltero	Permite segmentar la información por el estado civil del cliente.
Rango de Edades	Rango de Edad del cliente	Menores a 30	Permite segmentar la información por rango de edad del cliente.
Genero	Sexo	Femenino	Permite segmentar la información por el género del cliente.
Tipo de Institución	Tipo de Institución	Colegio	Permite segmentar la información por el tipo de institución.
Tipo de Cargo	Tipo de Cargo	Administrativo	Permite segmentar la información por tipo de cargo del cliente en la institución.
Estado	Estado Cliente	Activos (A)	Permite segmentar la información según el estado del cliente en la base.

La Tabla 4 detalla las medidas que intervienen en el proyecto.

Tabla 4: Listado de medidas

Nombre de Medida	Descripción
Número de clientes	Representa la cantidad de clientes en la empresa.
Número de clientes activos	Representa la cantidad de clientes activos.
Número de clientes que han actualizado datos	Representa la cantidad de clientes activos que han actualizado los datos en las campañas.
Número de clientes que han recibido premios	Representa la cantidad de clientes activos que han recibido premios en campañas de actualización de datos.
Número de clientes que han firmado cartas de descuento	Representa la cantidad de clientes activos que han firmado la carta de descuento en las campañas.

La Tabla 5 muestra la intersección de las dimensiones y medidas del proyecto.

Tabla 5: Diseño de Cubo

	DPA	NMG	Datos Generales	Datos Laborales	Datos empresa	Fecha registro
MEDIDA / DIMENSIÓN	Provincias/Cantón/Parroquia	Zona/ Distrito	Rango de edades / Género / Estado Civil	Relación Laboral / Jornada Laboral/Años de Servicios / Rango de Nivel de Ingreso	Tipo de Cargo/ Tipo de Establecimiento / Establecimiento/Estado del Cliente	Año/Mes /Día
Número de clientes	x	x	x	x	x	x

	DPA	NMG	Datos Generales	Datos Laborales	Datos empresa	Fecha registro
Número de clientes con datos actualizados	x	x	x	x	x	x
Número de clientes activos	x	x	x	x	x	x
Número de clientes con premios o regalos	x	x	x	x	x	x
Número de clientes con carta descuento	x	x	x	x	x	x

La Tabla 6 contiene todas las fórmulas que tiene el proyecto.

Tabla 6: Diseño de Fórmulas

Campos Calculados	Formulas	Metas Anuales
Antigüedad en años de Fecha de Ingreso	$\text{trunc}(\text{yearfrac}([\text{fecha}], \text{date}(\text{year}(\text{now}())), \text{month}(\text{now}()), \text{day}(\text{now}())), 1))$	N/A
Porcentaje Parcial de clientes que firmaron carta sobre el total de clientes con cartas	$\frac{\text{Número Clientes con Firma}}{\text{sum}(\text{Número Clientes con Firma})} * 100$	Bueno >3% Aceptable 1.5% a 3% Malo < 1.5%
Porcentaje Parcial de clientes con regalo sobre el total de clientes con regalo	$\frac{\text{Número Clientes con Regalo}}{\text{sum}(\text{Número Clientes con Regalo})} * 100$	Bueno >3% Aceptable 1.5% a 3% Malo < 1.5%
Porcentaje Parcial de clientes actualizados sobre el total de clientes actualizados	$\frac{\text{Número Clientes Actualizado}}{\text{sum}(\text{Número Clientes Actualizado})} * 100$	Bueno >4% Aceptable 2% a 4% Malo < 2%

Campos Calculados	Formulas	Metas Anuales
Porcentaje Parcial de clientes activos sobre el total de clientes activos	$\frac{\text{Número Clientes Activo}}{\text{sum(Número Clientes Activo)}} * 100$	Bueno >5% Aceptable 2.5% a 5% Malo < 2.5%
Porcentaje Parcial de clientes sobre el total de clientes	$\frac{\text{Número de Clientes}}{\text{sum(Número de Clientes)}} * 100$	Bueno >4% Aceptable 2% a 4% Malo < 2%
Porcentaje de adquisición de clientes	$\frac{\text{N° clientes nuevos en un periodo}}{\text{N° clientes en el periodo}} * 100$	Bueno >4% Aceptable 2% a 4% Malo < 2%
Porcentaje de retención de clientes	$\frac{\text{N° clientes activos en un periodo}}{\text{N° clientes en el periodo}} * 100$	Bueno >2% Aceptable 1% a 2% Malo < 1%
Porcentaje de clientes que actualizaron datos	$\frac{\text{N° clientes actualizados en periodo}}{\text{N° clientes en el periodo}} * 100$	Bueno >3% Aceptable 1.5% a 3% Malo < 1.5%
Porcentaje de clientes que han recibido premios o regalos en campañas	$\frac{\text{N° clientes con regalo en un periodo}}{\text{N° clientes en el periodo}} * 100$	Bueno > 1.5% Aceptable 0.75% a 1.5% Malo < 0.75%
Porcentaje de clientes que han firmado cartas de descuento bancario para el pago del servicio	$\frac{\text{N° clientes que firmaron en periodo}}{\text{N° clientes en el periodo}} * 100$	Bueno >2.5% Aceptable 1.5% a 2.5% Malo <1.5%

Reportes y validaciones

Los siguientes criterios se utilizarán para los reportes que espera el área de gestión de clientes.

Los informes contarán con los filtros de todas las dimensiones establecidas; a fin de que el usuario pueda navegar y ver la información de cada nivel según desee.

Los informes deben mostrar la información por año de ingreso, excepto casos puntuales que determine el usuario.

El diseño de posición de los filtros, gráficos y tablas en los informes debe ser fácil de interpretar, con el fin de que tanto administrativos, como jefes de áreas e incluso supervisores, puedan entender la información presentada.

La fuente de información principal proviene del sistema transaccional de la empresa, por lo cual será esta la fuente de validación de los datos presentados en los reportes, aclarando que no existe ninguna otra fuente externa que deba ser considerada.

A continuación en la Tabla 7 se mostrará el listado de todos los reportes que el usuario desea de la implementación.

Tabla 7: Listado de Reportes

Informe	Detalle	Rol
Reporte de Total de Clientes	Cantidad total de clientes, Cantidad total de clientes activos, Cantidad total de clientes actualizados, Cantidad total de clientes que recibieron regalo y Cantidad total de clientes que firmaron cartas.	Supervisor y Jefe de Gestión de Clientes

Informe	Detalle	Rol
Reporte de Clientes por ubicación geográfica	Cantidad Total de clientes, Cantidad total de clientes activos, Cantidad total de clientes actualizados, Cantidad total de clientes que recibieron regalo, Cantidad total de clientes que firmaron cartas, todos por provincias, cantones y parroquias.	Supervisor y Jefe de Gestión de Clientes
Reporte de Clientes por datos generales	Cantidad de Clientes por Estado Civil, Rango de edades y género.	Supervisor y Jefe de Gestión de Clientes
Reporte de Clientes por datos laborales	Cantidad de Clientes por Relación Laboral, Jornada Laboral, Rango Nivel de Ingreso y Años de Servicios.	Supervisor y Jefe de Gestión de Clientes
Reporte de Clientes por datos empresariales	Cantidad de Clientes por Tipo de Establecimiento, Tipo de Cargo y Estado del Cliente.	Supervisor y Jefe de Gestión de Clientes
Reporte de Clientes por fecha de ingreso	Total de clientes por Fecha de Ingreso, Crecimiento de clientes Anual y Tiempo de Antigüedad.	Supervisor y Jefe de Gestión de Clientes
Reporte de Clientes Activos por datos generales	Cantidad de Clientes Activos por Estado Civil, Rango de edades y género.	Supervisor y Jefe de Gestión de Clientes
Reporte de Clientes Activos por datos laborales	Cantidad de Clientes Activos por Relación Laboral, Jornada Laboral, Rango Nivel de Ingreso y Años de Servicios.	Supervisor y Jefe de Gestión de Clientes
Reporte de Clientes Activos por datos empresariales	Cantidad de Clientes Activos por Tipo de Establecimiento, Tipo de Cargo y Estado del Cliente.	Supervisor y Jefe de Gestión de Clientes

Informe	Detalle	Rol
Reporte de Clientes Activos por fecha de ingreso	Total de clientes Activos por Fecha de Ingreso, Crecimiento de clientes Anual y Tiempo de Antigüedad.	Supervisor y Jefe de Gestión de Clientes
Reporte de Clientes Actualizados, con firma y con regalo por Datos Generales	Cantidad de Clientes Actualizados, con firma y con regalo por Estado Civil, Rango de edades y género.	Supervisor y Jefe de Gestión de Clientes
Reporte de Clientes Actualizados, con firma y con regalo por Datos Laborales	Cantidad de Clientes Actualizados, con firma y con regalo por Relación Laboral, Jornada Laboral, Rango Nivel de Ingreso y Años de Servicios.	Supervisor y Jefe de Gestión de Clientes
Reporte de Clientes Actualizado, con firma y con regalos por Datos Empresariales	Cantidad de Clientes Actualizados, con firma y con regalo por Tipo de Establecimiento, Tipo de Cargo y Estado del Cliente.	Supervisor y Jefe de Gestión de Clientes
Reporte de Clientes actualizados por fecha de ingreso	Total de clientes Actualizados por Fecha de Ingreso, Crecimiento de clientes Anual y Tiempo de Antigüedad.	Supervisor y Jefe de Gestión de Clientes
Reporte de Clientes con regalo por fecha de ingreso	Total de clientes con regalo por Fecha de Ingreso, Crecimiento de clientes Anual y Tiempo de Antigüedad.	Supervisor y Jefe de Gestión de Clientes
Reporte de Clientes con firma por fecha de ingreso	Total de clientes con firma por Fecha de Ingreso, Crecimiento de clientes Anual y Tiempo de Antigüedad.	Supervisor y Jefe de Gestión de Clientes

Casos de Uso

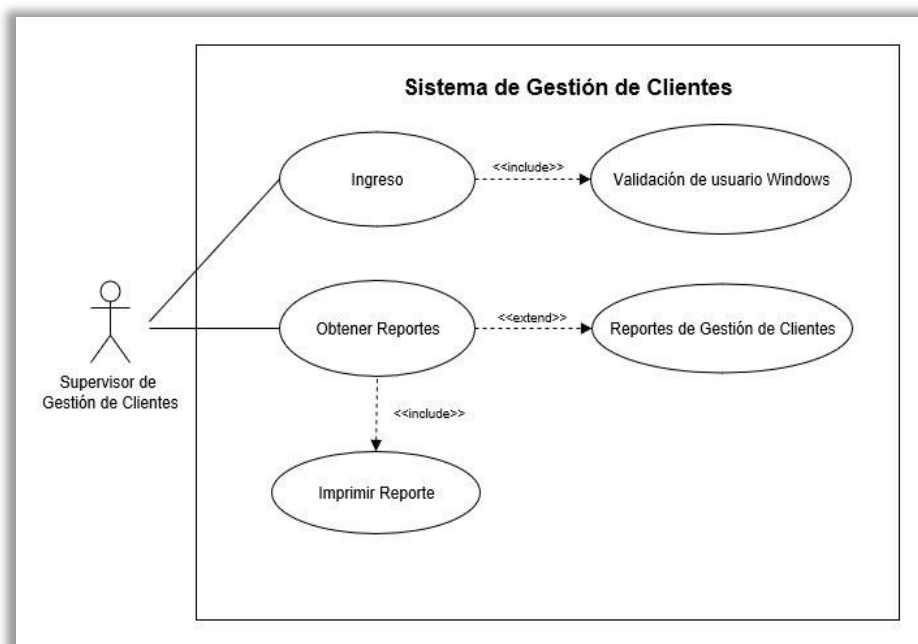


Figura 6.3: Caso de Uso FrontEnd

Caso de Uso FrontEnd.- El actor que interviene en este caso de uso, es el usuario final del Departamento de Gestión de Clientes, como lo son los Supervisores y el Jefe de área que pueden acceder y generar los reportes dependiendo de las necesidades. Es donde el usuario obtiene los datos que ayudan a tomar buenas decisiones.

- El usuario ingresará a la aplicación de Excel y automáticamente accederá con su usuario de Windows para actualizar o ejecutar los reportes para lo cual se debió dar autorización.

- Se le permitirá al usuario ingresar o modificar los parámetros necesarios de manera que pueda realizar las consultas del reporte como desee y lo puede hacer en cualquier momento.
- Se da la opción al usuario de imprimir los reportes generados.

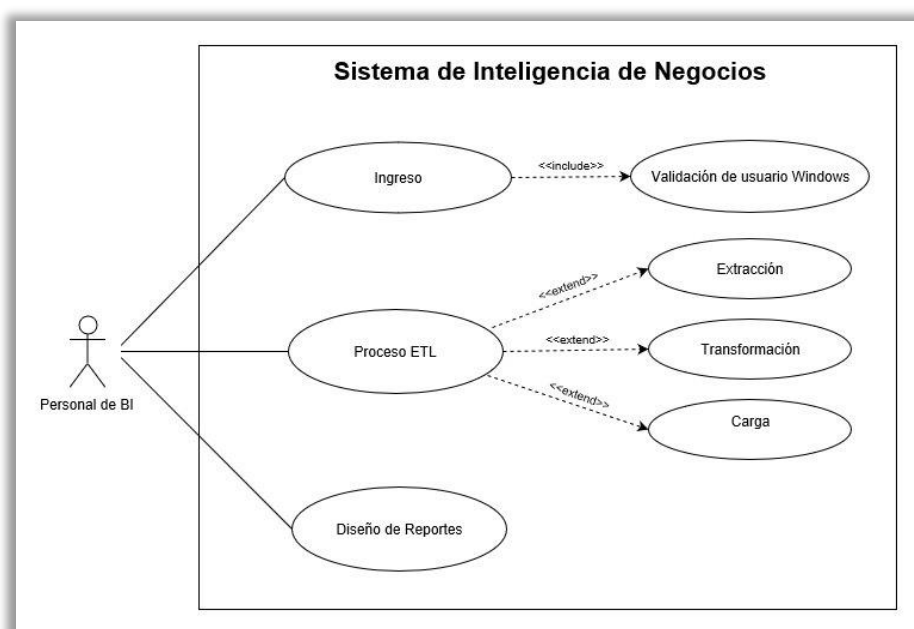


Figura 6.4: Caso de Uso Backend

Caso de Uso Backend.- Se encuentran los procesos responsables de la extracción de datos de diversos orígenes, que llegan al repositorio general eliminando posibles datos duplicados e inconsistencias. Conocido como extracción, transformación y carga.

El actor que interviene en este caso de uso es el Analista de BI, que es el que diseñara todo el proceso de BI, como diseño de reportes, permisos de acceso y extracción de datos.

- Agregará a los usuarios que tendrán permiso de acceso a los reportes.
- Programará la ejecución de los procesos de ETL en horarios nocturnos donde no afecte a la transaccionabilidad de la empresa.
- Realizar análisis de los datos disponibles para realizar el diseño de los reportes según el requerimiento.

3.3 Diagnóstico de Recursos Técnicos

Actualmente la empresa no cuenta con equipos totalmente óptimo para la implementación de un cubo de información para segmentación de clientes, debido a que los equipos actuales, aunque sean compatibles, el espacio en disco y memoria comprometerían los tiempos de respuestas, a nivel de software tenemos los aplicativos transaccionales, de los que no se realiza una intervención profunda, otro aplicativo es la base de datos SQL Server 2012 que es compatible con las funcionalidades requeridas.

Tabla 8: Recursos de Hardware y Software

Recursos Tecnológicos	Características
Hardware	Core i5 de 2da generación
	8Gb RAM
	500 GB Disco Duro de 7400rpm
Software	Excel 2013
	Aplicativo transaccional
	SQL Server 2012

Los equipos de los usuarios involucrados tienen el hardware indicado para la implementación del proyecto y acceso a los aplicativos de la empresa, así como la herramienta de Excel 2013 para los reportes requeridos.

CAPÍTULO 4

DISEÑO DE LA SOLUCIÓN

4.1 Diseño y definición de Modelo de Datos

Actualmente la fuente de datos para este proyecto no cumple con la estandarización necesaria, como un modelo entidad relación, por lo tanto los datos no tienen integridad, en ocasiones están duplicados y existe inserción de datos null, por lo que es necesario diseñar una nueva base de datos estandarizada, donde reconoceremos los campos que contienen la información necesaria para generar el cubo de información.

A continuación se podrá visualizar el esquema de la fuente de datos de manera general, donde se evidencia la escasa o nula relación de entidades:

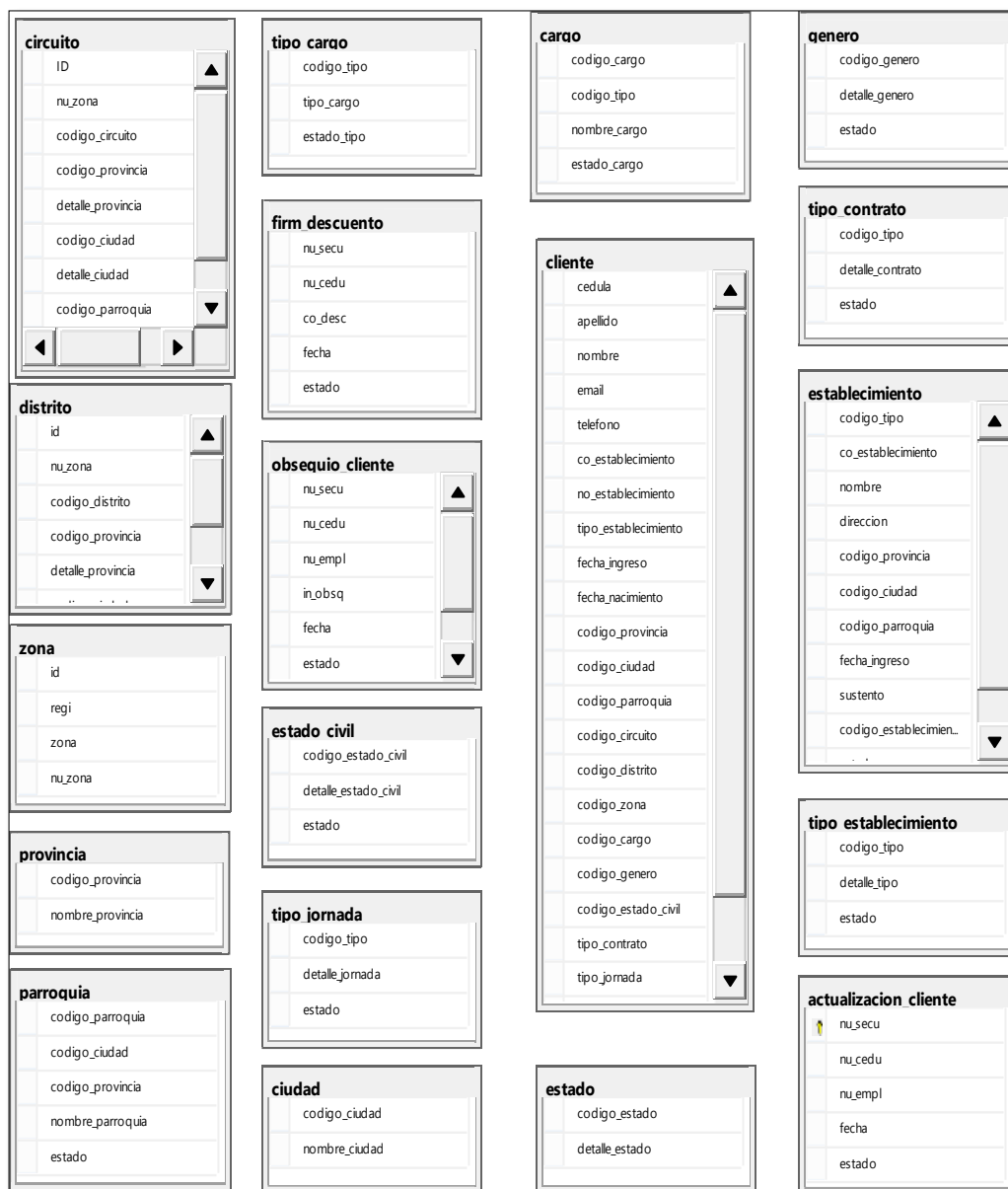


Figura 8.1: Esquema de Fuentes de Datos

Se debe recalcar que no existirán modificaciones al modelo inicial, debido que al realizar este paso se ven afectadas aplicaciones antiguas del sistema que no van a ser modificadas, ya que no pertenecen al desarrollo de este proyecto.

Entre las principales tablas se tienen:

- Tabla de Ciudad: Esta es una tabla catálogo que contiene todas las ciudades del Ecuador.
- Tabla de Parroquia: Tabla catálogo que contiene todas las parroquias del Ecuador.
- Tabla de Provincia: Es una tabla catálogo que contiene las provincias del Ecuador.
- Tabla Zona: Es una tabla catálogo que contiene las 8 zonas del Ecuador.
- Tabla Distrito: Es una tabla catálogo que contiene los distritos del Ecuador.
- Tabla Circuito: Es una tabla catálogo que contiene los circuitos del nuevo modelo de gestión del Ecuador.
- Tabla Cliente: Esta tabla contiene los datos personales del cliente como nombres, apellido, dirección, teléfono, etc.
- Tabla Género: Esta tabla es de tipo catálogo, contiene el sexo o género de un cliente.
- Tabla Estado Civil: Es una tabla catálogo que contiene el estado civil o condición social de un cliente, como por ejemplo: Soltero, Casado, Viudo, etc.
- Tabla tipo_cargo: Es una tabla catálogo que contiene el tipo de cargo, puesto que puede tener un cliente en un trabajo, como por ejemplo: Administrativo, Director, Servicio, etc.

- Tabla tipo_contrato: Es una tabla catálogo que contiene el tipo de contrato de un cliente, la forma en la que se relaciona con una empresa.
- Tabla obsequio cliente: Tabla que contiene la cédula de los clientes que recibieron obsequios en las campañas.
- Tabla firm_descuento: Tabla que contiene la cédula de los clientes que firmaron cartas de descuentos.
- Tabla tipo_establecimiento: Tabla que contiene el tipo de establecimiento al que pertenece un cliente.
- Tabla establecimiento: Tabla que contiene los datos del establecimiento al que pertenece un cliente.
- Tabla tipo_jornada: Tabla catálogo que contiene el tipo de jornada u horario laboral de un cliente.
- Tabla cargo: Tabla catálogo de los puestos de trabajo o cargo de los clientes.
- Tabla estado: Tabla que tiene los registros de los estados que tienen los clientes.
- Tabla actualización_cliente: Tabla que contienen las actualizaciones de datos de los clientes.

En la Tabla 9 se muestra las herramientas y plataforma que se utilizarán para el diseño del Datawarehouse del proyecto:

Tabla 9: Herramientas BI a utilizar para el DW

Herramientas BI a utilizar para el DW	
SQL Server	Base de datos para el DW
SQL Server Analysis Services (SSAS)	Plataforma de BI
SQL Server Integration Services (SSIS)	Procesos de ETL

4.2 Diseño de procesos de extracción, transformación y carga (ETL)

El proceso de ETL, es un paso necesario para extraer la información del datasource a una base de datos estandarizada antes de su extracción final al DW. En este proyecto se llevará a cabo dos extracciones debido a que se necesita consolidar la información en una sola fuente de datos en la que se encuentre exclusivamente la información solicitada para el proyecto según el requerimiento sin tener que acceder directamente a las diferentes base de datos transaccionales, permitiéndonos depurar la información y procesar el DW más fácil.

La fuente de datos del área de Gestión de clientes es una base que no está relacionada de la manera correcta, adicional se visualiza que para cada campaña crean tablas nuevas innecesarias y no posee una tabla unificada para dichos procesos, lo cual nos indica que existe un alto consumo de disco duro evidenciando el mal uso de la base de datos.

Con el diseño que se quiere implementar para la extracción de datos se busca obtener información consolidada y descubrir las inconsistencias y duplicidad existentes.

Para realizar la extracción se utilizarán dos tablas temporales que ayudarán a procesar la información, ya que servirá para actualizar los campos de las campañas de la tabla principal.

Las siguientes tablas son las involucradas para dicha extracción:


tqclie	temporal
 nsecu	ncedu
ncedu	
dapel	
dnomb	
dema	
dtele	
nestb	
destb	
desta	
nesta	
fingr	
nprov	
dprov	
nciud	
dciud	
ndist	
nparr	
dparr	
ncirc	
nzona	
ncarg	
tcarg	
dtcar	
dcarg	

Figura 8.2: Tablas temporales

La segunda extracción es la que se encargará de llenar el Datawarehouse de manera sencilla y rápida, al poseer la extracción intermedia permitirá ejecutar el Datawarehouse durante el día de ser necesario, sin comprometer la base de datos transaccional e interrumpir el servicio a los clientes.

Tabla 10: Fuente de datos para ETL Dimensiones

Nombre de la Dimensión	Tablas involucradas de la base de datos transaccional
Ubicación División Política Administrativa	Provincia
	Ciudad
	Parroquia
Ubicación nuevo modelo de gestión	Zona
	Distrito
	Circuito
Relación Laboral	Cliente
Rango Nivel de Ingreso	Cliente
Jornada Laboral	tipo_jornada y cliente
Rango Años Servicios	Cliente
Estado Civil	estado_civil y cliente
Rango de Edades	Cliente
Genero	Género y cliente
Tipo de Institución	establecimiento, tipo_establecimiento y cliente
Tipo de Cargo	tipo_cargo, cargo y cliente
Estado	Estado y cliente

En la Tabla 11 se detallan la fuente de datos para las medidas del diseño del ETL:

Tabla 11: Fuentes de datos para ETL Medidas

Nombre de Medida	Tablas Involucradas de la base de datos transaccional
Número de Clientes	Clientes
Número de Clientes activos	Clientes y estado
Número de clientes que han actualizado datos	actualizacion_clientes y clientes
Número de clientes que han recibido premios	obsequio_clientes y clientes
Número de clientes que han firmado cartas de descuento	firm_descuento y clientes

Las herramientas a utilizar en este diseño son:

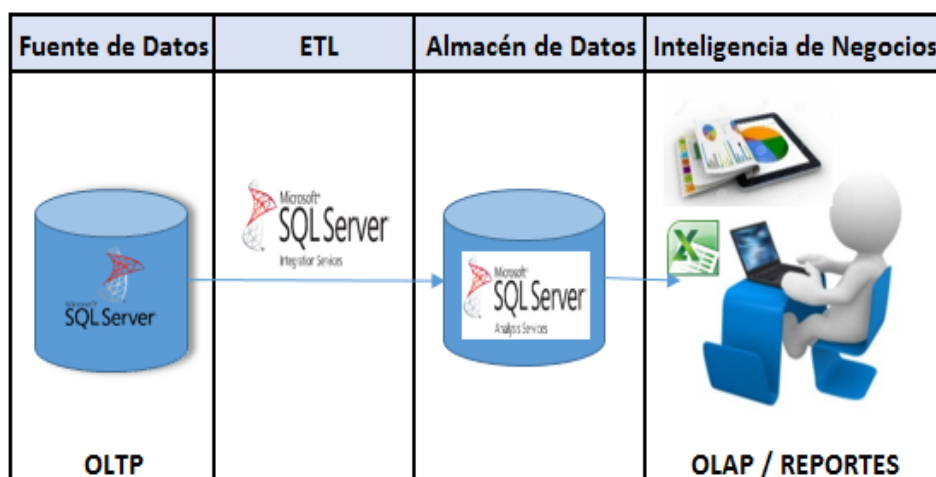


Figura 8.3: Arquitectura del DW

4.3 Diseño de Cubo OLAP

El diseño con el que se va a implementar el cubo de información es el modelo tipo estrella, este diseño centralizara toda la información, mejorando los tiempos de respuestas. A continuación se mostrara el diseño dimensional, tipo estrella, para el proceso de segmentación de clientes.

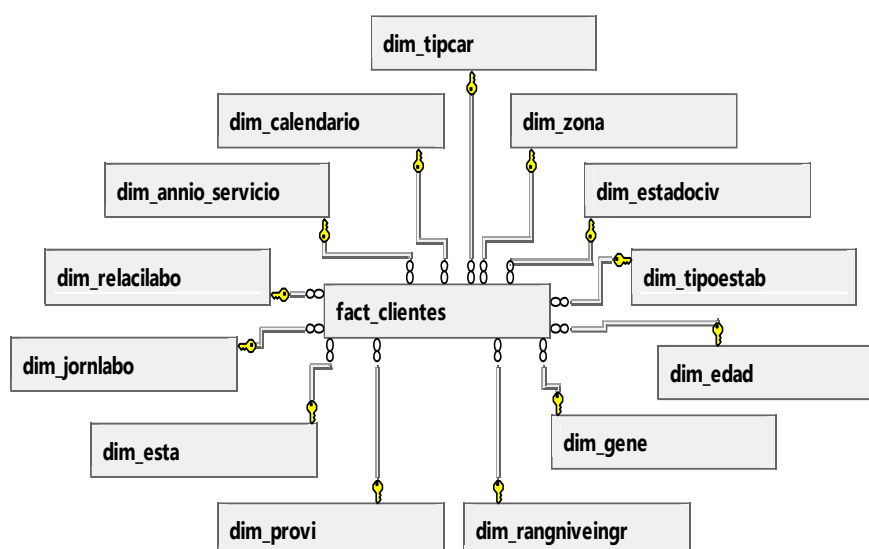


Figura 8.4: Diseño Estrella

La tabla de Hecho se centra en la de clientes la cual permitirá una visión amplia para la segmentación de clientes, y permitirá enfocarse en los distintos aspectos de análisis que el área de gestión de clientes necesita.

4.4 Diseño de reportes

El diseño de los informes se realiza en base al requerimiento del departamento de Gestión del cliente.

A continuación se detallara la funcionalidad de los diferentes reportes:

Reporte del Total de Clientes: Indican el número total de clientes segmentados por provincias, cantón y parroquia filtrado por Año y Mes.

Tabla 12: Definición de Reportes de Gestión de Clientes

Informe	Detalle
Reporte de Total de Clientes	Cantidad total de clientes, Cantidad total de clientes activos, Cantidad total de clientes actualizados, Cantidad total de clientes que recibieron regalo y Cantidad total de clientes que firmaron cartas.
Reporte de Clientes por ubicación geográfica	Cantidad Total de clientes, Cantidad total de clientes activos, Cantidad total de clientes actualizados, Cantidad total de clientes que recibieron regalo, Cantidad total de clientes que firmaron cartas, todos por provincias, cantones y parroquias.
Reporte de Clientes por datos generales	Cantidad de Clientes por Estado Civil, Rango de edades y género.
Reporte de Clientes por datos laborales	Cantidad de Clientes por Relación Laboral, Jornada Laboral, Rango Nivel de Ingreso y Años de Servicios.
Reporte de Clientes por datos empresariales	Cantidad de Clientes por Tipo de Establecimiento, Tipo de Cargo y Estado del Cliente.
Reporte de Clientes por fecha de ingreso	Total de clientes por Fecha de Ingreso, Crecimiento de clientes Anual y Tiempo de Antigüedad.
Reporte de Clientes Activos por datos generales	Cantidad de Clientes Activos por Estado Civil, Rango de edades y género.
Reporte de Clientes Activos por datos laborales	Cantidad de Clientes Activos por Relación Laboral, Jornada Laboral, Rango Nivel de Ingreso y Años de Servicios.
Reporte de Clientes Activos por datos empresariales	Cantidad de Clientes Activos por Tipo de Establecimiento, Tipo de Cargo y Estado del Cliente.
Reporte de Clientes Activos por fecha de ingreso	Total de clientes Activos por Fecha de Ingreso, Crecimiento de clientes Anual y Tiempo de Antigüedad.
Reporte de Clientes Actualizados, con firma y con regalo por Datos Generales	Cantidad de Clientes Actualizados, con firma y con regalo por Estado Civil, Rango de edades y género.

Informe	Detalle
Reporte de Clientes Actualizados, con firma y con regalo por Datos Laborales	Cantidad de Clientes Actualizados, con firma y con regalo por Relación Laboral, Jornada Laboral, Rango Nivel de Ingreso y Años de Servicios.
Reporte de Clientes Actualizado, con firma y con regalos por Datos Empresariales	Cantidad de Clientes Actualizados, con firma y con regalo por Tipo de Establecimiento, Tipo de Cargo y Estado del Cliente.
Reporte de Clientes actualizados por fecha de ingreso	Total de clientes Actualizados por Fecha de Ingreso, Crecimiento de clientes Anual y Tiempo de Antigüedad.
Reporte de Clientes con regalo por fecha de ingreso	Total de clientes con regalo por Fecha de Ingreso, Crecimiento de clientes Anual y Tiempo de Antigüedad.
Reporte de Clientes con firma por fecha de ingreso	Total de clientes con firma por Fecha de Ingreso, Crecimiento de clientes Anual y Tiempo de Antigüedad.

Número de clientes total, actualizados, activos, que recibieron regalo y firmaron carta.- Muestran el alcance logrado de clientes en la empresa, cuantos en total recibieron regalo, cuantos firmaron carta, están activos o actualizaron datos.

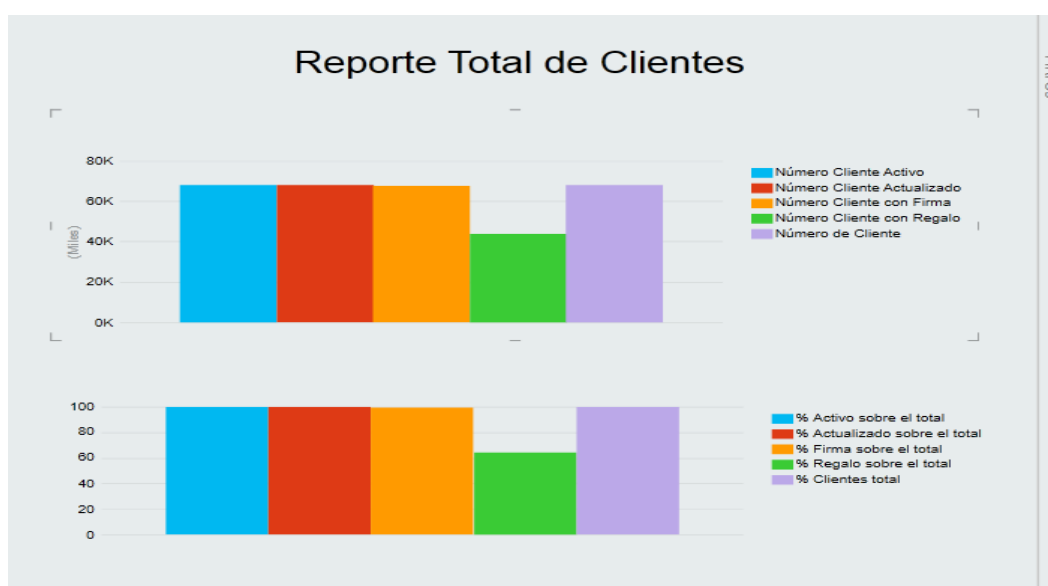
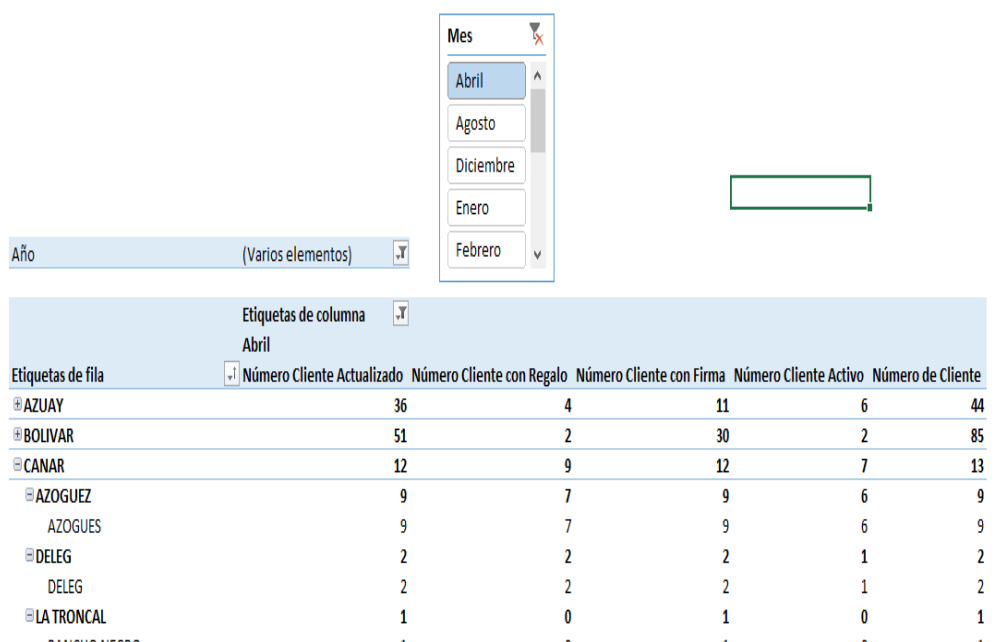


Figura 8.5: Reporte Total de Clientes

Número de clientes por provincias: Muestran el alcance logrado en cada provincia, cantón o parroquia por un mes específico y año.



Mes:

 Año: (Varios elementos)

Etiquetas de fila	Número Cliente Actualizado	Número Cliente con Regalo	Número Cliente con Firma	Número Cliente Activo	Número de Cliente
AZUJAY	36	4	11	6	44
BOLIVAR	51	2	30	2	85
CANAR	12	9	12	7	13
AZOGUEZ	9	7	9	6	9
AZOGUES	9	7	9	6	9
DELEG	2	2	2	1	2
DELEG	2	2	2	1	2
LA TRONCAL	1	0	1	0	1
LA TRONCAL	1	0	1	0	1

Figura 8.6: Reporte de Clientes por Ubicación Geográfica

Número de clientes por Datos generales: Muestra el reporte de clientes por género, estado civil y Rango de Nivel de ingreso, también lo puede filtrar por año.

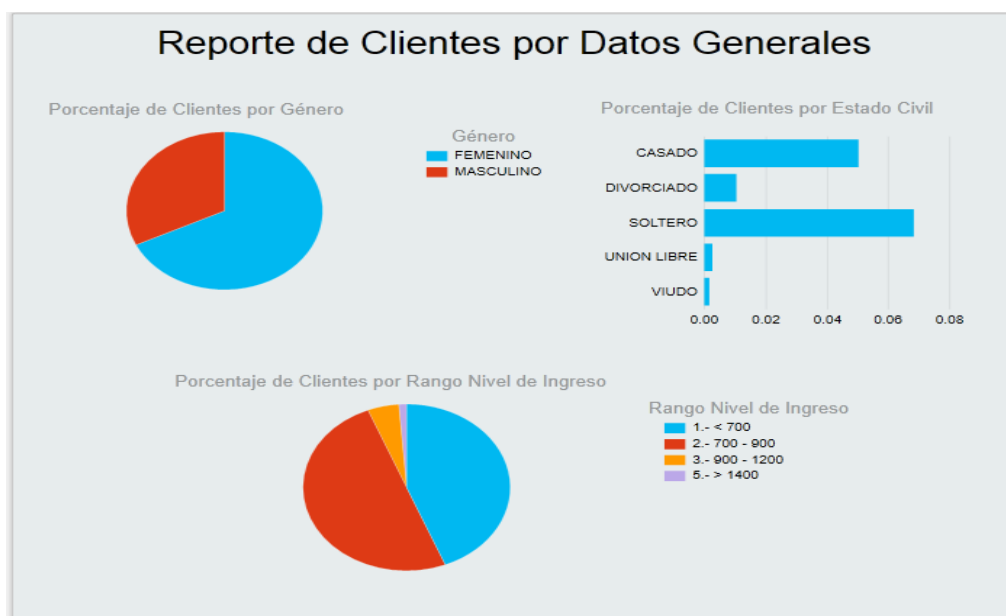


Figura 8.7: Reporte de Clientes por Datos Generales

Número de clientes por Datos Laborales: Muestra el reporte de clientes por Jornada Laboral, Rango de años de servicios, relación laboral también lo puede filtrar por año.

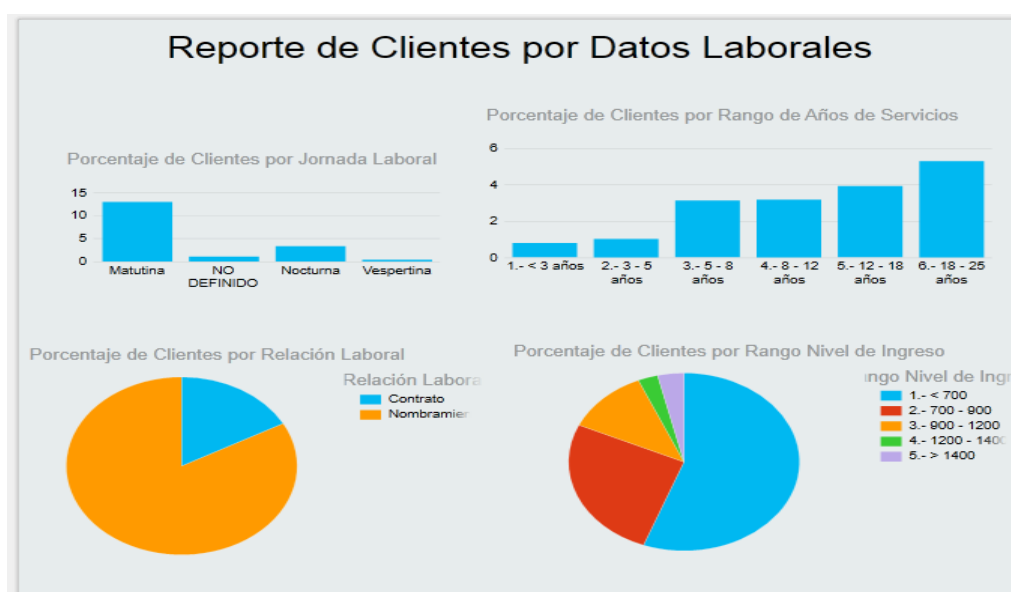


Figura 8.8: Reporte de Clientes por Datos Laborales

Número de clientes por Datos Empresariales: Muestra el reporte de clientes por Tipo de cargo, estado del cliente, tipo de establecimiento también lo puede filtrar por año.

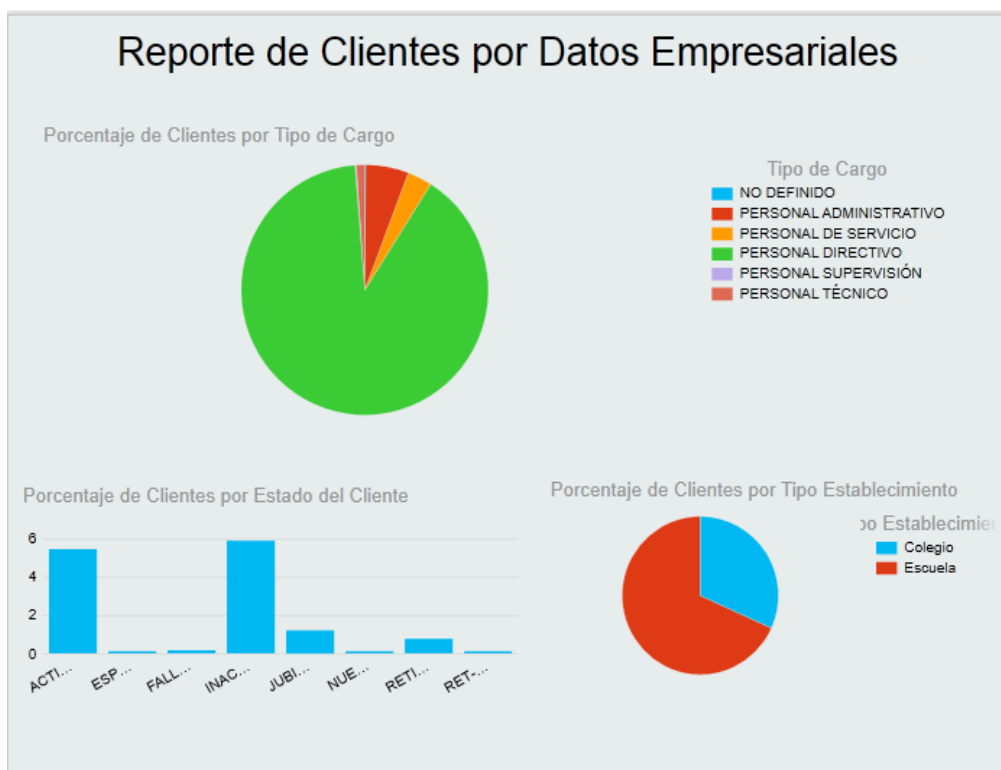


Figura 8.9: Reporte de Clientes por Datos Empresariales

Reporte de clientes por fecha de ingreso: Muestra la cantidad de clientes ingresados en cada año, el crecimiento de clientes cada año y la antigüedad de los clientes desde su fecha de ingreso.

Total de Clientes por Fecha de Ingreso		Crecimiento de clientes Anual		
Etiquetas de fila	Recuento de Número de Cliente	Año	Total de Clientes que Ingresan cada Año	Total de Clientes en el año
2010	4253	2010	12052	12052
2011	3502	2011	1208	13260
2012	2535	2012	825	14085
2013	1253	2013	2003	16088
2014	2535	2014	3525	19613
2015	352	2015	2054	21667
2016	1205	2016	758	22425
2017	2053	2017	3562	25987
2018	2568	2018	2538	28525
Total general	20256	Total general	28525	

Figura 8.10: Reporte de Clientes por fecha de ingreso

Número de clientes activos por Datos generales: Muestra el reporte de clientes activos por género, estado civil y Rango de Nivel de ingreso, también lo puede filtrar por año.

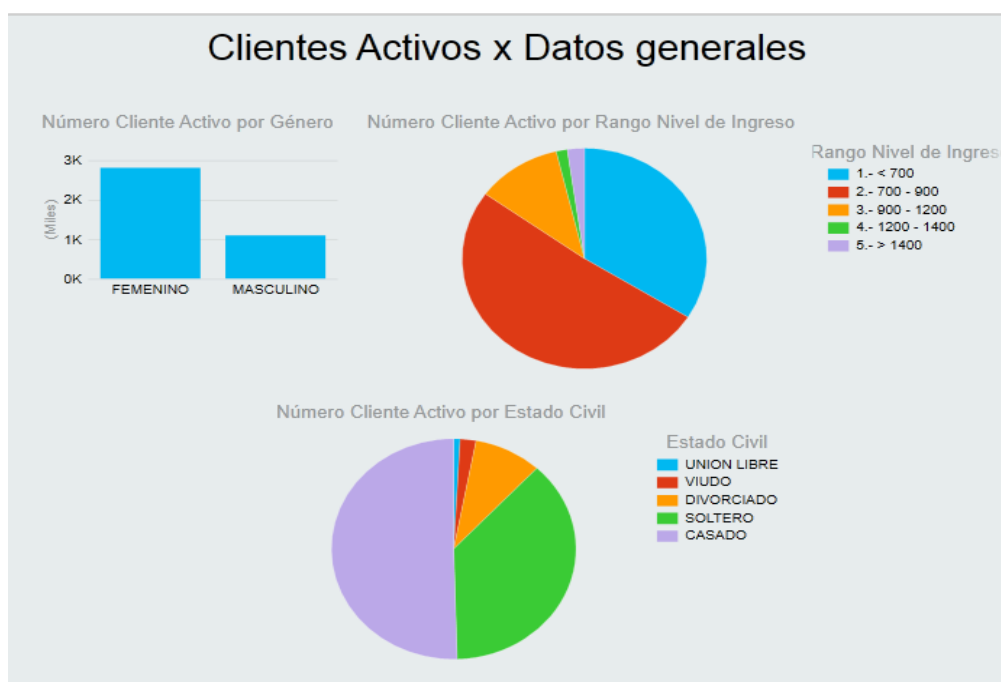


Figura 8.11: Reporte de Clientes activos por datos generales

Número de clientes activos por Datos Laborales: Muestra el reporte de clientes activos por Jornada Laboral, Rango de años de servicios, relación laboral también lo puede filtrar por año.

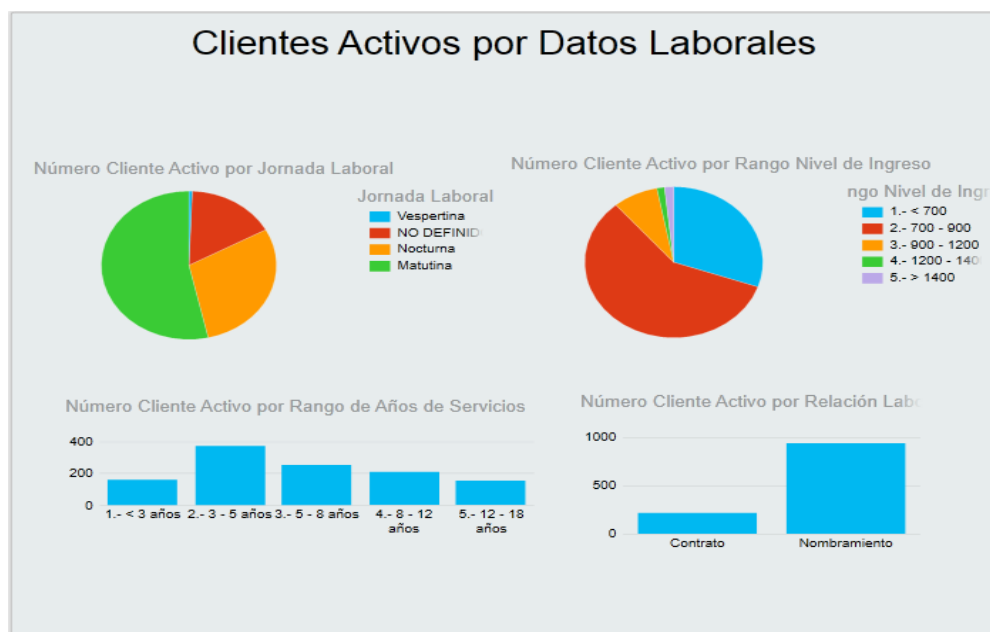


Figura 8.12: Reporte de Clientes activos por Datos Laborales

Número de clientes Activos por Datos Empresariales: Muestra el reporte de clientes activos, por Tipo de cargo, estado del cliente, tipo de establecimiento también lo puede filtrar por año.

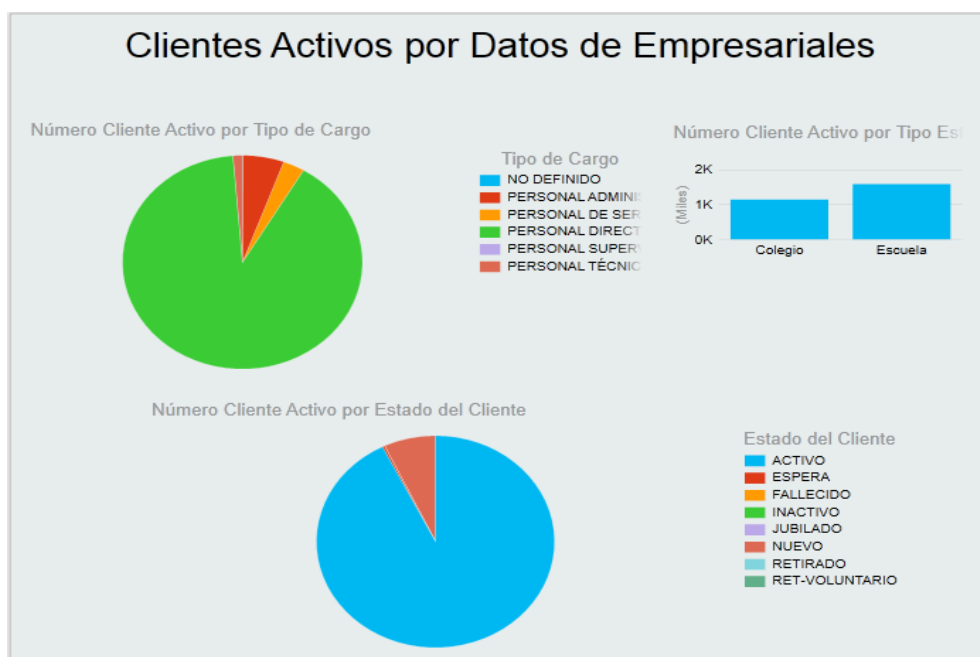


Figura 8.13: Reporte de Clientes activos por Datos Empresariales

Reporte de clientes activos por fecha de ingreso: Muestra la cantidad de clientes activos ingresados en cada año, el crecimiento de clientes cada año y la antigüedad de los clientes desde su fecha de ingreso.

Total de Clientes por Fecha de Ingreso		Crecimiento de clientes Anual		
Etiquetas de fila	Número Cliente Activo	Año	Cliente Activo cada Año	Total de Cliente Activo x Año
2010	2578	2010	2578	2578
2011	2124	2011	2124	4702
2012	524	2012	524	5226
2013	1258	2013	1258	6484
2014	2535	2014	2535	9019
2015	352	2015	352	9371
2016	2582	2016	2582	11953
2017	658	2017	658	12611
2018	2547	2018	2547	15158
Total general	15158	Total general	15158	

Figura 8.14: Reporte de Clientes activos por fecha de Ingreso

Reporte de datos generales de clientes que firmaron descuento, recibieron regalo y actualizaron datos: Este reporte da la visión total de los clientes que confían en la empresa.

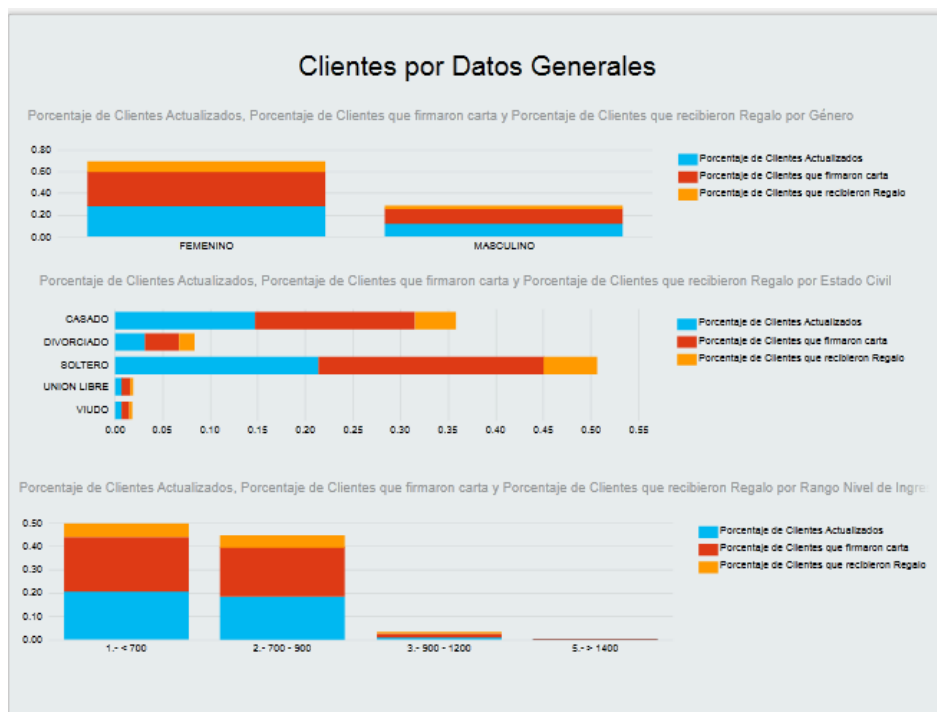


Figura 8.15: Clientes Actualizados, con firma y regalo por Datos Generales

Número de clientes actualizados, con firma y regalo por Datos Laborales: Muestra el reporte de clientes actualizados, con firma y regalo por Jornada Laboral, Rango de años de servicios, relación laboral también lo puede filtrar por año.

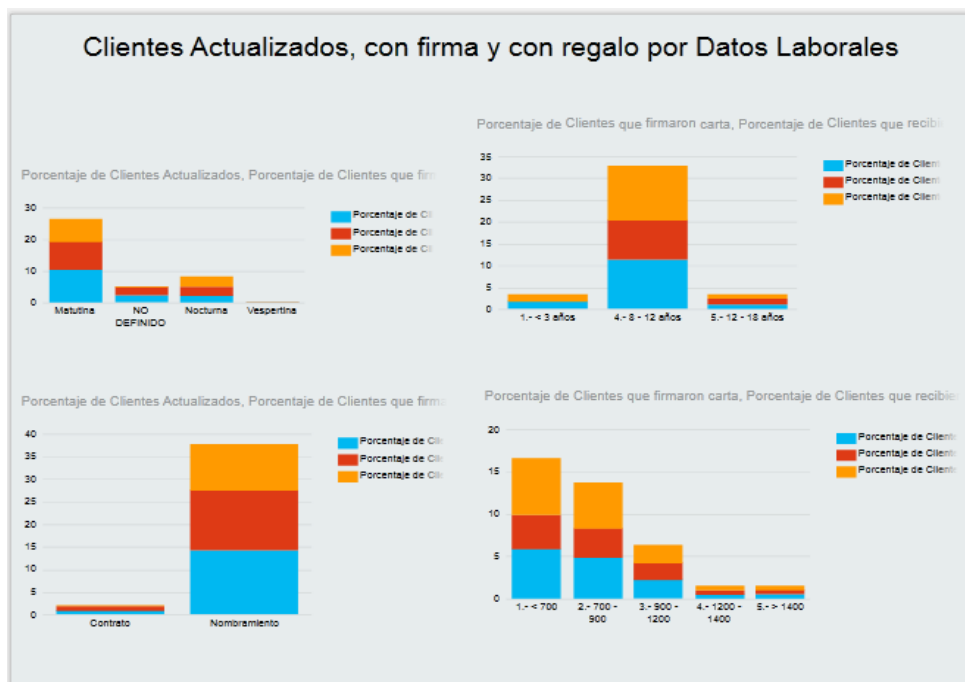


Figura 8.16: Clientes Actualizados, con firma y regalo por Datos Laborales

Número de clientes actualizados, con firma y regalo por Datos Empresariales:
 Muestra el reporte de clientes actualizados, con firma y regalo por Tipo de cargo, estado del cliente, tipo de establecimiento también lo puede filtrar por año.

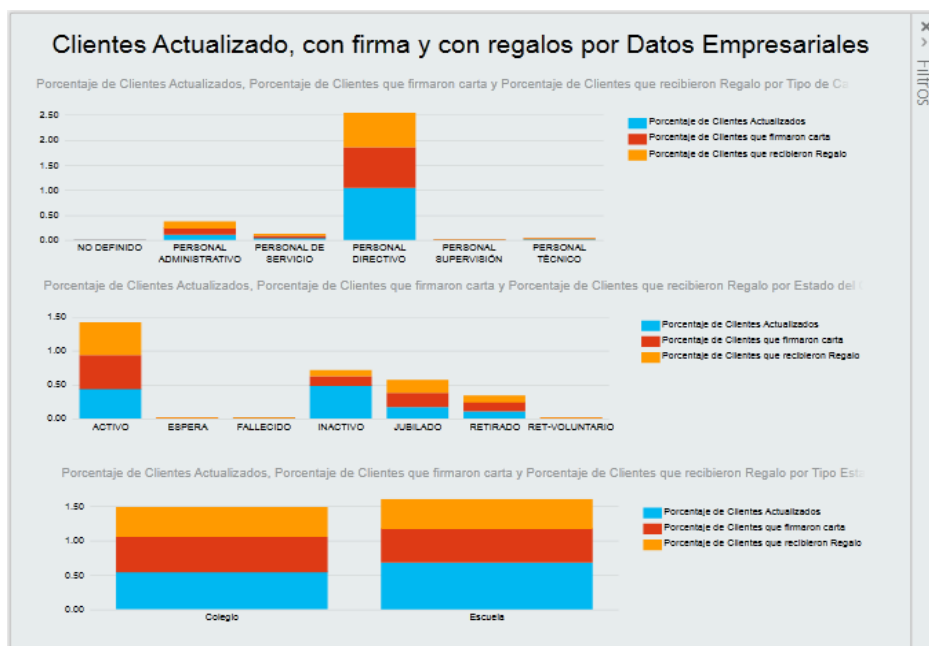


Figura 8.17: Clientes Actualizados, con firma y regalo por datos Empresariales

Reporte de clientes actualizados por fecha de ingreso: Muestra la cantidad de clientes actualizados en cada año, el crecimiento de clientes cada año y la antigüedad de los clientes desde su fecha de ingreso.

Total de Clientes por Fecha de Ingreso	
Etiquetas de fila	Suma de Número Cliente Actualizado
+ 2010	6253
+ 2011	4257
+ 2012	758
+ 2013	1425
+ 2014	2158
+ 2015	2535
+ 2016	1854
+ 2017	652
+ 2018	2152
Total general	22044

Figura 8.18: Clientes Actualizados por fecha de ingreso

Reporte de clientes con regalo por fecha de ingreso: Muestra la cantidad de clientes con regalo ingresados en cada año, el crecimiento de clientes cada año y la antigüedad de los clientes desde su fecha de ingreso.

Total de Clientes por Fecha de Ingreso	
Etiquetas de fila	Suma de Número Cliente con Regalo
+ 2010	1254
+ 2011	542
+ 2012	0
+ 2013	483
+ 2014	102
+ 2015	57
+ 2016	85
+ 2017	0
+ 2018	0
Total general	2523

Figura 8.19: Clientes con regalo por fechas de Ingreso

Reporte de clientes con firma por fecha de ingreso: Muestra la cantidad de clientes que firmaron cartas débitos ingresados en cada año, el crecimiento de clientes con cartas cada año y la antigüedad de los clientes con cartas desde su fecha de ingreso.

Crecimiento de clientes Anual		
Año	Suma de Número Cliente con Firma	Suma de Número Cliente con Firma2
2010	1487	1487
2011	2452	3939
2012	258	4197
2013	352	4549
2014	1245	5794
2015	2587	8381
2016	1781	10162
2017	352	10514
2018	826	11340
Total general	11340	

Figura 8.20: Clientes con firma por fechas de ingreso

Se detectaron un sin números de reportes que pueden surgir en los diferentes filtros y segmentos, por tal motivo, se fijó con el usuario que se les ofrecería los diseños solicitados, y que los demás los pueden elaborar con la herramienta, ya que da la facilidad de modificarlos a su gusto. Adicional muchos de los reportes solicitados, estaban contenidos en otros por lo que queda eliminada la redundancia y se unifica para obtener una mejor visualización de informes, logrando una amplia visión para la toma de decisiones oportuna.

CAPÍTULO 5

DESARROLLO E IMPLEMENTACIÓN DE LA SOLUCIÓN

5.1 Plan de Desarrollo e Implementación

Las tareas para la implementación de la solución, comprenden: preparación del servidor para la bodega de datos, instalación de las herramientas de Inteligencia de Negocios, así como el desarrollo y base de datos. A continuación en la Tabla 13 detalla los pasos antes mencionados:

Tabla 13: Actividades de Implementación de Bodega de datos

Tareas de Implementación de la bodega de datos	
Actividades	Descripción
Preparación de Infraestructura	Servidor para Bodega de datos.
Instalación de herramientas de Inteligencia de Negocios	SQL Server Integration Services.
	SQL Server Analysis Services.

Actividades	Descripción
Creación de Tablas en Base de datos	Creación de Dimensiones y tablas de hecho.
Creación de modelos ETL	Se desarrollan 2 modelos, una para la consolidación de datos en una sola fuente y la segunda para llenar la estructura de dimensiones y hechos.

El servidor de base de datos de datawarehouse cuenta con las siguientes instancias instaladas:

- Motor de Base de Datos.
- Integration Services.
- Analysis Services.
- SQL Server Data Tools.
- SSMS (SQL Server Management Studio).

5.1.1. Proceso de Extracción para Segmentación de Clientes

El desarrollo del ETL fue realizado con la herramienta de SQL Server Integration Services, instalado en una de las máquinas, a partir de este paso, se inicia con el desarrollo del proceso para consolidar la información en una sola fuente de información, llenando de esta manera tablas principales de clientes y tablas temporales, de esta manera se realizó con más facilidad el siguiente paso de extracción, en el segundo proceso se

implementa la extracción, transformación y carga de las diferentes dimensiones generales y de tiempo, además de la tabla de hecho.

Adicional se realizó un proceso de limpieza de información de datos que se fundamentó en:

- Eliminar tablas que ya no se utilizan y solo fueron creadas para actividades puntuales.
- Eliminación de datos inconsistentes (caracteres especiales en algunos campos de ciudad, provincia, nombres de clientes nulos, campos de género vacíos).
- Normalización de tipo de datos de algunos campos como la identificación del cliente que estaba numérico y se pasó a texto para almacenar la información de forma correcta.
- Creación de índices para realizar consultas más rápidas.

La eliminación de tablas se realizó con el apoyo del área de desarrollo, las cuales definieron las tablas que ya no se utilizaban y se podían eliminar para optimizar el espacio de disco en el servidor de base de datos.

El proceso de eliminación de datos inconsistentes se realizó con ayuda de los usuarios, se les envió un reporte de todas las inconsistencias y ellos se encargaron de actualizar la información correcta en algunos casos, en los registros que no se pudo actualizar definieron cuáles serían los datos por default para los campos que continuarán vacíos. En el caso de seguir llenando datos vacíos en las transacciones diarias, se estableció que el

proceso de extracción coloque como valor por default a dichos campos como NO DEFINIDO.

La normalización se realizó en algunas tablas que manejaban el campo de cédula del cliente como numérico, para este proceso se contó con el equipo de desarrollo y base de datos para dicha actualización.

La creación de índices se realizó en las tablas transaccionales, ya que no existían y este proceso ayudó a que algunas consultas sean más rápida para realizar la carga principal.

En las dimensiones generales, se extrae la información de la base consolidada de SQL Server 2012 a la base de datos Datawarehouse, entre las dimensiones tenemos, la de Jornada Laboral, Zona, Provincia, Rango de nivel de ingreso, estado de cliente, Años de Servicio, Género, Tipo de establecimiento, Rango de edad, Estado Civil, Relación Laboral, Tipo de Cargo y Calendario. La Figura 5.1 presenta la perspectiva realizada:

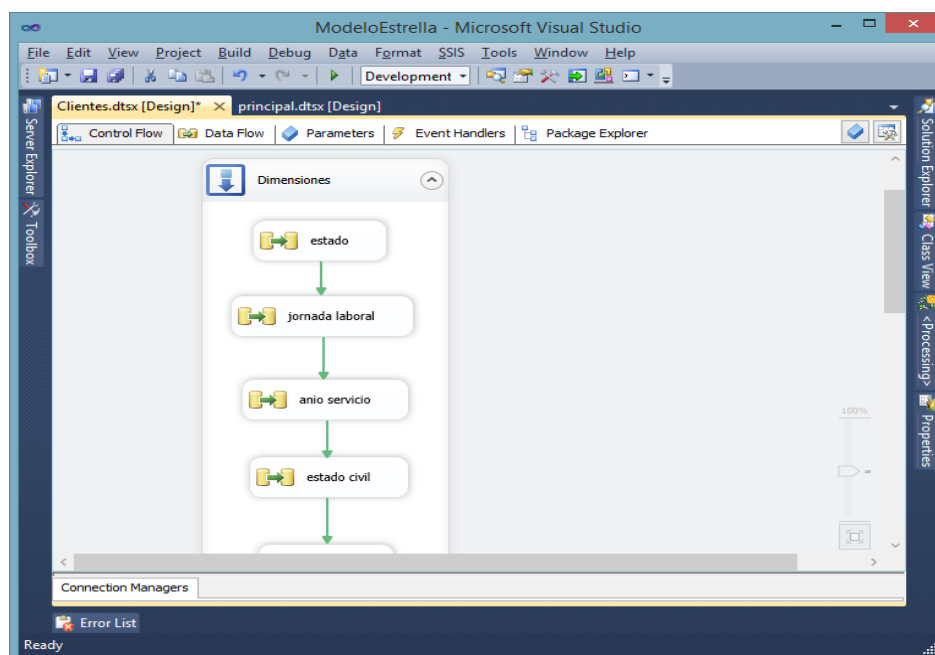


Figura 10.1: Dimensiones Generales

Para la dimensión calendario, se creó una tarea de ejecución de Sql para generar los registros de las fechas y luego almacenarlos en la bodega de datos de Sql Server 2012. En la Figura 5.2, se muestra la perspectiva desarrollada:

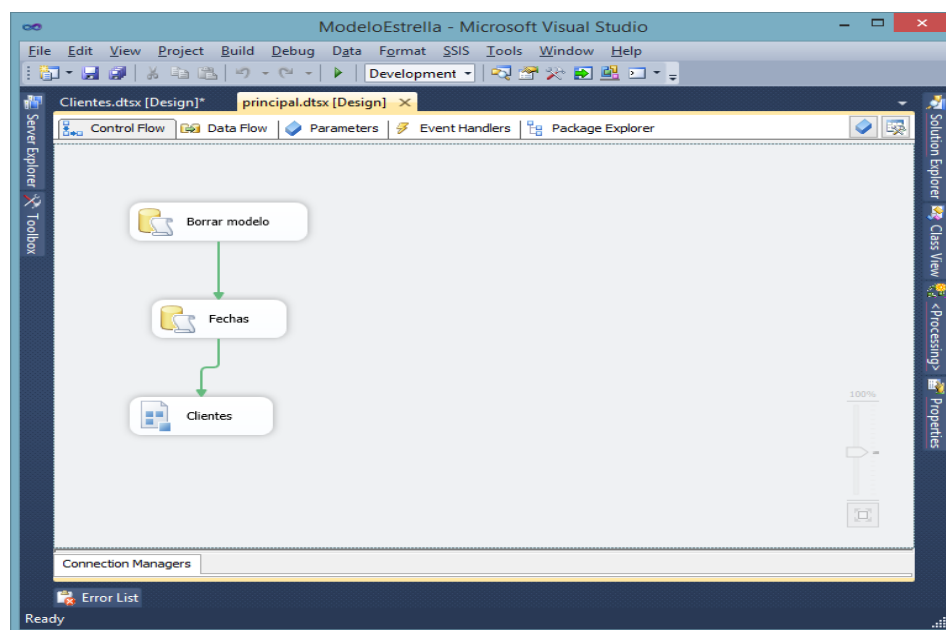


Figura 10.2: ETL-Dimensión Calendario (Tiempo)

En la perspectiva de hecho, se desarrolló el proceso ETL para cargar los registros de información de medidas, de acuerdo al modelo del área de Gestión de clientes. Dicho proceso toma información de la base de datos transaccional para trasladarla al modelo de base de datos datawarehouse. A continuación en la Figura 5.3 podemos visualizar lo antes mencionado:

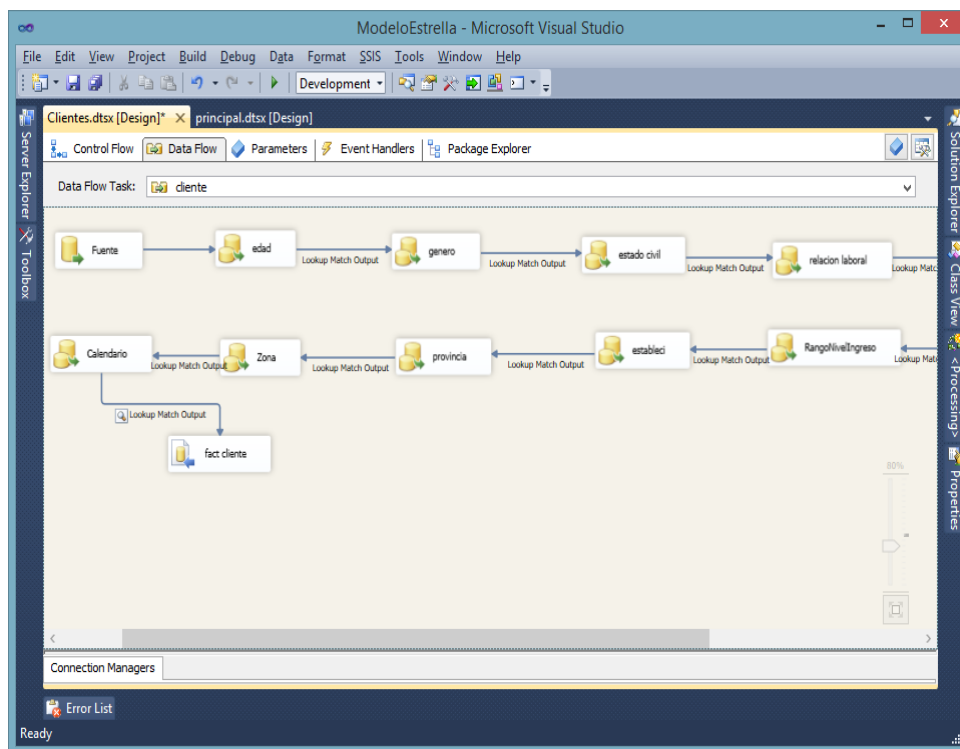


Figura 10.3: ETL-Hechos

5.1.2. Carga del proceso de segmentación de clientes

La carga inicial es la del proceso de llenar el repositorio que consolida toda la información para el procesamiento del modelo estrella, dicho proceso elimina la información duplicada, e inconsistencia encontradas, dejando en un repositorio de fácil acceso que no afecte a las transacciones del día a día, permitiendo consultar la información en cualquier momento del día sin que se vea afectado el negocio o alterada la información a cada minuto, por lo que se obtendrá información hasta el día anterior.

El proceso de carga inicia realizando shrink de la base de datos permitiendo truncar y reducir el fichero de log, reduciendo espacio

ocupado por el archivo log de la base de datos, como se realiza una carga total no se requiere realizar un respaldo del archivo, después continua limpiando la base de dato que se va a llenar, el siguiente paso es la carga de la tabla principal de clientes, y la actualización de los campos que son extraídos de otras bases que ayudarán a complementar los datos de la tabla de clientes como por ejemplo la actualización de clientes, quienes firmaron carta de compromiso de descuento, entre otros datos que están en diferente fuentes. A continuación la Figura 5.4 muestra la carga de la base consolidada:

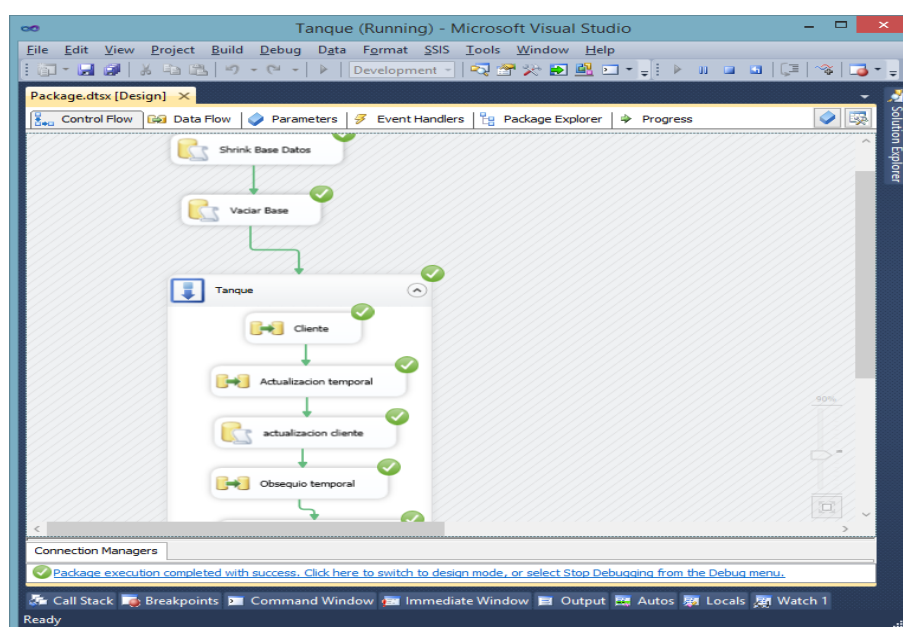


Figura 10.4: Carga de datos a base consolidada

El proceso de carga del modelo estrella, se realiza en 2 paquetes: el paquete principal que se encarga de limpiar las tablas del modelo estrella, llenar las tablas de fechas y llamar al paquete de clientes que es el

segundo paquete que se encarga de realizar la carga de toda la información del modelo estrella establecido para el proyecto.

A continuación en las Figuras 5.5 y 5.6 se muestran la ejecución exitosa de la carga del modelo estrella del proyecto:

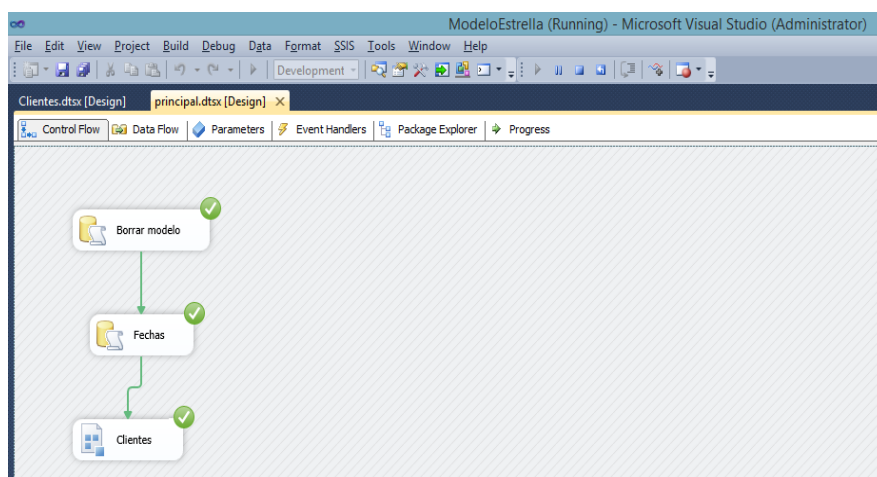


Figura 10.5: Carga principal del modelo Estrella

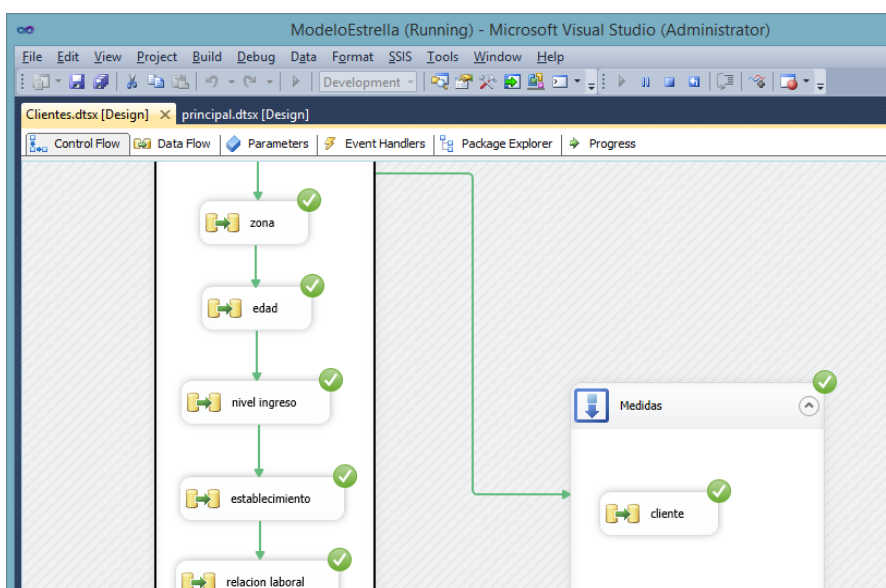


Figura 10.6: Carga del paquete de clientes

5.1.3. Cubo de Información de Segmentación de Clientes

Conforme al esquema diseñado en el capítulo anterior, tenemos todas las tablas de dimensiones que se unen con la tabla centralizada de hecho clientes, de la siguiente manera:

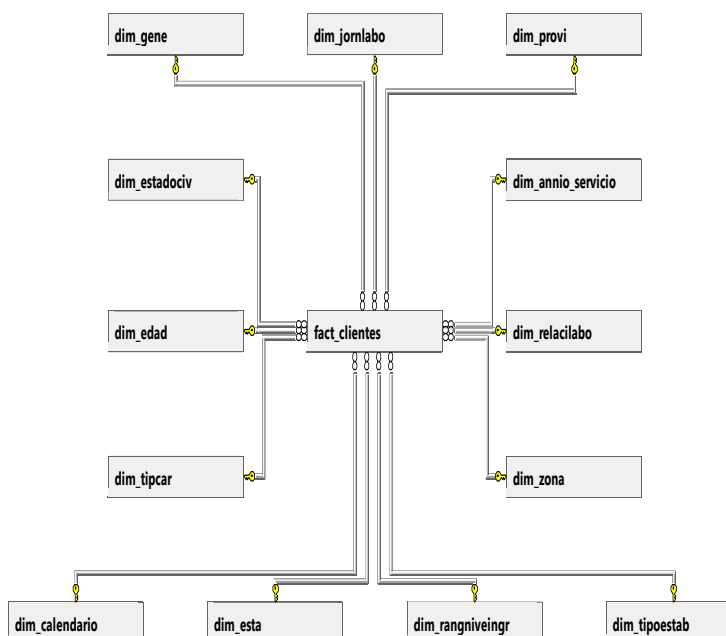


Figura 10.7: Implementación Esquema Estrella (Staging Database)

Para la generación del cubo se consideró lo siguiente:

- Se tomó las tablas principales del modelo entidad relación.
- Se centralizó todas las medidas o métrica en una sola tabla.
- La tabla centralizada se denomina tabla de hechos, y las demás son tablas dimensiones.
- Se definió las métricas (edad, género, provincia, ciudad, parroquia, etc.)

A continuación se mostrará la imagen del desarrollo del cubo de información:

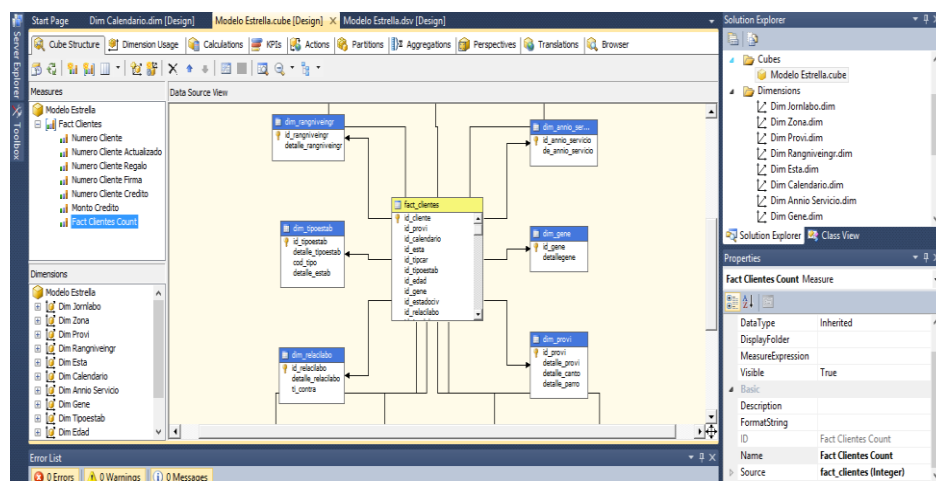


Figura 10.8: Desarrollo del cubo de información

Para esta implementación del cubo de información del área de Gestión de clientes, se utilizó la herramienta de Analysis Services de SQL Server, como servidor OLAP.

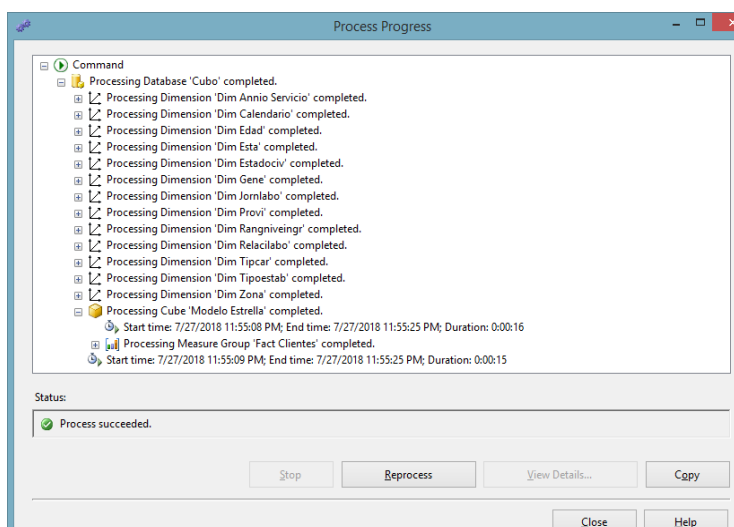


Figura 10.9: Procesamiento del cubo de información

Para la ejecución del cubo de información se coordinó con el administrador de base de datos, la configuración de un job que se ejecute todos los días, permitiendo obtener información actualizada para los usuarios.

Se solicitó permisos de select/insert/update/delete/alter de las tablas del modelo estrella para la ejecución de los paquetes de extracción en el servidor de base de datos de datawarehouse con el usuario definido para los procesos de BI.

Se definieron los siguientes pasos en el Job para que se ejecute la extracción, transporte y carga de la información del modelo estrella y el del procesamiento del cubo.

Tabla 14: Pasos de Ejecución

Pasos	Descripción
1	Ejecuta script EliminarConstraints
2	Ejecuta paquete de carga Inicial
3	Ejecuta paquete del Modelo Estrella
4	Ejecuta script CrearConstraints
5	Ejecuta paquete 'Carga Total Procesamiento Cubo'

Entre las actividades para la implementación del cubo de información se habilitaron permiso de acceso completo a los usuarios del cubo definidos en el proyecto.

La visualización del cubo de información desde la máquina del usuario se realiza desde Excel agregando una conexión a la fuente de datos donde residen los datos de la Bodega de datos, producto de la ejecución de los modelos de ETL.

Implementación de Reportes

Para la implementación de reportes se utiliza Excel 2013, herramienta que el usuario maneja a la perfección, para iniciar el proceso, se debe verificar que el Excel tenga habilitado las opciones de Power Pivot y Power View herramientas que se van a utilizar para la reportería, para esto debemos ir al menú de Archivo – Opciones.



Figura 10.10: Activación de Power Pivot en Excel

Luego en la siguiente pantalla nos vamos a la pestaña de complementos opción administrar y escoger “Complementos COM” y dar clic en el botón Ir.

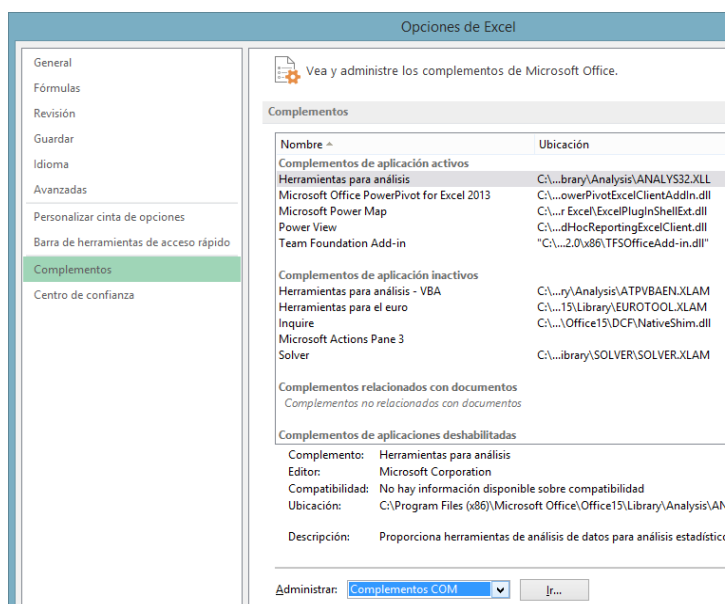


Figura 10.11: Opciones de Excel

A continuación aparecerá la siguiente pantalla y si no están activas las siguientes opciones la chequeamos, Microsoft Office PowerPivot for Excel 2013 y PowerView, damos clic en Aceptar y continuamos con la conexión.

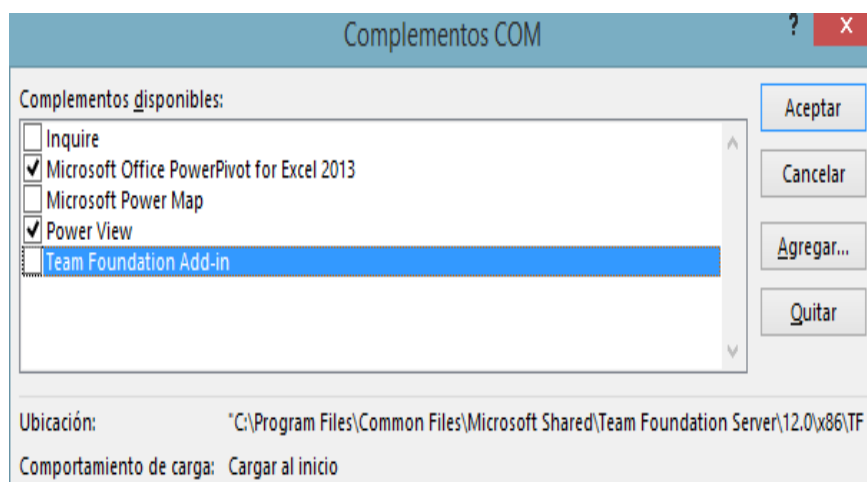


Figura 10.12: Activación Power Pivot y Power View

Luego de realizar los pasos anteriores, nos aparecerá la opción de PowerPivot en la pestaña de Excel, en esta pantalla dar clic en el botón Administrar.

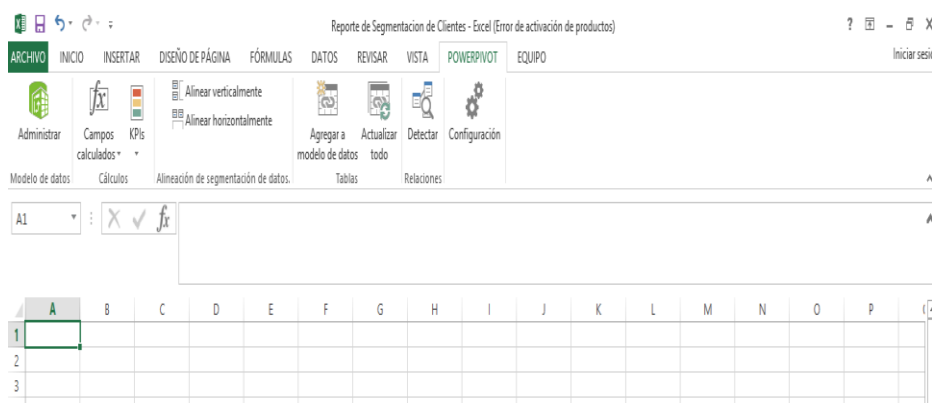


Figura 10.13: Vista de Power Pivot en Excel

A continuación nos aparece la siguiente pantalla:

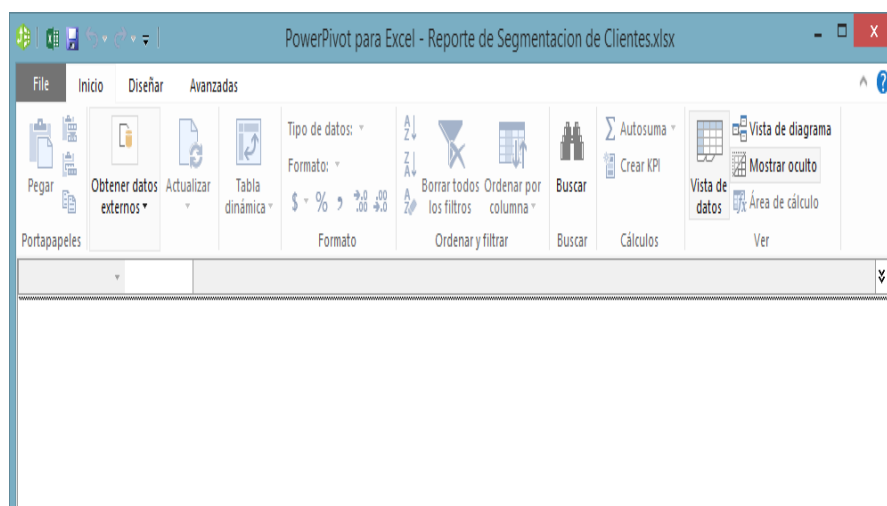


Figura 10.14: Conexión de datos externos

Nos vamos a dirigir al botón de Obtener datos externos - De base da de datos – De Analysis Services o PowerPivot.

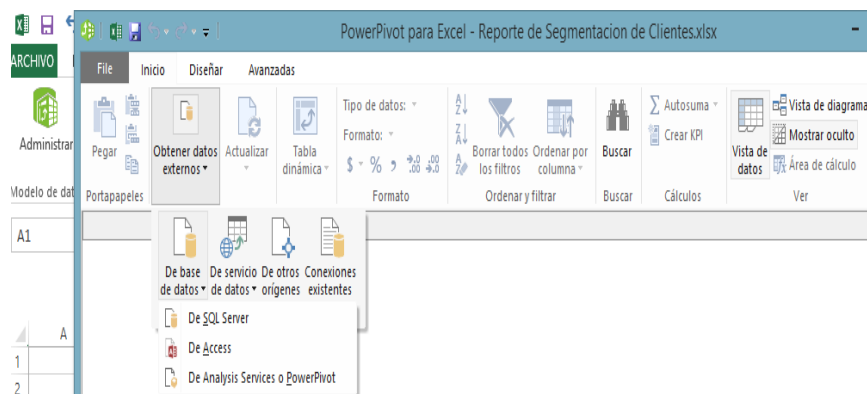


Figura 10.15: Conexión Analysis Services

Se debe usar la opción de usar la autenticación de Windows y escribir el nombre del servidor al cual se va acceder (servidor que contiene el cubo de información), en el nombre de la base de datos, se escribe el nombre del cubo y se da clic en el botón Siguiente.

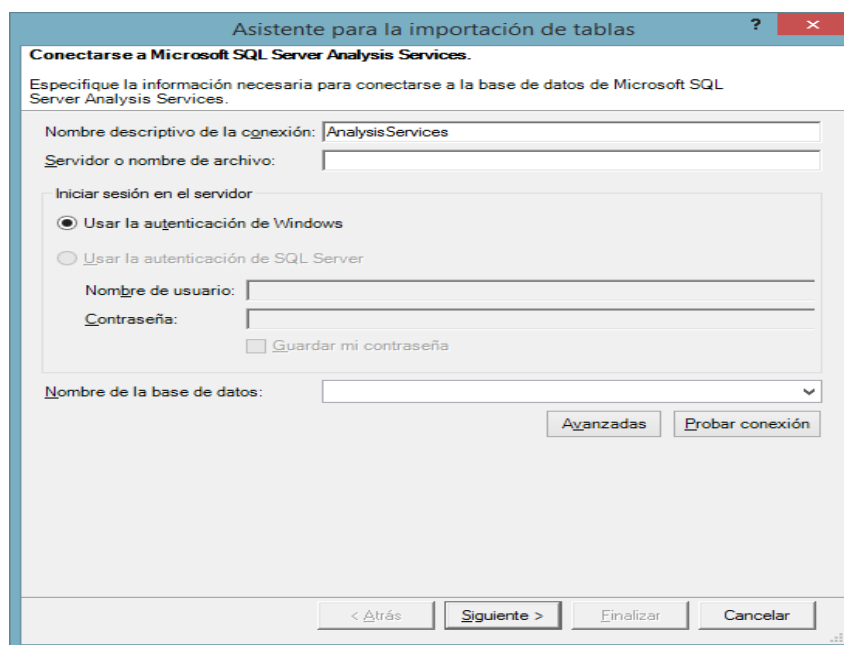


Figura 10.16: Asistente de Importación de tablas

Aparecerá la siguiente pantalla que realiza la conexión y el diseño de los campos que serán utilizados en los diferentes reportes.

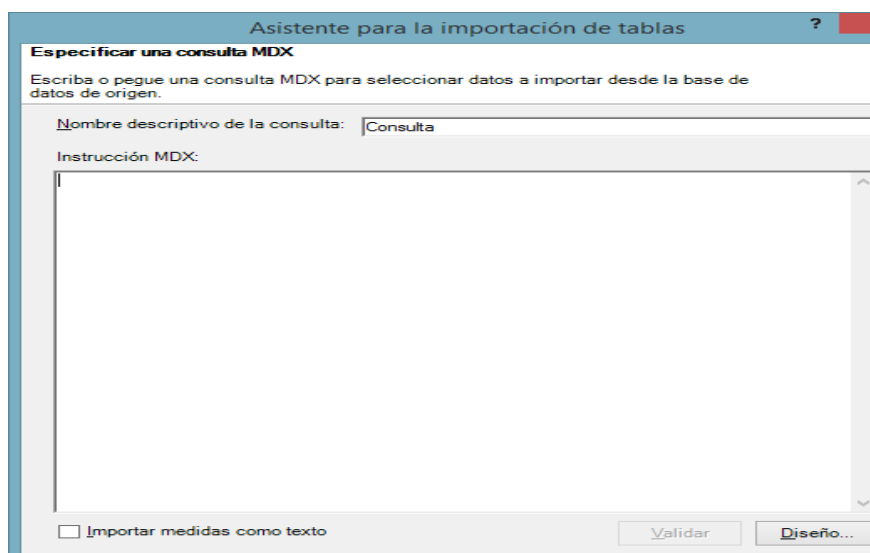


Figura 10.17: Especificación de consulta MDX

Al dar clic en el botón Diseño, aparece la siguiente pantalla donde se escogen los campos que aparecerán en el reporte.

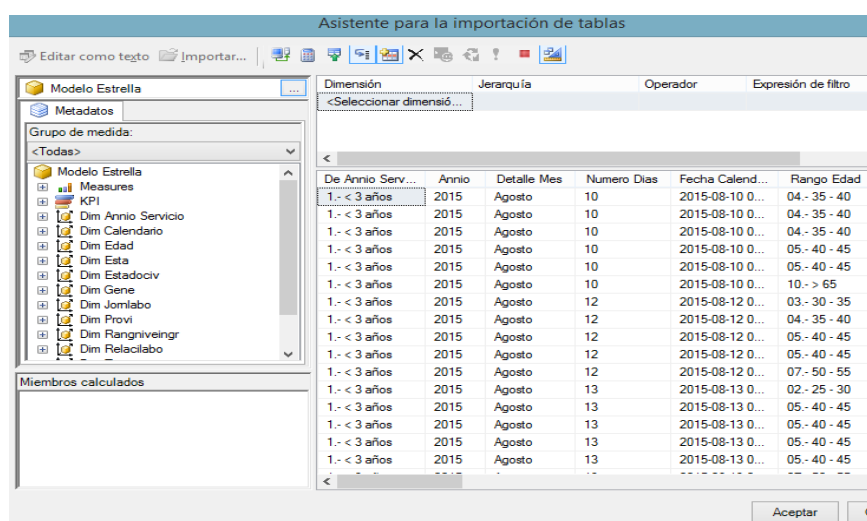


Figura 10.18: Pantalla de Diseño de Consultas

Cuando se termina se da clic en Aceptar y en Siguiente de la pantalla anterior para que inicie el procesamiento.

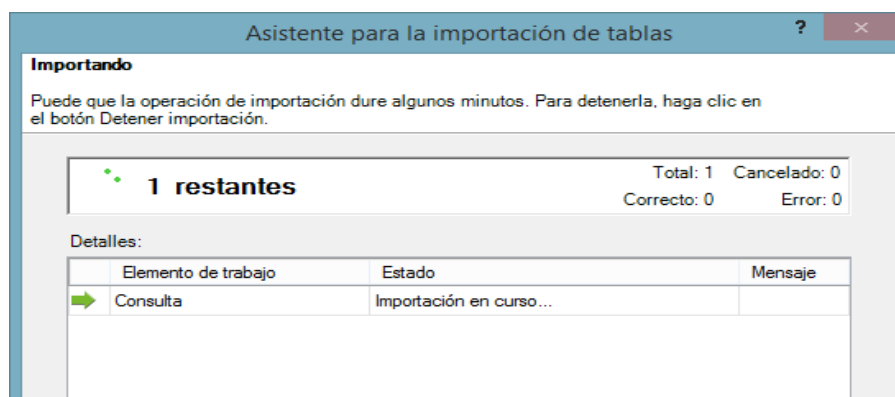


Figura 10.19: Importación de datos

Se da clic en Cerrar y visualizaremos los datos con los que se van a trabajar, se da clic en grabar y listo para realizar los diseños de reporte en PowerView.

Dim CalendarioAnno	Dim CalendarioDetalle Mes	Dim CalendarioNumero Dias	Dim CalendarioFecha	Dim EdadRango
1992	Noviembre	5	1992-11-05 00:00:00	07- 50 - 55
1992	Agosto	20	1992-08-20 00:00:00	08- 55 - 60
1992	Agosto	24	1992-08-24 00:00:00	09- 60 - 65
1992	Agosto	3	1992-08-03 00:00:00	08- 55 - 60
1992	Agosto	5	1992-08-05 00:00:00	08- 55 - 60
1992	Diciembre	20	1992-12-20 00:00:00	06- 45 - 50
1992	Diciembre	20	1992-12-20 00:00:00	07- 50 - 55
1992	Diciembre	22	1992-12-22 00:00:00	07- 50 - 55
1992	Diciembre	22	1992-12-22 00:00:00	09- 60 - 65
1992	Diciembre	23	1992-12-23 00:00:00	09- 60 - 65
1992	Julio	24	1992-07-24 00:00:00	07- 50 - 55
1992	Julio	24	1992-07-24 00:00:00	09- 60 - 65
1992	Julio	24	1992-07-24 00:00:00	09- 60 - 65
1992	Junio	12	1992-06-12 00:00:00	08- 55 - 60
1992	Junio	13	1992-06-13 00:00:00	09- 60 - 65
1992	Junio	14	1992-06-14 00:00:00	08- 55 - 60
1992	Junio	14	1992-06-14 00:00:00	08- 55 - 60
1992	Junio	14	1992-06-14 00:00:00	09- 60 - 65

Figura 10.20: Datos cargados en Power Pivot

El diseño de PowerView es sencillo y solo se arrastran los datos que se desean analizar.

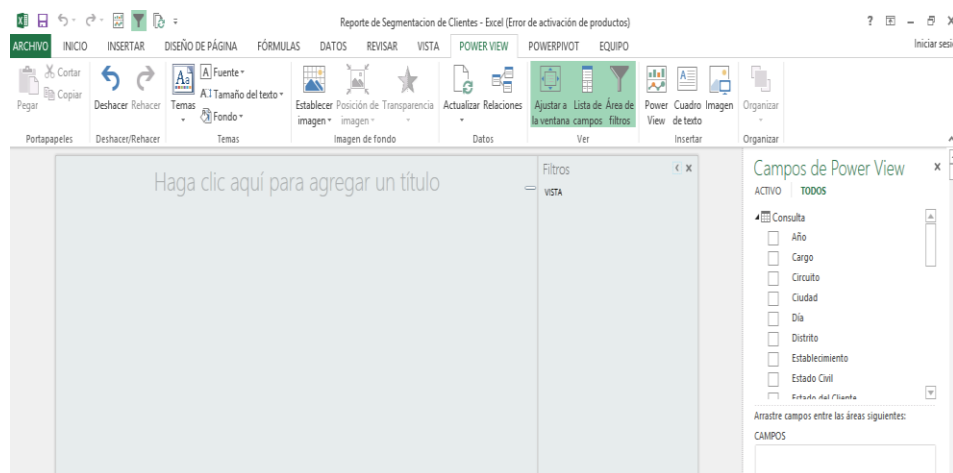


Figura 10.21: Pantalla de Diseño Power View

La implementación de reportes se realizó acorde a la información recopilada del levantamiento de requerimiento, los esquemas fueron diseñados con los involucrados, a continuación se detalla alguno de ellos:

Etiquetas de columna						
Abril						
Etiquetas de fila	Número Cliente Actualizado	Número Cliente con Regalo	Número Cliente con Firma	Número Cliente Activo	Número de Cliente	
AZUAY	36	4	11	6	44	
BOLIVAR	51	2	30	2	85	
CANAR	12	9	12	7	13	
AZOGUEZ	9	7	9	6	9	
AZOGUES	9	7	9	6	9	
DELEG	2	2	2	1	2	
DELEG	2	2	2	1	2	
LA TRONCAL	1	0	1	0	1	

Figura 10.22: Reporte de Clientes por Ubicación Geográfica

Uno de los reportes solicitados, fue el de reporte de clientes por ubicación geográfica, en el cual se observa la cantidad y porcentaje de clientes que tiene la empresa por zona y por provincias en el cual se visualiza donde se concentra el mayor número de clientes, dando una visión general de cómo está posicionado en cada provincia.

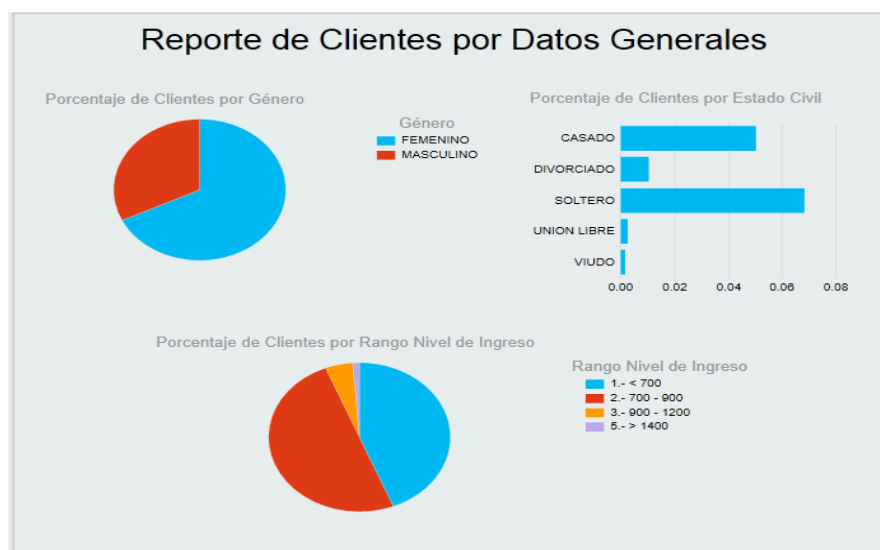


Figura 10.23: Reporte de Clientes por Datos Generales

El reporte de clientes por datos generales, da una visión amplia sobre el porcentaje de clientes por género, por estado civil y por el nivel de ingreso. Este informe ayuda a ver una amplia visión a la gerencia del porcentaje de clientes que tenemos en cada segmento y donde hay que seguir trabajando para captar nuevos clientes o retenerlos.

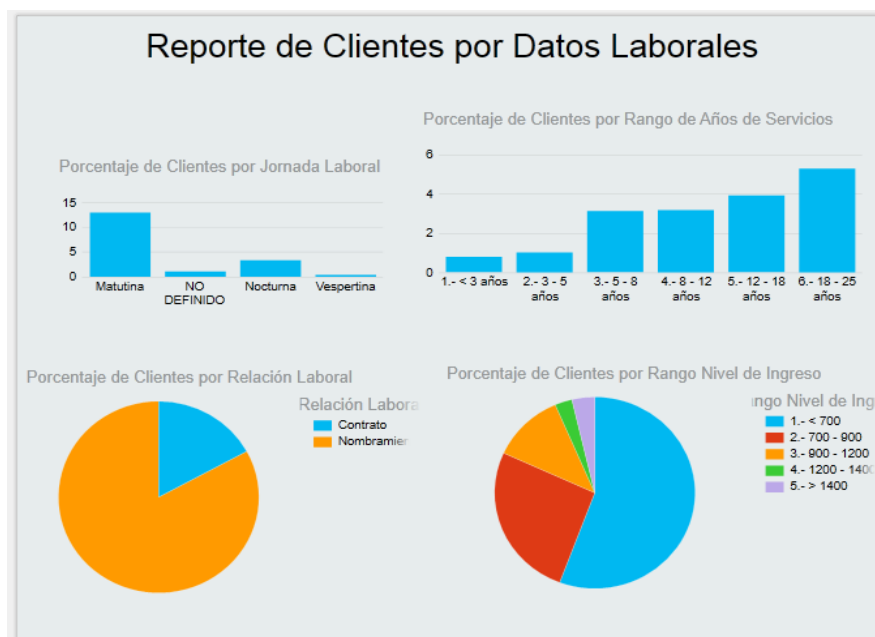


Figura 10.24: Reporte de Clientes por Datos Laborales

El reporte de clientes datos laborales, abarca la segmentación por Jornada Laboral, Rango de años de servicios, la relación laboral y el rango de nivel de ingreso salarial, que permite una visión amplia de dichos segmentos para la posterior toma de decisiones gerencial.

La reportería brinda una forma visual amigable y dinámica para el análisis de la información que facilita el análisis gerencial.

5.2 Plan de Pruebas

Para la realización de las pruebas del cubo de información, se implementó un ambiente de prueba denominado pre-producción, un ambiente similar al de producción con un mes de diferencia de información. Se establecieron en los modelos de ETL la conexión a dicho servidor designado para pruebas.

El ambiente de pruebas mantiene información al año 2018, lo que nos ayuda a una ejecución de fase de pruebas, esta actividad involucra el proceso de ETL conectada a desarrollo.

Pruebas unitarias

En dichas pruebas se realizó la ejecución de los modelos dimensionales, con información de ambiente de pre-producción, verificando los siguientes aspectos:

- Carga de datos a las tablas de dimensiones.
- Verificación de los datos de las tablas de dimensiones.
- Carga de datos de las tablas de hecho.
- Verificación de los datos de la tabla de hecho.
- Navegación de los niveles de jerarquías de dimensiones.
- Carga de datos del Tanque, Modelo Estrella y el cubo de Información.
- Consistencia de datos de todas las cargas.

A continuación se detalla la matriz de pruebas realizadas:

Tabla 15: Resultados de Pruebas Unitarias

Resultados de Casos de Uso para Pruebas Unitarias		
Tipo	Caso de Uso	Categoría
Funcional	Carga de Tanque (repositorio de unificación de datos).	Alto
Funcional	Verificación de conexión a SQL Server.	Alto
Funcional	Verificación de existencia de datos numéricos en campos alfabéticos.	Alto
Funcional	Verificación de campos alfanuméricos en campos de fecha.	Alto

Tipo	Caso de Uso	Categoría
Funcional	Verificación de consistencia de datos, verificando la existencia de campos nulos.	Alto
Tipo	Carga de Modelo Estrella	Categoría
Funcional	Verificación de conexión a SQL Server.	Alto
Funcional	Revisión de carga de datos de las tablas de dimensiones.	Alto
Funcional	Verificación de los datos de las tablas de dimensiones, que no se carguen datos duplicadas.	Alto
Funcional	Revisión de carga de datos de las tablas de hecho.	Alto
Funcional	Verificación de carga correcta de las tablas de hechos, sin duplicidad de datos, y todos los campos llenos.	Alto
Tipo	Cubo de Información	Categoría
Funcional	Verificación de conexión a SQL Server.	Alto
Funcional	Verificación de las jerarquías creadas en el cubo de información.	Alto
Funcional	Verificación de la coherencia de información reflejada en el cubo de información.	Alto
Funcional	Verificación de la estructura creada en el cubo de información.	Alto
Funcional	Verificación entre las medidas y dimensiones.	Alto
Funcional	Verificación de duplicidad de datos.	Alto

Pruebas con usuarios

Para las pruebas con usuarios se validó los siguientes aspectos:

- Tiempos de ejecución de informes.
- Resultados numéricos de informes.
- Consistencia con la data del sistema transaccional.
- Nivel de jerarquización de las dimensiones.

Para las pruebas se consideraron los siguientes puntos de vista técnicos:

Funcionalidad: Utilidad de la implementación para el usuario.

Rendimiento: Nivel de procesamiento del sistema frente a la demanda del requerimiento del usuario.

Las métricas para evaluar los siguientes parámetros son:

Alto: La funcionalidad y rendimiento son muy buena o exitosa.

Medio: La funcionalidad y rendimiento responde a un escenario normal, ni muy bueno, ni muy malo, es aceptable.

Bajo: La funcionalidad y el rendimiento a una alta demanda es muy bajo, se debe mejorar.

Nula: Funcionalidad nula, rendimiento inaceptable.

A continuación se detalla los resultados obtenidos de la fase de pruebas con los usuarios:

Tabla 16: Resultados de Pruebas con Usuarios

Resultados de Casos de Uso para Pruebas		
Tipo	Caso de Uso	Categoría
Funcional	Jerarquías de dimensiones que sean pertinentes.	Alto
Funcional	Datos de informes igual al de la base transaccional.	Alto
Funcional	Validación general de resultados de informes.	Alto
Rendimiento	Ejecución de informes por usuarios diferentes.	Alto
Rendimiento	Rendimiento de ejecución de informes.	Medio

El manejo de errores que pueda existir en el proceso diario, se maneja de la siguiente manera:

- Todos los días el administrador de base de datos realiza una revisión general de los jobs, que son los encargados de mandar a ejecutar todas las cargas que intervienen como carga principal, carga de modelo estrella y procesamiento de cubo de información, en el caso de existir un error en estos procesos, el administrador de base de datos comunica inmediatamente al encargado del mantenimiento y actualización del proceso del cubo de información.
- En el caso de existir un error de inconsistencia de datos no detectado en pruebas, el usuario afectado, debe comunicarse con el departamento de Sistemas, para indicar cuál es el inconveniente para que sea solucionado.

CAPÍTULO 6

ANÁLISIS DE RESULTADOS

Una vez terminado el proceso de diseño, desarrollo e implementación del cubo de información, detallamos en este capítulo el análisis de los resultados obtenidos con la culminación de este proyecto.

6.1 Información Integrada y Actualizada

Con el proceso del cubo de información de segmentación de clientes, se logró consolidar en una sola fuente de información todas las categorizaciones asociadas al cliente, de esta forma es más fácil de atender todas las solicitudes de información realizadas, dado a que la información se encuentra en un solo lugar. Adicional se efectuó el análisis de los datos respectivos, para la eliminación de inconsistencias existentes. Se organizó la información que se obtuvo en los diferentes niveles de detalle según los requerimientos de los usuarios, considerando solo datos puntuales para la generación de conocimiento del área

de negocio de Gestión de Clientes. De esta manera se tiene información actualizada en todo momento.

6.2 Revisión de resultados obtenidos

En la implementación del cubo de información para presentar los resultados que requiere el área de Servicio al Cliente de la empresa, se logró mejorar los tiempos de entrega de informes del equipo de supervisores a la Gerencia Administrativa.

A continuación se detallara los resultados obtenidos, enfocados en los problemas mencionados:

Tabla 17: Matriz de resultados Obtenidos (Problemas del Giro del Negocio)

Problemas	Mejora (%)	Antes	Después
Riesgos de error en los reportes	86.67	Información no estructurada en modelo de entidad relación.	Se crea modelo entidad relación.
Retrasos en la toma de decisión	86.67	Dependencia del personal de sistema para que obtenga la información y la presente en hoja de cálculo, según la necesidad del usuario.	Disponibilidad de la información del Datawarehouse, dependiendo únicamente del administrador de base de datos para actualizaciones y verificación del buen funcionamiento.

Problemas	Mejora (%)	Antes	Después
Costos Operativos	86.67	Se designaban un recurso para la generación de información que provocaba horas adicionales para la entrega oportuna, así como atrasos en tareas.	Se designa un recurso para la generación de información mediante la herramienta, se elimina las horas adicionales y atrasos de tareas.
No disponibilidad del recurso existente	86.67	Se depende del horario laboral disponible del personal de sistema para la entrega de la información al departamento de Gestión de clientes.	La información está disponible debido a que esta automatizada en el cubo de información, dependiendo únicamente de la supervisión del administrador de base de datos.

Revisando la tabla correspondiente encontramos que los problemas se presentaban en los tiempos de entrega de reportes, en los costos operativos que generan los recursos mal utilizados, y la información errónea que provocaba el reprocesamiento de informes. Ahora con el proceso actual, el riesgo de error en los reportes se ha reducido, al igual que ha mejorado los tiempos de retrasos en la entrega de reportes lo que mejora la toma de decisiones a tiempo, en los costos operativos también mejoró debido a que se disminuye la carga operativa, evitando la generación de horas de trabajo adicionales y retrasos de tareas.

Tabla 18: Matriz de resultados Obtenidos (Problemas Tecnológicos)

Problemas	Mejora (%)	Antes	Después
Falta de Seguridad	93.33	No existe control de acceso para la información.	Se crean grupos de usuarios con accesos al cubo.
Duplicidad de información	93.33	La misma información se encuentra en distintas tablas.	Se consolida la información en una sola fuente de información.
Reportes con información errónea	93.33	Por el nivel de rotación del personal de sistema no siempre se tenía expertos del negocio para la generación de los reportes, por lo que la información entregada en ocasiones era errónea.	Se eliminó la dependencia de personal sin conocimientos del negocio debido a la automatización del proceso de extracción de información.

Con respecto a los problemas tecnológicos tenemos la falta de seguridad de los archivos ya que no cuentan con control de acceso, adicional la duplicidad de información y reportes con información errónea es muy frecuente. Ahora con el proceso actual se maneja grupos de usuarios con accesos al cubo, se consolida la información en una sola fuente de datos, y se elimina la dependencia de personal sin conocimientos, debido a la automatización del proceso de extracción de información, dejando atrás el problema de rotación de personal.

Tabla 19: Matriz de Resultados Obtenidos (Nuevos Problemas Encontrados)

Problemas	Mejora (%)	Antes	Después
Personal no capacitado	86.67	Sobrecarga de trabajo en expertos del negocio.	Procesos automatizados (cubo de información).
No se da prioridades en las tareas y se atrasa la entrega de información	86.67	Se quitaba de proyectos al personal de sistema para solventar solicitudes de usuario.	Con el ETL y el datawarehouse se disminuye la carga laboral.
Información limitada del cliente	73.33	No se provee de todas las aristas de un perfil de clientes, por lo que el insumo de trabajo es escaso.	Se provee todas las perspectivas de información de clientes, pero depende del usuario su uso.
Requerimientos encolados	73.33	El tiempo de requerimiento se disminuye por atender los reportes solicitados por usuarios.	El cubo de información disminuye la carga laboral.

Adicional se detectaron nuevos problemas como sobrecarga de trabajo por personal no capacitado, atrasos en entrega de proyectos debido a la prioridad de tareas, no se proveía toda la información necesaria del cliente, todos estos problemas se presentan por el poco tiempo de atención a los requerimientos. Con la implementación del proyecto se logra la mejoría de estos puntos, debido a la automatización de procesos, permitiendo tener el tiempo necesario para atender requerimientos encolados, dar prioridades a proyectos de otras áreas y de capacitar al personal para que pueda atender mejor los procesos.

La encuesta fue realizada a los analistas y supervisor encargados de la entrega de informes al área de gestión de clientes.

6.3 Mejor uso de Recursos

El cubo de información de segmentación de clientes les brindó un mejor acceso a la información a todos los usuarios involucrados en el proceso con un lenguaje de negocio, ayudándoles a construir informes, realizar consultas y a efectuar un mejor análisis del negocio, incrementando su eficiencia con el aprovechamiento de los recursos tecnológicos de la empresa.

El problema de acceso a la información fue solucionado con la concentración de las diferentes fuentes de información de negocio en un solo entorno centralizado, obteniendo de esta manera rapidez en el análisis, consultas y menor tiempo de uso de la información.

La explotación de datos es más fácil y practica a través del datawarehouse aumentando una mayor eficacia del negocio, que no se logra con el uso de datos que proceden de aplicaciones operacionales, donde la información se obtiene de procesos independientes lo que hace que el proceso sea mucho más complejo.

El mayor beneficio de este proyecto es la optimización de tiempos, en cuanto a la visualización de reportes de información consolidada, y que está orientada para la toma de decisiones. Por otro lado, desde el punto de vista tecnológico, tenemos la optimización de recursos en cuanto a la entrega de reportes, ya que ahora la solicitud es al sistema dimensional y no a un recurso humano.

A continuación, se presenta la Tabla 20, que resume el alcance propuesto y los tiempos optimizados:

Tabla 20: Alcances Propuestos y tiempos optimizados

Área	Informes / Reportes	Tiempo Manual (min)	Tiempo automatizado (min)	Mejora (%)
Gestión de Clientes	Reporte de clientes por ubicación geográfica.	10	1	90.00
	Reporte de clientes por fecha de ingreso, por datos generales, laborales y empresariales.	15	1	93.33
	Reporte de clientes con datos actualizados por ubicación geográfica.	11	1	90.91
	Reporte de clientes con datos actualizados por fecha de ingreso, por datos generales, laborales y empresariales.	16	1	93.75
	Reporte de clientes activos por ubicación geográfica.	10	1	90.00
	Reporte de clientes activos por fecha de ingreso, por datos generales, laborales y empresariales.	15	1	93.33
	Reporte de clientes con premios o regalo por ubicación geográfica.	12	1	91.67
	Reporte de clientes con premios o regalo por fecha de ingreso, por datos generales, laborales y empresariales.	18	1	94.44
	Reporte de clientes con carta descuento por ubicación geográfica.	12	1	91.67
	Reporte de clientes con carta descuento por fecha de ingreso, por datos generales, laborales y empresariales.	17	1	94.12

La implementación del modelo dimensional para el área de Gestión de Clientes, proporcionó reportes de información consolidada e integrada que antes se realizaban de manera manual mediante hojas electrónicas, y que actualmente ayudan a la toma de decisiones debido a los criterios de negocio identificados.

A continuación, en la Tabla 21 se detallan algunos criterios que se consideraron para la toma de decisiones:

Tabla 21: Criterios para la Toma de Decisiones

Criterios para la toma de decisiones
Niveles de actualización de datos según el sector, sexo, género, edad y profesión, sueldos.
Niveles de actualización de datos bajo incentivos (obsequios, regalos).
Recuperación de clientes gracias a la actualización de datos.
Recuperación de clientes gracias a la firma de carta de autorización de descuento según los sectores.
Niveles de clientes según el sector, datos personal, profesionales y laborales.
Niveles de clientes activos por sector.

CONCLUSIONES Y RECOMENDACIONES

CONCLUSIONES

1. La disponibilidad y rapidez de la información de negocio, es de vital importancia para la empresa ya que se obtiene a tiempo para poder tomar decisiones de manera oportuna y poder tomar acciones en el momento indicado.
2. La presencia de una herramienta de inteligencia de negocios que brinde información de manera eficaz y desde diferentes perspectivas en unos minutos, permite realizar mejores análisis de la situación actual de la empresa.
3. La implementación de la solución propuesta, permite visualizar mejoras en los procesos, con respecto a los tiempos, ya que ahora no se depende del tiempo del persona de tecnología para la entrega de informes, obteniendo de esta manera un importante logro por las áreas involucradas que tienen disponibilidad de la información en todo momento para su análisis.

4. La calidad de la información se valora en el proyecto, debido a la eliminación de inconsistencia de datos, impactando directamente en la mejora de la toma de decisiones, reduciendo el sesgo de datos incorrectos.

RECOMENDACIONES

1. Usar los modelos desarrollados, de acuerdo al alcance como proyecto inicial para incursionar en el área de Inteligencia de Negocios.
2. Socializar el proyecto con otras áreas no definidas en el proyecto, a fin de mostrar la utilidad de la herramienta, evidenciando las ventajas y utilidad en la toma de decisiones.
3. Efectuar un proceso de medición del nivel de madurez que permite a la empresa ver el nivel de uso de la herramienta tecnológica, post-implementación a fin de verificar el acoplamiento del personal ante las soluciones de Inteligencia de Negocios.
4. Actualizar el software de base de datos, de SQL Server 2012 a 2016, capacitando al personal interno, para el mejor manejo y administración de la base de datos.
5. La administración del cubo de información debe quedar bajo custodia del área de base de datos.

BIBLIOGRAFÍA

- [1] Díaz, J. C. , Introducción al business intelligence. Editorial UOC, 2012.
- [2] ROSADO GOMEZ, A. A. y Rico Bautista, D. W. , «Inteligencia de negocios: Estado del arte», Sci. Tech., vol. 16, n.o 44, 2010.
- [3] Fernández, J. , Mayol, E. y Pastor, J. A. , «Agile business intelligence governance: Su justificación y presentación», 2008.
- [4] Ramos, S. , «Microsoft Business Intelligence: vea el cubo medio lleno», Microsoft Bus. Intell. , 2011.
- [5] Alcalá, G. C. , C. de Pablos Heredero, y Lozano, I. A. , «El Proceso de Implantación del Data Warehouse en la Organización: Análisis de un caso», Investig. Eur. Dir. Econ. Empresa, vol. 4, n.o 3, 1998, pp. 73–92.
- [6] Alexander, C. T. C. , «Elaboracion de un data mart para evidenciar el retraso académico en los alumnos de pregrado de la fii-unp», 2015.
- [7] Zabala A. y Jimena, F. , «Datamart para la toma de decisiones en los procesos de evaluación de la Universidad Regional Autónoma de Los Andes Uniandes», 2017.
- [8] El-Sappagh, S. H. A. , Hendawi, A. M. A. , y El Bastawissy, A. H. , «A proposed model for data warehouse ETL processes», J. King Saud Univ.-Comput. Inf. Sci., vol. 23, n.o 2, 2011, pp. 91–104.
- [9] De Lourdes R. A. M. , «Datawarehouse», 2015.

- [10] Alcívar Espinoza, L. M. , «Aplicación de un Cubo de información para la Recopilación del Área de Cobranzas vía Interfaz del Dispositivo Móvil», Universidad de Guayaquil. Facultad de Ciencias Matemáticas y Físicas. Carrera de Ingeniería en Sistemas Computacionales. Carrera de Ingeniería en Networking y Telecomunicaciones., 2011.
- [11] Cervantes Hernández, B. A. , «Propuesta de extracción automatizada de información para la predicción de ventas de la industria relojera», 2012.
- [12] Kimball, R. , Ross, M. , y others, «The data warehouse toolkit: the complete guide to dimensional modelling», Nachdr N. Y. Ua Wiley, 2002, pp. 1–447.
- [13] Llombart, Ó. A. , «BI: Inteligencia aplicada al negocio», DAA Conten. Digit. CMS-Spain ComEn Línea Madr. Dispon. En [Httpwww Eldiarioexterior Comconocimientodocs](http://www.eldiarioexterior.com/conocimiento/docs), 2003.
- [14] Villarreal Rosero, R. X. , «Estudio de metodologías de Data Warehouse para la implementación de repositorios de información para la toma de decisiones gerenciales.», B.S. thesis, 2014.
- [15] Arnoldo Márquez, A. G. , «Segmentación de clientes para empresas comercializadoras de alimentos para animales», 2007.

ANEXO - CRONOGRAMA

	TAREA	DURACIÓN	COMIENZO	FIN	PREDECESOR
1	Creación de Cubo de Información	154 días	11/12/2017	24/07/2018	
2	Inicialización	20 días	11/12/2017	9/1/2018	
3	Identificación y recopilación de información	15 días	11/12/2017	2/1/2018	
4	Identificar Fuente de Datos	3 días	3/1/2018	5/1/2018	3
5	Elaboración de Cronograma	2 días	8/1/2018	9/1/2018	4
6	Análisis	8 días	10/1/2018	19/01/2018	
7	Análisis de las nuevas dimensiones y medidas	5 días	10/1/2018	16/01/2018	5
8	Reunión con Especialistas	3 días	17/01/2018	19/01/2018	7
9	Diseño	32 días	22/01/2018	8/3/2018	
10	Diseño de arquitectura	5 días	22/01/2018	26/01/2018	8
11	Diseño de Extracción, Transformación y carga	10 días	29/01/2018	9/2/2018	10
12	Diseño de Dimensiones de cubo	5 días	14/02/2018	20/02/2018	11
13	Diseño de tabla de hechos del cubo	2 días	21/02/2018	22/02/2018	12
14	Diseño de reportes	10 días	23/02/2018	8/3/2018	13
15	Implementación	14 días	9/3/2018	28/03/2018	
16	Limpieza de información	8 días	9/3/2018	20/03/2018	14
17	Preparación de entorno DEV	6 días	21/03/2018	28/03/2018	16
18	Tablas Principales	12 días	29/03/2018	16/04/2018	
19	Creación o construcción del datawarehouse	10 días	29/03/2018	12/4/2018	17

	TAREA	DURACIÓN	COMIENZO	FIN	PREDECESOR
20	Creación de relaciones	1 día	13/04/2018	13/04/2018	19
21	Creación de índices	1 día	16/04/2018	16/04/2018	20
22	Procesos ETLs	19 días	17/04/2018	15/05/2018	
23	Definición de queries de extracción de información considerando las dimensiones definidas	12 días	17/04/2018	4/5/2018	21
24	Construcción de proceso ETL	7 días	7/5/2018	15/05/2018	23
25	Cubo de Información	11 días	16/05/2018	31/05/2018	
26	Creación de DataSource y vista	1 día	16/05/2018	16/05/2018	24
27	Creación del cubo	10 días	17/05/2018	31/05/2018	26
28	Reportes	8 días	1/6/2018	12/6/2018	
29	Creación de reportes	8 días	1/6/2018	12/6/2018	27
30	Pruebas Internas	16 días	13/06/2018	4/7/2018	
31	Preparación de Entorno de pruebas	5 días	13/06/2018	19/06/2018	29
32	Limpieza de datos y ordenamiento de información	3 días	20/06/2018	22/06/2018	31
33	Ejecución de pruebas integrales	5 días	25/06/2018	29/06/2018	32
34	Ajustes de procesos	3 días	2/7/2018	4/7/2018	33
35	Pruebas con Usuarios	13 días	5/7/2018	23/07/2018	
36	Preparación de Entorno de pruebas	2 días	5/7/2018	6/7/2018	34
37	Limpieza de datos y ordenamiento de información	3 días	9/7/2018	11/7/2018	36
38	Ejecución de pruebas integrales	5 días	12/7/2018	18/07/2018	37
39	Ajustes e Implementación de procesos	3 días	19/07/2018	23/07/2018	38
40	Finalización	1 día	24/07/2018	24/07/2018	
41	Cierre de documentos	1 día	24/07/2018	24/07/2018	39

ANEXO - ENCUESTA

Objetivo: Medir el impacto de la implementación del Cubo de Información en el Área de Gestión de Clientes.

Negocio

¿Usted considera que el proyecto ayudó a disminuir los errores en los reportes?

- a) Totalmente en desacuerdo
- b) En desacuerdo
- c) Ni de acuerdo, ni en desacuerdo
- d) De acuerdo
- e) Totalmente de acuerdo

¿Usted considera que el proyecto ayudó a disminuir el tiempo de demora en la toma de decisiones?

- a) Totalmente en desacuerdo
- b) En desacuerdo
- c) Ni de acuerdo, ni en desacuerdo
- d) De acuerdo

e) Totalmente de acuerdo

¿Está usted de acuerdo con que el proyecto ayudó a la eliminación de costos Operativos?

a) Totalmente en desacuerdo

b) En desacuerdo

c) Ni de acuerdo, ni en desacuerdo

d) De acuerdo

e) Totalmente de acuerdo

¿Está usted de acuerdo que el proyecto ayudó a eliminar la dependencia de recursos de tecnología para obtener reportes?

a) Totalmente en desacuerdo

b) En desacuerdo

c) Ni de acuerdo, ni en desacuerdo

d) De acuerdo

e) Totalmente de acuerdo

Problemas Tecnológicos

¿Considera que el proyecto contribuyó a mejorar la seguridad de la información?

- a) Totalmente en desacuerdo
- b) En desacuerdo
- c) Ni de acuerdo, ni en desacuerdo
- d) De acuerdo
- e) Totalmente de acuerdo

¿Considera que el proyecto contribuyó a la eliminación de duplicidad de datos?

- a) Totalmente en desacuerdo
- b) En desacuerdo
- c) Ni de acuerdo, ni en desacuerdo
- d) De acuerdo
- e) Totalmente de acuerdo

¿Considera que el proyecto contribuyó a mejorar la calidad de reportes, eliminando la cantidad de información errónea?

- a) Totalmente en desacuerdo
- b) En desacuerdo
- c) Ni de acuerdo, ni en desacuerdo
- d) De acuerdo
- e) Totalmente de acuerdo

Nuevos Problemas Encontrados

¿Considera que el proyecto ayudará a mitigar la falta de personal no capacitado?

- a) Totalmente en desacuerdo
- b) En desacuerdo
- c) Ni de acuerdo, ni en desacuerdo
- d) De acuerdo
- e) Totalmente de acuerdo

¿Considera que el proyecto ayuda a evitar que se sobrecargue de trabajo al personal?

- a) Totalmente en desacuerdo
- b) En desacuerdo
- c) Ni de acuerdo, ni en desacuerdo
- d) De acuerdo
- e) Totalmente de acuerdo

¿Considera que el proyecto contribuye a obtener un mejor perfil de los clientes?

- a) Totalmente en desacuerdo
- b) En desacuerdo
- c) Ni de acuerdo, ni en desacuerdo

- d) De acuerdo
- e) Totalmente de acuerdo

¿Considera que el proyecto ayuda a evitar atrasos en los requerimientos del negocio, evitando generar encolamientos?

- a) Totalmente en desacuerdo
- b) En desacuerdo
- c) Ni de acuerdo, ni en desacuerdo
- d) De acuerdo
- e) Totalmente de acuerdo