



UNIVERSIDAD CENTRAL DE VENEZUELA
FACULTAD DE CIENCIAS
ESCUELA DE COMPUTACIÓN
CENTRO DE INVESTIGACIÓN EN SISTEMAS DE INFORMACIÓN

Desarrollo de una solución de inteligencia de negocio para los indicadores del proceso de inspección de vehículos de empresas de seguros

Trabajo Especial de Grado presentado ante la ilustre
Universidad Central de Venezuela por

Br. Luis Hernández.

Para optar al título de Licenciado en Computación

Tutor:

Prof. Franky Uzcátegui

Caracas, 16/10/2018

Agradecimientos

Es impresionante como se puede resumir la vida académica que con tanto esfuerzo, resistencia, constancia, trabajo y dedicación llevamos a cabo para llegar a este momento que para todos aquellos quienes deseamos ser egresados de la casa que vence las sombras es tan anhelado, pero es un logro compartido y que va sobre los hombros y el apoyo de muchas otras personas que durante todo este período se han convertido en parte de la esencia que nos hace quienes somos y la vida que hemos construido.

Primero ante todo debo agradecerle a Dios, por ser mi apoyo espiritual durante toda la vida y llenarme de sabiduría en todos los momentos que lo he necesitado. En segundo lugar, a mi familia, mi mamá, mi abuela y mi papá que se convirtieron en los pilares fundamentales del hogar y que hicieron de mis hermanas Jhohana, Alejandra y mi persona los seres humanos y profesionales que se han forjado bajo su cauteloso cuidado. A mis hermanas Alejandra, Jhohana y Dixia por ser seres únicos y admirables, cada una en su forma única. Quiero agradecerle de igual forma a mi compañero de vida Nisson García por haber sido mi apoyo incondicional durante los últimos siete años, por creer en mi esfuerzo y haber pujado hasta el final para llegar a este logro que hoy comparto con ustedes.

Por último, y no menos importante, quiero agradecerle muy especialmente a mi tutor, el profesor Franky Uzcátegui por ser un excelente profesional, íntegro y consecuente guía quien dedicó mucho de su esfuerzo y aprendizajes al logro de este trabajo y sin el cual no habría sido posible llevar a cabo el mismo.

Universidad Central de Venezuela.
Facultad de Ciencias
Escuela de Computación
Centro de Investigación de Sistemas de Información

Desarrollo de una solución de inteligencia de negocio para los indicadores del proceso de inspección de vehículos de empresas de seguros

Autor: Luis Hernández.

Tutor: Prof. Franky Uzcátegui[Manager].

Fecha: 16-10-2018.

RESUMEN

La inspección de vehículos para la solicitud de pólizas o reporte de siniestros asociados al parque automotor de los propietarios de estos requiere el manejo de indicadores para funcionar como un proceso eficiente y eficaz que brinde a sus beneficiarios un servicio de calidad a la altura que demandan sus usuarios, adicionalmente se pretende brindar el manejo y control de los recursos que dispone la empresa para llevar a cabo exitosamente dicho proceso. Por esta razón se busca, con este trabajo, diseñar, establecer y desarrollar una solución con enfoque sobre inteligencia de negocios a fin de obtener indicadores que permitan dar el soporte necesario para ejecutar acciones en beneficio de la empresa. Por esto se presentan conceptos, metodologías y arquitecturas de desarrollo que permitan garantizar la ejecución y eventual obtención de los indicadores que brinden un apoyo oportuno a la toma de decisiones con la finalidad de manejar adecuadamente los recursos de los cuales dispone la organización. Esta solución de inteligencia de negocios está fundamentada en la metodología de desarrollo ascendente (Ralph Kimball) y el apoyo de herramientas de desarrollo para software libre (PostgreSQL, Pentaho BI Server (PBI), Pentaho Schema Workbench (PSW), Pentaho Data Integration (PDI) y Pentaho Report Designer (PRD))

Palabras Claves: Inteligencia de negocio, Metodología de Ralph Kimball, Póliza de seguros, Siniestro, Indicadores.

UNIVERSIDAD CENTRAL DE VENEZUELA
FACULTAD DE CIENCIAS
ESCUELA DE COMPUTACIÓN

ACTA

Quienes suscriben, miembros del jurado designado por el Consejo de la Escuela de Computación, para examinar el Trabajo Especial de Grado titulado "Desarrollo de una solución de inteligencia de negocio para los indicadores del proceso de inspección de vehículos de empresas de seguros" y presentado por el bachiller: Br. Luis Hernández titular de la cédula de identidad V-18.363.852, a los fines de optar al título de **Licenciado en Computación**, dejamos constancia de lo siguiente:

Leído como fue dicho trabajo, por cada uno de los miembros del jurado, se fijó el día 16 de OCTUBRE de 2018, a las 8:00 AM horas, para que el autor lo defendiera en forma pública, lo que estos hicieron en la Sala de Conferencias del Centro de Computación, mediante una presentación oral de su contenido, luego de lo cual respondió a las preguntas formuladas. Finalizada la defensa pública del Trabajo Especial de Grado, el jurado decidió aprobar con la nota de 19 puntos.

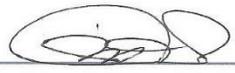
En fe de lo cual se levanta la presente Acta, en Caracas el día 16 de OCTUBRE de 2018.



Prof. Franky Uzcátegui
(Tutor)



Prof. Franklin Sandoval
(Jurado)



Prof. (a) Mercy Ospina
(Jurado)

Índice de Contenido

Introducción	i
CAPÍTULO 1 PROBLEMA DE INVESTIGACIÓN	1
1.1. Situación Actual	1
1.2. Planteamiento del Problema	2
1.3. Objetivos	4
1.3.1. Objetivo General	4
1.3.2. Objetivos Específicos	4
1.4. Solución Propuesta	4
1.5. Justificación	6
1.6. Alcance	8
Capítulo 2 Marco Conceptual	9
2.1. Dato, información y conocimiento	9
2.1.1. Dato	9
2.1.2. Información	9
2.1.3. Conocimiento	11
2.2. Sistemas de Información	12
2.2.1. Definición	12
2.2.2. Características de los Sistemas de Información	15
2.2.3. Tipos de Sistemas de Información	15
2.2.3.1. Enfocados a la Organización	15
2.2.3.2. Sistemas de Procesamiento de Transacciones (TPS)	17
2.2.3.3. Sistemas de Información Gerencial (MIS)	17
2.2.3.4. Sistemas de apoyo a la toma de decisiones (DSS)	18
2.2.3.5. Sistemas de apoyo a ejecutivos (ESS)	18
2.2.3.6. Enfocados a la mejora de la toma de decisiones	19
2.2.3.7. Sistemas de Procesamiento de Transacciones en Línea (OLTP)	19
2.2.3.8. Sistemas de Procesamiento Analítico en Línea (OLAP)	20
2.2.3.9. OLTP vs OLAP	21
2.2.4. Objetivos de un Sistema de Información	22
2.3. Bases de datos	23
2.3.1. Bases de datos Relacional	24
2.4. Almacén de Datos	26
2.4.1. Objetivos de un Almacén de datos	27
2.4.2. Características de un Almacén de Datos	29
2.4.3. Bodega de Datos (Data Mart)	30
2.4.4. Diferencias entre una Bodega de Datos y un Almacén de Datos	31
2.4.5. Modelo Dimensional	33
2.4.6. Ventajas y Desventajas de los Almacenes de Datos	41
2.5. Inteligencia de Negocios	42

2.5.1. Características de una Solución de Inteligencia de Negocio	43
2.5.2. Funciones de una Solución de Inteligencia de Negocio	44
2.5.3. Arquitectura de una Solución de Inteligencia de Negocio	44
2.6. Tecnologías y Herramientas para la construcción de una Solución de Inteligencia de Negocio.....	46
2.6.1. Pentaho Business Intelligence.....	47
2.6.2. Oracle Business Intelligence (OBI)	52
2.6.3. Tableau.....	58
2.6.4. Comparación entra las herramientas de inteligencia de Negocios.....	61
Capítulo 3 Marco Metodológico	73
3.1. Metodologías de desarrollo según Ralph Kimball (Bottom-up)	73
3.1.1. Planificación del proyecto	75
3.1.2. Definición de los Requerimientos del negocio	76
3.1.3. Diseño de la arquitectura	77
3.1.4. Modelo Dimensional.....	78
3.1.5. Diseño de la aplicación BI.....	78
3.1.6. Selección de productos e instalación	78
3.1.7. Diseño físico	79
3.1.8. Diseño y desarrollo del ETL.....	79
3.1.9. Desarrollo de la aplicación BI	80
3.1.10. Implementación	80
3.1.11. Expansión.....	80
3.1.12. Mantenimiento	81
capítulo 4 Marco Aplicativo	82
4.1. Fases del Proyecto	82
4.1.1. Planificación del Proyecto	82
4.1.2. Definición de los Requerimientos del Negocio	83
4.1.3. Diseño de la Arquitectura	86
4.1.4. Selección de Productos e Instalación.....	88
4.1.5. Modelado Dimensional	90
4.1.5.1. Definir el Proceso de Negocio	90
4.1.5.2. Identificar el Nivel de Granularidad	90
4.1.5.3. Definir las Dimensiones	92
4.1.5.4. Identificación de los Hechos y las Tablas de Hechos	93
4.1.6. Diseño Físico	94
4.1.7. Diseño y Desarrollo de Procesos ETL.....	98
4.1.8. Especificaciones de las Aplicaciones Analíticas.....	104
4.1.9. Desarrollo de las Aplicaciones Analíticas	106
4.1.10. Implementación	107
Conclusiones y Recomendaciones	116
Bibliografía	118

Índice de Figuras

FIGURA 1 ARQUITECTURA PARA LA SOLUCIÓN DE INTELIGENCIA DE NEGOCIO PROPUESTA	6
FIGURA 2 PROCESO DE TRANSFORMACIÓN DE DATOS EN INFORMACIÓN (STAIR & REYNOLDS) ...	10
FIGURA 3 SISTEMAS DE INFORMACIÓN (LAUDON & LAUDON)	12
FIGURA 4 ACTIVIDADES DE UN SISTEMA DE INFORMACIÓN (STAIR & REYNOLDS)	14
FIGURA 5 TIPOS DE SISTEMAS DE INFORMACIÓN (LAUDON & LAUDON)	16
FIGURA 6 EJEMPLO DE UN DIAGRAMA ENTIDAD-RELACIÓN (SILBERSCHATZ)	25
FIGURA 7 EJEMPLO DE UN DIAGRAMA DE BASE DE DATOS RELACIONAL (SILBERSCHATZ)	26
FIGURA 8 EJEMPLO DE UNA JERARQUÍA GEOGRÁFICA	36
FIGURA 9 EJEMPLO DE UN ESQUEMA ESTRELLA	38
FIGURA 10 EJEMPLO DE UN ESQUEMA COPO DE NIEVE	38
FIGURA 11 EJEMPLO DE UN ESQUEMA CONSTELACIÓN	39
FIGURA 12 ARQUITECTURA DE UNA SOLUCIÓN DE INTELIGENCIA DE NEGOCIOS	45
FIGURA 13 INTERFAZ DE PENTAHO REPORT DESIGNER	50
FIGURA 14 EJEMPLO DE DASHBOARDS	51
FIGURA 15 INTERFAZ DE DATA INTEGRATOR	52
FIGURA 16 ARQUITECTURA DE ORACLE BUSINESS INTELLIGENCE.....	53
FIGURA 17 INTERFAZ DE ORACLE WAREHOUSE BUILDER.....	55
FIGURA 18 INTERFAZ INTERACTIVA DEL DASHBOARDS.....	56
FIGURA 19 INTERFAZ DE OBI ANSWERS.....	57
FIGURA 20 EJEMPLO DE INTERFAZ BI PUBLISHER	58
FIGURA 21 EJEMPLO DE INTERFAZ TABLEAU DESKTOP	60
FIGURA 22 EJEMPLO DE INTERFAZ TABLEAU SERVER.....	61
FIGURA 23 BUSINESS DIMENSIONAL LIFECYCLE DIAGRAM (KIMBALL & ROSS, 2002)	75
FIGURA 24 ARQUITECTURA DE SOLUCIÓN DE INTELIGENCIA DE NEGOCIOS	86
FIGURA 25 RELACIONES JERÁRQUICAS DEL MODELO DIMENSIONAL EN LA SOLUCIÓN PLANTEADA	93
FIGURA 26 MODELO DIMENSIONAL PROPUESTO	94
FIGURA 27 DISEÑO FÍSICO DE LA DIMENSIÓN TIEMPO	95
FIGURA 28 DISEÑO FÍSICO DE LA DIMENSIÓN CENTRO DE INSPECCIÓN	96
FIGURA 29 DISEÑO FÍSICO DE LA DIMENSIÓN ESTATUS	96
FIGURA 30 DISEÑO FÍSICO DE LA DIMENSIÓN PERITO	96
FIGURA 31 DISEÑO FÍSICO DE LA DIMENSIÓN SOLICITUD DE INSPECCIÓN.....	97
FIGURA 32 DISEÑO FÍSICO DE LA DIMENSIÓN TIPO DE INSPECCIÓN.....	97
FIGURA 33 MODELO FÍSICO DEL ALMACÉN DE DATOS PROPUESTO COMO SOLUCIÓN	98
FIGURA 34 DIRECTORIO DE ARCHIVOS.....	101
FIGURA 35 ETL PARA CARGAR LA DIMENSIÓN CENTRO DE INSPECCIÓN.....	101
FIGURA 36 ETL PARA CARGAR LA DIMENSIÓN ESTATUS	102
FIGURA 37 ETL PARA CARGAR LA DIMENSIÓN PERITO	102
FIGURA 38 ETL PARA CARGAR LA DIMENSIÓN SOLICITUD DE INSPECCIÓN	103
FIGURA 39 ETL PARA CARGAR LA DIMENSIÓN TIEMPO	103
FIGURA 40 ETL PARA CARGAR LA DIMENSIÓN TIPO DE INSPECCIÓN.....	103

FIGURA 41 ETL PARA CARGAR LA TABLA DE HECHOS	104
FIGURA 42 TRABAJO (JOB) PARA AUTOMATIZAR TODOS LOS ETL'S.....	104
FIGURA 43 ESQUEMA DE DATOS DESARROLLADO CON PENTAHO SCHEMA WORKBENCH	107
FIGURA 44 VISTA DINÁMICA DE LA CANTIDAD DE SOLICITUDES POR TIPO	108
FIGURA 45 VISTA DINÁMICA DE LA CANTIDAD DE SOLICITUDES POR ESTATUS	109
FIGURA 46 VISTA DINÁMICA DE LA CANTIDAD DE SOLICITUDES POR PERITO	109

Índice de Tablas

TABLA 1 OLTP VS OLAP.....	22
TABLA 2 ALMACÉN DE DATOS VS BODEGA DE DATOS	31
TABLA 3 COMPARACIÓN ENTRE LAS HERRAMIENTAS DE INTELIGENCIA DE NEGOCIOS	61
TABLA 4 COMPARACIÓN ENTRE LAS METODOLOGÍAS DE DESARROLLO DE UN ALMACÉN DE DATOS	74
TABLA 5 ACTIVIDADES ASOCIADAS A LA IMPLEMENTACIÓN DE LA SOLUCIÓN DE INTELIGENCIA DE NEGOCIOS	83
TABLA 6 HERRAMIENTAS PARA EL DESARROLLO DE LA SOLUCIÓN DE INTELIGENCIA DE NEGOCIOS.....	90
TABLA 7 DIMENSIONES ASOCIADAS AL MODELO DIMENSIONAL EN LA SOLUCIÓN PLANTEADA	92
TABLA 8 PRINCIPALES PASOS UTILIZADOS EN LOS PROCESOS DE ETL	99

INTRODUCCIÓN

Una empresa de seguros es una organización especializada en prestar una extensa gama de servicios y productos destinados a la protección y bienestar de los clientes que pertenecen o forman parte de estas, a través de la atención y satisfacción de sus necesidades en temas de garantía de los consumidores que se encuentran afiliados a estos. Es por esta razón que, progresivamente, las empresas de seguros aumentan y diversifican los tipos de servicios o productos que pretenden ofrecer a sus clientes. Como consecuencia de esto las empresas de seguros se han visto envueltas en un crecimiento progresivo que ha originado un manejo masivo de datos e información producida por todos los procesos que conviven o ejecutan simultáneamente dentro de estas y dada la naturaleza de todo este flujo de información dichas organizaciones se han visto en la necesidad de recurrir, en su mayoría, a la automatización para poder manejarlos.

Uno de estos procesos es la solicitud de inspección o peritaje de un vehículo con la finalidad de reportar algún evento ocurrido con el automóvil o bien, solicitar la afiliación de algún servicio. Sin embargo, la automatización es solo uno de los objetivos que se pretenden alcanzar en la actualidad ya que al tratarse de recursos finitos se pretende maximizar la ganancia obtenida a partir de estos a través de cualquier proceso de optimización que haga evolucionar al proceso original. El proceso actual se subdivide en otros dos conformados por un proceso de solicitud, en donde el cliente completa un formulario de datos personales y adicionalmente solicita hora, fecha y locación donde desea ser atendido, y un proceso de atención, donde el solicitante se dirige al sitio donde fue acordada, de acuerdo con el cliente, la previa cita para la inspección del vehículo.

Parte del proceso de inspección se realiza a través de un sistema automatizado vía web y otra parte dentro de un centro de inspección de vehículos que no es más que la unidad responsable de supervisar y velar por la prestación de dicho servicio de peritaje a los clientes afiliados o no a la compañía de seguros. En este ente se encuentra el personal encargado de realizar las verificaciones asociadas al proceso de peritaje, recibir, evaluar, avalar la autorización de

suscripción de póliza, llamar a los clientes (para realizar la verificación), avalar las indemnizaciones asociadas al reporte de un siniestro, coordinar y ejecutar la comprobación.

Es por esta razón que, al tratarse de un sistema transaccional de grandes cantidades de datos que son procesados y resguardados para eventualmente convertirse en información útil para la organización, se pretende realizar el debido control y seguimiento de estos a través de un sistema de análisis dedicado para tal propósito, ya que al no contar con esto se generan problemas para realizar los estudios analíticos y de monitoreo correspondientes en función de los datos que se recopilan trayendo como consecuencia la afección sobre la toma de decisiones. Gracias a este escenario se hace necesario implementar una solución de inteligencia de negocios como alternativa tecnológica que permita manejar la información desde la fase de extracción, depuración y transformación de los datos hasta la exploración y visualización de la información a través de herramientas fáciles de usar que brinden el apoyo necesario y suficiente al proceso de toma de decisiones.

Es por ello, que este Trabajo Especial de Grado (TEG) se enfoca fundamentalmente en el desarrollo de una Solución de Inteligencia de Negocio debido a la necesidad de obtener indicadores de control y seguimiento del proceso de inspección de vehículos de una empresa de seguros.

La estructura de este trabajo especial de grado es la siguiente:

El Capítulo I – problema de investigación: Contiene el planteamiento del problema que dio origen al presente trabajo, el objetivo general, los objetivos específicos, la solución propuesta y la justificación de la solución que se presenta.

Capítulo II – Marco conceptual: Se presentan los fundamentos teóricos y principales conceptos que servirán de base para dar soporte al desarrollo del presente T.E.G. y se indicará algunas herramientas para generar soluciones de Inteligencia de Negocio existentes en el mercado actual. En primer lugar, se exponen las nociones básicas de los sistemas de información, los tipos, características y objetivos. En segundo lugar, los conceptos relacionados a los

almacenes de datos (Data Warehouse), seguido de los relacionados a las herramientas de BI, con la particularidad de la herramienta Pentaho BI Suite.

Capítulo III – Marco Metodológico: En este capítulo se describen las distintas fases que componen la metodología propuesta por Ralph Kimball para el desarrollo de una solución de inteligencia de negocio, la cual se utilizará como guía para la realización de esta solución.

Capítulo IV – Marco Aplicativo: En este capítulo se muestran las actividades realizadas en las distintas fases del desarrollo del software, según la metodología escogida: Planificación del Proyecto, Definición de los Requerimientos del Negocio, Selección de las Herramientas, Diseño Técnico de la Arquitectura, Modelo Dimensional, Diseño Físico, Diseño de Procesos de Extracción Transformación y Carga (ETL), Especificación de la Aplicación de Usuario Final y el Despliegue.

Capítulo V – Conclusiones y Recomendaciones: Se presentan las conclusiones sobre el trabajo logrado, como parte del desarrollo de dicho sistema de inteligencia de negocio. Adicionalmente, incluye las recomendaciones para trabajos que se presenten en un futuro.

CAPÍTULO 1

PROBLEMA DE INVESTIGACIÓN

En el presente capítulo se detallan los puntos asociados con el problema planteado en el documento de tesis, adicionalmente se presenta la justificación que da origen a dicha investigación. A continuación, se presenta la solución que pretende dar respuesta al presente problema, se especifica el objetivo general del proyecto acompañado de los objetivos específicos y para finalizar se determina el alcance y limitaciones de dicho trabajo de investigación y desarrollo.

1.1. Situación Actual

Hoy en día alrededor del mundo al igual que en Venezuela las empresas que se encargan de prestar servicios al público tienden a desarrollar y aplicar procesos automatizados para suplir las demandas de los clientes en todo momento y tener una forma adicional de acceso a estos. Pero ya no se trata solo de brindar un servicio automatizado que supla en todo momento las necesidades de los clientes que integran la organización, hoy en día se trata, adicionalmente, de prestar un servicio de calidad, a la altura de lo que demandan los clientes, con estándares eficientes y que le permita a la organización hacer un uso adecuado de los recursos con los que cuenta. Esto aplica para cualquier tipo de servicio que requiera de una reservación previa, entre los cuales se puede tener como ejemplo las reservas que se emplean para el apartado de hoteles, pasajes de viaje, funciones cinematográficas, citas médicas, entre otras.

En la actualidad, la mayor parte de estos procesos se encuentran automatizados a partir del uso de tecnologías web que nos permiten desarrollar plataformas especializadas para hacer la captura de los datos. Sin embargo, ya no es solo de vital importancia para cualquier empresa el poder facilitar a sus clientes el acceso a estos servicios ofrecidos, adicionalmente, se pretende que estos sean usados en forma óptima y los datos recopilados provenientes de estos sistemas transaccionales nos brindan una oportunidad de llevar un análisis profundo para lograr este objetivo.

Es importante resaltar que, para el desarrollo de este tipo de proyectos, es indispensable conocer el proceso automatizado que se pretende abordar ya que de ello dependerá el manejo eficiente de los recursos asociados y se podrá solventar en forma eficaz la aparición de cualquier inconveniente en función de los indicadores que se pretenden generar. Por esta razón, se procedió a estudiar uno de los casos en donde se evidencie un gran nivel de concurrencia al momento de realizar la solicitud de un servicio. Para ello se desarrolló un proyecto dirigido a la empresa seguros Horizonte con la finalidad de crear un sistema de inteligencia de negocios para el proceso de trámite para la inspección de vehículos relacionado con el área de seguros automotriz. Este es un proceso que se encuentra automatizado y forma parte de un sistema transaccional en donde los clientes pueden determinar el estado actual de sus vehículos y verificar si se encuentran en condiciones para realizar una suscripción a una póliza de seguros o efectuar una indemnización en caso de un siniestro. Sin embargo, se pretende generar indicadores de gestión en función de los datos obtenidos a partir de este sistema para darle soporte a las decisiones que se tomen con la finalidad de brindar un servicio no solo eficaz pero además eficiente y ofrecerle a la organización una mejora en el uso de los recursos que esta maneja.

1.2. Planteamiento del Problema

En la actualidad las organizaciones no solo buscan la automatización de los procesos que se llevan a cabo dentro de estas, se plantea adicionalmente la optimización de estos y el manejo de los recursos al tratarse de elementos cuantificables y finitos por lo cual se debe maximizar cualquier beneficio que provean y hacer un uso inteligente de los mismos.

Los sistemas transaccionales tienden a ser una solución que proporciona la captura masiva de los datos dentro de las organizaciones, sin embargo, forman parte de una solución ajustada al contexto sobre el cual fueron diseñados y eventualmente creados. Por esta razón, en el proceso de captura de los datos, tiende a dificultarse la obtención de la información gracias a los excesivos tiempos de espera tanto para la extracción como el procesamiento de esta, y de igual forma para reagrupar e interpretar la información a usuarios de la alta gerencia dentro de la organización. Todo esto acarrea la generación de problemas orientados al análisis y el monitoreo en el contexto de los datos que se recolectan y por tanto pueden afectar el proceso de toma de decisiones.

Sin embargo, las empresas de seguros no están exentas de este escenario ya que dentro de ellas también conviven procesos que necesitan del adecuado control y seguimiento con la finalidad de brindar una oportunidad de mejora en cualquiera de las etapas sobre las cuales se desenvuelvan estos, bien sea al comienzo, intermedio o final de cualquiera de dichos procesos. Ahora bien, si deseamos lograr esto debemos tener en cuenta que existen muchas formas, instrumentos y métodos para lograr dicho control y seguimiento en cualquiera de estas etapas, pero puede llegar a ser un proceso lento y engorroso o no estar automatizado.

De igual forma, en cuanto a la inspección de vehículos se trata debe existir una correcta planificación entre los centros de inspección, la cantidad de peritos en estos y los horarios disponibles para atender a los clientes con la finalidad de cumplir con los tiempos de atención y ajustarse a los horarios previamente reservados por estos. Es por esta razón que de mantener un constante control y seguimiento sobre dicho proceso se podrán evaluar todas las posibles mejoras que disminuyan o eliminen las posibles pérdidas de recursos para, de esta forma, ser usados eficientemente.

Como consecuencia del control y seguimiento del proceso de inspección de vehículos podemos llegar a toparnos con una gran cantidad de inconvenientes que van desde prestarle atención a los detalles que involucran a este hasta el manejo de grandes cantidades de datos para lograr una adaptación al contexto actual sobre el cual se pretende desenvolver este proyecto. Seguido de esto se pretende crear los respectivos informes que contengan la información sobre los indicadores en función de los datos que hayan sido recopilados, sin embargo, de no ser un proceso que se encuentre de igual forma automatizado puede ocurrir un retraso muy pronunciado, más aún para quien es importante dicha información y venir acompañada de un sinnúmero de errores producto del manejo humano.

Adicionalmente, y gracias a que, tanto los recursos como el alcance del proyecto en cuestión son limitados se hace indispensable la implementación de indicadores con la finalidad de incrementar la eficiencia en líneas generales y la distribución de los recursos al igual que llevar el respectivo seguimiento para lograr los objetivos y metas planteadas, respetar las limitaciones del mismo y hacer un correcto uso de este. Dicho sistema analítico se encontrará

en un espacio centralizado para el posterior análisis de los datos, sobre los cuales se podrán generar los respectivos informes estadísticos con el objetivo de evaluar cualquier asunto relacionado con la solicitud de citas para la inspección de vehículos del parque automotor de los clientes o no de la compañía de seguros haciendo uso de herramientas tecnológicas que faciliten la elaboración de reportes e indicadores que den el respectivo soporte a la toma de decisiones.

1.3. Objetivos

1.3.1. Objetivo General

Desarrollar una solución de inteligencia de negocio para los indicadores del proceso de inspección de vehículos de empresas de seguros

1.3.2. Objetivos Específicos

- Analizar los requerimientos de indicadores
- Diseñar el modelo dimensional y el panel de control
- Desarrollar el almacén de datos
- Desarrollar los indicadores
- Realizar las pruebas funcionales, no funcionales e integradas

1.4. Solución Propuesta

Como solución al problema planteado previamente se establece, la implementación de un modelo dimensional y almacén de datos históricos con base en el sistema transaccional del cual dispone la empresa de seguros con la finalidad de obtener los respectivos indicadores de control y seguimiento al proceso de inspección de vehículos, los cuales se podrán visualizar a través de reportes, un panel de control y un cubo de información a fin de que los usuarios puedan obtener de forma rápida y flexible la información relevante y en forma oportuna que brinde el respectivo soporte al proceso de toma de decisiones, para de esta forma evaluar la gestión de citas para la inspección de vehículos dentro de la empresa de seguros.

Por esta razón tanto el panel de control como los reportes y el cubo serán construidos con herramientas de inteligencia de negocios, ya que facilitan la emisión de reportes y construcción de indicadores ajustados a la medida del proyecto para, eventualmente, emitir la publicación de los mismos en forma automatizada y que estos vayan dirigidos u orientados a usuarios de la alta gerencia de acuerdo con el perfil de acceso que estos poseen. De igual forma al hacer uso de herramientas de inteligencia de negocios podemos obtener contenidos dinámicos como el uso de filtros, tablas, búsquedas y diversos criterios de selección que no requieren de la atención de personal especializado. Adicionalmente, es importante destacar que la emisión de reportes puede darse en una importante gama de formatos que facilitan su presentación para quienes serán eventualmente entregados. Esta solución cuenta con una serie de elementos que podemos contemplar en la figura 1 de la siguiente forma:

- La fuente de datos provenientes del sistema transaccional de la organización y constituidas básicamente por bases de datos, las cuales son administradas por el Sistema Manejador de Base de Datos PostgreSQL.
- Un conjunto de procesos de Extracción, Transformación y Carga (ETL o ETC) que permiten extraer los datos desde las fuentes de datos (base de datos transaccional), realizar procesos de depuración, transformación y operaciones sobre los datos y por último cargarlos en el almacén de datos.
- El Almacén de Datos, basado en un esquema estrella como forma de implementación del modelo dimensional utiliza el Sistema Manejador de Base de Datos PostgreSQL para su administración.
- Un Portal Web, que contiene el tablero de control, un grupo de reportes y el cubo de información, los cuales permitirán la visualización de los indicadores e información solicitados por la alta gerencia.



FIGURA 1 ARQUITECTURA PARA LA SOLUCIÓN DE INTELIGENCIA DE NEGOCIO PROPUESTA

Utilizaremos como plataforma de tecnológica el sistema manejador de base de datos (SMBD) PostgreSQL para el almacenamiento de los datos dentro del almacén de datos, adicionalmente las herramientas Pentaho Data Integration para la implementación de los distintos procesos de extracción, transformación y carga de los datos desde la fuente de datos al almacén de datos y las herramientas Pentaho BI server, Pentaho Schema Workbench y Pentaho Report Designer tanto para la manejo, presentación y estudio de los datos como para la construcción del panel de control, los diferentes reportes y el cubo de información.

1.5. Justificación

Vivimos en la actualidad un período donde la información se ha vuelto significativamente un elemento clave que permite a cualquier organización o ente lograr los objetivos que se plantea como misión y visión. Debido a esto, quienes forman parte de dichas organizaciones demandan que el acceso a la información sea fácil y rápida con la intención de apoyarse en esto para tomar decisiones en beneficio de dicho ente en períodos de tiempo bastante cortos a fin de mantener un desempeño eficaz y óptimo dentro de dichas organizaciones.

Por esta razón, contamos con un limitado tiempo para realizar el debido análisis a toda la información y a esto se le suma el hecho de tomar las decisiones adecuadas en un entorno que se sufre constantes cambios, por lo cual se deben tomar acciones que estén ajustadas en

tiempo y asertividad ya que estas serán el norte que determinara las acciones que se tomaran en un futuro, por lo tanto es de vital importancia contar con información histórica y correctamente procesada de todo el proceso de solicitud de citas para la inspección de vehículos que nos permita realizar una comparación entre los distintos períodos además de identificar posibles patrones de comportamiento.

Ya que, a diario se producen grandes cantidades de solicitudes para la inspección de vehículos, quienes se encuentren encargados de realizar el monitoreo, control y seguimiento sobre dicho proceso podrán realizar todas estas tareas en forma constante y adicionalmente identificar las áreas involucradas en el proceso que requieran de una mayor atención o prioridad para ser atendidas o mejoradas, adicionalmente se tiene acceso inmediato a información de vital importancia que permite tomar decisiones efectivas para ejecutar correctamente todo lo que se ha planificado con anterioridad y superar cualquier eventualidad que se pueda presentar con la finalidad de optimizar el manejo y uso de los recursos.

En adición a lo antes mencionado, el hecho de tener indicadores de gestión permite lograr una mejor administración dentro del proceso de negocio, debido a que, quien es responsable del control y seguimiento de este no usará más del tiempo necesario para realizar el análisis de cada uno de estos, y podrá observar reflejados en cada uno de estos indicadores las diferentes áreas que requieran de mayor atención, sumado a esto la emisión de reportes permite realizar un estudio muchos más detallado de toda la información con la finalidad de buscar una solución que se ajuste mucho mejor aún al problema que se presente. Al ser reportes generados en forma automática brindan información en tiempo real a quienes sea útil y de interés, permitiendo que el análisis sea efectivo y ajustado al momento en que se genera.

Por tal motivo la implementación de una solución de inteligencia de negocios es una alternativa que permite la toma de decisiones en forma acertada a cualquier nivel dentro de la organización sobre la cual se implementa, a través del manejo de la información desde su extracción, depuración y transformación como datos hasta su exploración y distribución como información gracias al uso de herramientas de fácil uso para los usuarios, a fin de compensar sus necesidades para poder realizar un aporte dentro del proceso de solicitud de citas para la inspección de vehículos del parque automotor de los clientes afiliados o no.

Es importante resaltar que las soluciones de inteligencia de negocio les permiten a las personas encargadas del proceso ser más productivas debido a lo siguiente:

- Reduce el tiempo requerido para recolectar toda la información necesaria del negocio al evitar la ejecución de consultas directas sobre los sistemas transaccionales.
- Descubrir información no evidente a partir de los análisis realizados sobre los datos operativos.
- Evitamos que se manipulen manualmente los datos, hecho que puede causar que el resultado de los análisis efectuados sea incorrecto.
- Se le permite al usuario final realizar análisis rápidos, personalizados y fiables, en los que se obtenga la información que es requerida en un momento dado a través de panel de control.

1.6. Alcance

La intención de crear una solución de inteligencia de negocios tiene un camino orientado al diseño y despliegue de los indicadores relacionados al proceso solicitud de citas para la inspección de vehículos, con un particular enfoque sobre el control y seguimiento que involucra tanto al proceso de solicitud como al proceso de atención de los clientes. Con dicha solución quien se responsabilice de llevar el seguimiento eventualmente podrá dar el soporte adecuado para tomar las acciones necesarias que permitan corregir alguna falla sobre el manejo adecuado de los recursos que se encuentran disponibles, apoyándose en toda la información que genera dicha aplicación, la cual será mostrada tanto en el panel de control, los informe emitidos por esta y de igual forma el cubo de información, todo esto soportado por los datos que el sistema transaccional genera como fuente para ser cargados en nuestro almacén de datos.

CAPÍTULO 2

MARCO CONCEPTUAL

2.1. Dato, información y conocimiento

2.1.1. Dato

De acuerdo con Cohen & Asín (2009) el dato "puede ser un número, una palabra o una imagen; en la que el habla cotidiana no se utiliza el plural "datos" que en términos formales son la materia prima para la producción de información". Ahora bien, para los autores Davenport & Prusak, (1998), "Los datos son un conjunto de hechos discretos y objetivos sobre los acontecimientos. En un Contexto organizacional, los datos son más usualmente descritos como registros estructurados de transacciones". Y para los autores Stair & Reynolds (2010) "Un dato consiste en un hecho aislado, por ejemplo, un número de empleado, las horas totales trabajadas a la semana, los números de parte en un inventario o las órdenes de venta"

Si bien es cierto, los datos representan solo parte de todo aquello que nos rodea, percepciones de la realidad o una forma de reflejar todos los hechos reales, sin embargo, en su forma más esencial no proporcionan un análisis profundo o definición de lo que pretenden representar ya que son unidades fundamentales y adicionalmente esenciales en la construcción de la información.

2.1.2. Información

Suele suceder que los términos dato e información se manejan como iguales sin percatarse del error cometido. Esto, debido a que, como ya lo mencionamos en el concepto previo, un dato es una representación atómica que puede presentarse como un número, palabra o imagen y que conforma la esencia de la información.

Para los autores Cohen & Asín (2009) "La información, por su parte, son datos que en un contexto dado tienen significado para alguien." Es de vital importancia el manejo de la información ya que este da pie a la toma de decisiones y "en los negocios hay una característica relevante de la información: su calidad. Para entender que significa información de "calidad" es indispensable considerar tres dimensiones: la dimensión de tiempo, la de contenido y la de forma", a continuación, explicamos con detenimiento la perspectiva de dichos autores.

- "Dimensión de tiempo. La información debe estar disponible cuando se necesita, estar actualizada, proveerse con la periodicidad requerida y representar el pasado, el presente y el futuro."
- "Dimensión de contenido. La información no debe contener errores, debe ser relevante respecto a lo que se analiza, ser completa, concisa, interna y externa, con un enfoque amplio o centrado y medir el desempeño."
- "Dimensión de forma. La información se debe integrar en una forma sencilla, sea detallada o en forma de resumen, debe estar ordenada con base en cierto criterio, y se puede presentar en formatos diferentes: tablas, gráficas, listas y, finalmente, aparecer en diferentes medios: papel, medios digitales, etcétera."

Por otra parte, para Stair & Reynolds (2010) "la información es un conjunto de hechos organizados de tal manera que poseen un valor adicional más allá del valor que se les puede atribuir como hechos individuales".

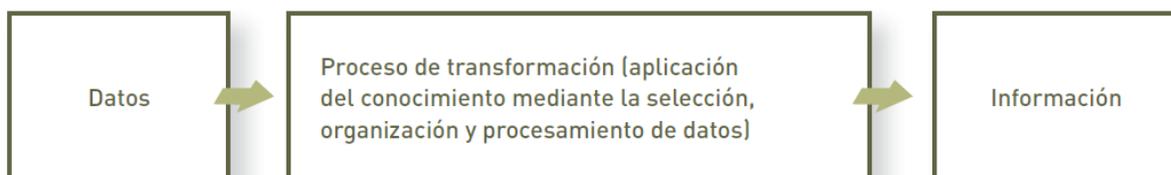


FIGURA 2 PROCESO DE TRANSFORMACIÓN DE DATOS EN INFORMACIÓN (STAIR & REYNOLDS)

Sin embargo, para los autores Davenport & Prusak (1998) existe una perspectiva diferente en la que "como muchos investigadores que han estudiado la información, lo describiremos como un mensaje, generalmente en forma de un documento, comunicación audible o visible. Como cualquier mensaje, tiene un remitente y un receptor. La información está destinada a cambiar la forma en que el receptor percibe algo, para tener un impacto en su juicio y el comportamiento. Es la información lo que marca la diferencia. La palabra "informar" originalmente significaba "dar forma a" y la información está destinada a dar forma a la persona que lo recibe, para hacer alguna diferencia en su perspectiva o visión. Hablando estrictamente, entonces, el receptor, no el remitente, decide si el mensaje que recibe es realmente información, es decir, si realmente le informa. Un memo completo de divagaciones no conectadas puede considerarse "información" por el escritor, pero juzgado como ruido por el receptor. El único mensaje que puede comunicarse con éxito es uno no intencional sobre la calidad de la inteligencia o del juicio del remitente."

2.1.3. Conocimiento

De acuerdo con Stair & Reynolds (2010) "El conocimiento es la comprensión de un conjunto de información y de las formas en que ésta puede convertirse en algo útil para realizar una tarea específica o tomar una decisión". Es irrefutable que el conocimiento es un invaluable recurso e indispensable para el desarrollo de cualquier organización y para Laudon & Laudon (2012) "el conocimiento es un conjunto de conceptos, experiencias y perspicacias que proporcionan un marco de trabajo para crear, evaluar y utilizar información". Ineludiblemente el conocimiento solo se transforma en tal hecho cuando se ha formado parte de una vivencia o experiencia que conlleva a esto, tener conocimiento.

Ahora bien, para los autores Cohen & Asín (2009) "en contraste con los conceptos de datos e información, el conocimiento sugiere alternativas de acción o guías de actuación específicas relacionadas con la información", en consecuencia, este último concepto nos sugiere que el conocimiento no es más que una forma de actuar o un conjunto de pautas que están estrechamente relacionadas a la información que se tiene respecto de un hecho, evento o registro.

2.2. Sistemas de Información

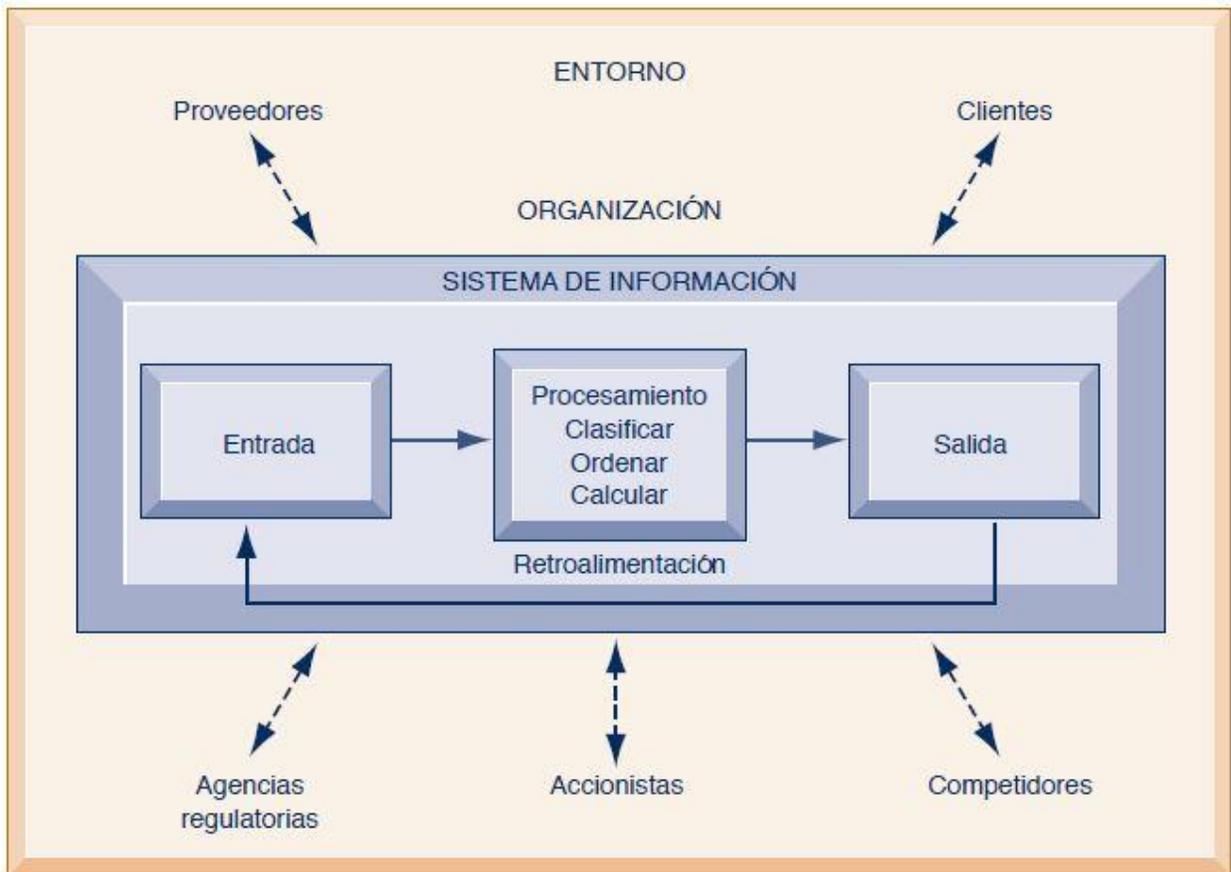


FIGURA 3 SISTEMAS DE INFORMACIÓN (LAUDON & LAUDON)

2.2.1. Definición

Actualmente la información se ha convertido en el recurso de mayor relevancia para las organizaciones. Cuando pensamos sobre el porqué del uso de sistemas que permitan el procesamiento dentro de estas, la respuesta parece bastante obvia, para registrar los eventos que se van suscitando. Sin embargo, existe un motivo aún más importante y no tan obvio, y es que la información puede usarse para respaldar la toma de decisiones. Sin importar el tamaño de una empresa u organización, la conducción exitosa de la misma estará dramáticamente influenciada por la precisión de sus registros y la adopción de decisiones acertadas.

De acuerdo con Laudon & Laudon en su décimo segunda edición (2012) "Podemos plantear la definición técnica de un sistema de información como un conjunto de componentes interrelacionados que recolectan (o recuperan), procesan, almacenan y distribuyen información para apoyar los procesos de toma de decisiones y de control en una organización".

Apoyados en este concepto podemos expresar que un sistema de información actualmente conforma una parte de vital importancia para cualquier organización u empresa si uno de sus objetivos es conocer su entorno a fin de generar una verdadera y sensata ventaja competitiva que le permita posicionarse estratégicamente.

De acuerdo con Cohen & Asín (2009) en su quinta edición un sistema de información "no necesariamente incluye equipo electrónico", sin embargo, realiza cuatro actividades básicas incluyendo:

- **Entrada de información:** Es el ingreso de los datos al sistema de información. Las entradas pueden ser manuales, proporcionadas por el usuario o automáticas, datos o información que provienen o se toman de otros sistemas o módulos, lo que se denomina interfaces automáticas.

Los típicos dispositivos de entrada de datos a los sistemas de información son las cintas magnéticas, las unidades de disquete, los lectores de código de barras, las plumas ópticas, el escáner, los monitores sensibles al tacto, los lectores de CD-ROM, los lectores de DVD, dispositivos USB, el mouse y el teclado.

- **Almacenamiento de información:** Es una de las capacidades más importantes de una computadora, ya que permite al sistema recordar la información guardada en la sesión anterior. Esta información se almacena en estructuras de información denominadas archivos, en su versión simple, y bases de datos, en su modalidad compleja.

Ejemplo de estos son los dispositivos de almacenamiento como los discos magnéticos o duros, los discos flexibles o disquetes, los discos compactos (CD-ROM), los discos de alta capacidad (Zip), los discos de video (DVD) y los USB. Existen otros mecanismos de almacenamiento, que más adelante se detallan.

- **Procesamiento de información:** Es la capacidad del sistema de información para efectuar cálculos de acuerdo con una secuencia de operaciones preestablecida. Estos cálculos trabajan con datos de recién ingreso o con datos ya almacenados. Esta característica de los sistemas permite la transformación de datos fuente en información útil, lo que posibilita, entre otras cosas, que quien toma decisiones genere una proyección financiera a partir de los datos que contiene un estado de resultados o un balance general de un año base.
- **Salida de información:** Es la capacidad de un sistema de información para convertir la información procesada o los datos de entrada en información para el exterior. Las típicas unidades de salida son las impresoras, disquetes, cintas magnéticas, la voz, los graficadores y los plotters, entre otros. Es importante aclarar que la salida de un sistema de información puede constituir la entrada a otro sistema de información o módulo. En este caso también existe una interfaz automática de salida. Por ejemplo, el sistema de control de clientes, al que alimenta el sistema de facturación, tiene una interfaz automática de salida con el sistema de contabilidad, que genera las pólizas contables de los movimientos de los clientes.

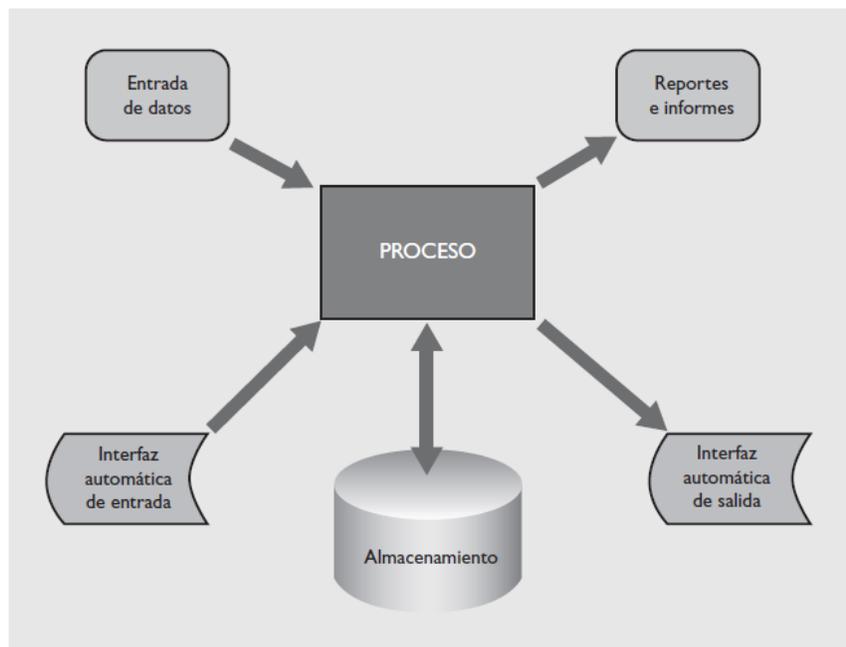


FIGURA 4 ACTIVIDADES DE UN SISTEMA DE INFORMACIÓN (STAIR & REYNOLDS)

2.2.2. Características de los Sistemas de Información

Para Cohen & Asín (2009) un sistema de información "eficaz" debe contar con las siguientes características:

- Su estructura y funciones deben ser coherentes con la estrategia competitiva de la empresa.
- La planeación y el diseño de un sistema de información debe servir a los objetivos estratégicos de la empresa.
- El sistema de información se debe planear, diseñar e implantar para guardar un equilibrio con el resto de los sistemas integrantes de la infraestructura de la organización.
- Debe evaluar y revisarse desde una perspectiva amplia, es decir, con la colaboración de la alta gerencia de la compañía.

2.2.3. Tipos de Sistemas de Información

2.2.3.1. Enfocados a la Organización

Es inimaginable pensar en una organización sin ningún tipo de jerarquía posible, tan arraigado está este concepto que el significado de organización por si sola brinda una idea de ésta ya que de acuerdo al diccionario de la real academia española se define como "Disposición, arreglo u orden" o "Acción y efecto de organizar u organizarse", de igual forma es natural pensar que no todas son iguales ya que están definidas de acuerdo a las necesidades que necesitan suplir cada una de ellas, bien sea de acuerdo a su estructuras, procesos o políticas.

Sin importar el tamaño de estas organizaciones siempre es posible visualizar algún tipo de estructura jerárquica conformada por diferentes niveles y para cada una de ellas existe un sistema de información que permite brindar el soporte adecuado en las actividades que desempeñan diariamente.

De acuerdo con la figura N° 5, las organizaciones pueden estructurarse en 4 posibles niveles donde se hace intuitivo pensar que a medida que escalamos a un nivel superior las decisiones suelen ser menos estructuradas, la autoridad y responsabilidades en cualquier empresa u organización se suelen constituir de igual forma jerarquizada donde los niveles superiores están compuestos por trabajadores gerenciales, profesionales y técnicos mientras que el nivel base se compone de personal operacional.

La administración de nivel superior se encarga de la toma de decisiones estratégicas abarcando un amplio alcance sobre los bienes y servicios que ésta presta, y asegurando su bienestar económico, la gerencia de nivel medio u operacional se encuentra encargada de ejecutar los planes de acción, mesas de trabajo que demanda la administración de nivel superior e inspección de las actividades cotidianas dentro de la organización, por último, pero no por eso menos importante, tenemos a los trabajadores del conocimiento, los datos o servicios, constituido por profesionales del área en cuestión.

Ésta pudiese constituir una forma de organización dentro de cualquier empresa sin embargo no es la única, a continuación, presentaremos una alternativa que muestra cómo se puede manejar la información concerniente a la toma de decisiones en los diferentes niveles que plantea y que tipos de sistemas de información generan la base del conocimiento y dan soporte a estas.



FIGURA 5 TIPOS DE SISTEMAS DE INFORMACIÓN (LAUDON & LAUDON)

2.2.3.2. Sistemas de Procesamiento de Transacciones (TPS)

Los sistemas de procesamiento de transacciones (TPS, Transaction Processing Systems) son sistemas de información computarizada creados para procesar grandes cantidades de datos relacionadas con transacciones rutinarias de negocios, como las nóminas y los inventarios. Un TPS elimina el fastidio que representa la realización de transacciones operativas necesarias y reduce el tiempo que una vez fue requerido para llevarlas a cabo de manera manual, aunque los usuarios aún tienen que capturar datos en los sistemas computarizados. (Kendall & Kendall, 2005).

Los gerentes operacionales necesitan sistemas que lleven el registro de las actividades y transacciones elementales de la organización, como ventas, recibos, depósitos en efectivo, nóminas, decisiones de créditos y el flujo de materiales en una fábrica. Los Sistemas de Procesamiento de Transacciones (TPS) proveen este tipo de información. (Laudon & Laudon)

2.2.3.3. Sistemas de Información Gerencial (MIS)

La gerencia de nivel medio necesita sistemas que les ayuden con las actividades de supervisión, control, administración, y toma de decisiones proporcionando informes semanales, mensuales y anuales sobre el desempeño de la organización. Los sistemas de información gerencial (MIS) utilizan los datos provenientes de los sistemas de procesamiento de transacciones (TPS) y responden a esta pregunta: ¿Funcionan bien las cosas? (Laudon & Laudon, 2012).

Los sistemas de información gerencial (MIS, Management Information Systems] no reemplazan a los sistemas de procesamiento de transacciones, más bien, incluyen el procesamiento de transacciones. Los MIS son sistemas de información computarizados cuyo propósito es contribuir a la correcta interacción entre los usuarios y las computadoras. Debido a que requieren que los usuarios, el software [los programas de cómputo] y el hardware (las computadoras, impresoras, etc.), funcionen de manera coordinada, los sistemas de información gerencial dan apoyo a un espectro de tareas organizacionales mucho más amplio que los sistemas de procesamiento de transacciones, como el análisis y la toma de decisiones. (Kendall & Kendall, 2005).

2.2.3.4. Sistemas de apoyo a la toma de decisiones (DSS)

Los sistemas de apoyo a la toma de decisiones (DSS, Decisión Support Systems) constituyen una clase de alto nivel de sistemas de información computarizada. Los DSS coinciden con los sistemas de información gerencial en que ambos dependen de una base de datos para abastecerse de datos. Sin embargo, difieren en que el DSS pone énfasis en el apoyo a la toma de decisiones en todas sus fases, aunque la decisión definitiva es responsabilidad exclusiva del encargado de tomarla. Los sistemas de apoyo a la toma de decisiones se ajustan más al gusto de la persona o grupo que los utiliza que a los sistemas de información gerencial tradicionales. En ocasiones se hace referencia a ellos como sistemas que se enfocan en la inteligencia de negocios. (Kendall & Kendall, 2005).

Los sistemas de soporte de decisiones (DSS) brindan apoyo a la gerencia intermedia en la toma de decisiones en ocasiones donde los problemas son únicos, cambiantes y poco fáciles de anticipar. Los DSS responden a preguntas como estas: ¿Cuál sería el impacto en los calendarios de producción si tuviéramos que duplicar las ventas de diciembre? ¿Qué pasaría con nuestro rendimiento sobre la inversión si el calendario de producción de una fábrica se retrasara seis meses? Se abastecen tanto de datos provenientes de los TPS y MIS como fuentes externas. Estos sistemas condensan grandes cantidades de información de tal forma que su análisis sea sencillo de entender para los encargados de tomar las decisiones. (Laudon y Laudon, 2012).

2.2.3.5. Sistemas de apoyo a ejecutivos (ESS)

Estos sistemas ayudan a la gerencia de alto nivel en a las actividades de planificación a largo plazo de las organizaciones ayudan a los directores de las organizaciones enfrentar y resolver aspectos estratégicos; dando así los cambios necesarios del entorno externo de la organización. Los ESS generan herramientas graficas dándole la posibilidad a los ejecutivos de comprender sus entornos y resolver problemas no estructurados y con visión a largo plazo, sin embargo, al igual que los DSS dependen de la información producida por los TPS y los MIS. (Kendall & Kendall, 2005).

Cuando los ejecutivos recurren a la computadora, por lo general lo hacen en busca de métodos que los auxilien en la toma de decisiones de nivel estratégico. Los sistemas de apoyo a

ejecutivos (ESS, Executive Support Systems) ayudan a estos últimos a organizar sus actividades relacionadas con el entorno externo mediante herramientas gráficas y de comunicaciones, que por lo general se encuentran en salas de juntas o en oficinas corporativas personales. A pesar de que los ESS dependen de la información producida por los TPS y los MIS, ayudan a los usuarios a resolver problemas de toma de decisiones no estructuradas, que no tienen una aplicación específica, mediante la creación de un entorno que contribuye a pensar en problemas estratégicos de una manera bien informada. Los ESS amplían y apoyan las capacidades de los ejecutivos al darles la posibilidad de comprender sus entornos. (Kendall & Kendall, 2005).

2.2.3.6. Enfocados a la mejora de la toma de decisiones

2.2.3.7. Sistemas de Procesamiento de Transacciones en Línea (OLTP)

Los sistemas OLTP del acrónimo en inglés Online Transaction Processing o procesamiento de transacciones en línea, de acuerdo con Ponniah (2001, p.12) "son sistemas diseñados para el procesamiento de transacciones atómicas que se utilizan para dirigir el negocio cotidiano del día a día de la empresa", por lo que se puede inferir que manejan una carga grande de lecturas y escrituras de datos. "Los sistemas operacionales hacen girar las ruedas del negocio. Apoyan los procesos básicos de negocio de la empresa", por lo tanto, abarcan múltiples sistemas operacionales dentro de la organización donde se encuentren implementados. "Estos sistemas normalmente obtienen los datos en la base de datos. Cada transacción procesa información sobre una sola entidad, como una sola orden, una sola factura o un solo cliente". De acuerdo con Cohen & Asín (2009) son sistemas de información encargados de automatizar los procesos operativos dentro de una organización, "ya que su función principal consiste en procesar transacciones, tales como pagos, cobros, pólizas, entradas, salidas, etc."

Adicionalmente se puede catalogar según Kimball como "la descripción original de todas las actividades y sistemas asociados con la introducción fiable de datos en una base de datos. Se utiliza con mayor frecuencia con referencia a bases de datos relacionales, aunque OLTP se puede utilizar genéricamente para describir cualquier entorno de procesamiento de transacciones".

2.2.3.8. Sistemas de Procesamiento Analítico en Línea (OLAP)

Los sistemas OLAP del acrónimo en inglés On-Line Analytical Processing, es según Ponniah (2001) es una categoría de tecnología de software que permite a los analistas, administradores y ejecutivos ganar percepción dentro de la data a partir de rapidez, consistencia y acceso interactivo en una amplia variedad de posibles vistas de la información que ha sido transformada desde una data pura para reflejar la dimensionalidad de la compañía de forma que se entendida por el usuario". De acuerdo con Cohen & Asín (2009) son sistemas complementarios a los sistemas transaccionales y se encuentran enfocados a la mejora de la toma de decisiones dentro de una organización.

Los sistemas OLAP están diseñados para la media y alta gerencia, que requieren una manipulación de la data generada por los sistemas OLTP. Por tanto, los sistemas OLAP podemos verlo como una evolución de los sistemas OLTP, pasando del registro de operaciones a la utilización de los eventos de la organización para la toma de decisiones. Debido a que son sistemas que soportan consultas complejas para realizar agregación y análisis sobre los datos, son sistemas donde solo se realizan operaciones de lectura (salvo en las fases donde se carga la data al sistema).

OLAP soporta el análisis de datos multidimensional, el cual permite a los usuarios ver los mismos datos de distintas formas mediante el uso de varias dimensiones. Cada aspecto de información —producto, precios, costo, región o periodo de tiempo— representa una dimensión distinta. Así, un gerente de productos podría usar una herramienta de análisis de datos multidimensional para saber cuántas arandelas se vendieron en el Este en junio, cómo se compara esa cifra con la del mes anterior y con la de junio del año anterior, y cómo se compara con el pronóstico de ventas. OLAP permite a los usuarios obtener respuestas en línea a las preguntas ad hoc tales como éstas en un periodo de tiempo bastante corto, incluso cuando los datos se almacenan en bases de datos muy grandes, como las cifras de ventas de varios años.

En aras de encontrar una forma de representar la información proveniente de las diversas fuentes en una organización, los sistemas OLAP hacen uso de los cubos OLAP, que son la

forma por excelencia de representar en dimensiones los datos contenidos en las bases de datos fuente de los sistemas transaccionales.

Características de los sistemas OLAP:

- El acceso a los datos suele ser de sólo lectura ya que la acción más común es la consulta, con muy pocas inserciones, actualizaciones o eliminaciones.
- Los datos se estructuran según las áreas de negocio o preferencias de estudios sobre los datos, y los formatos de los datos están integrados de manera uniforme en toda la organización.
- El historial de datos es a largo plazo, hablamos de períodos usualmente extensos, meses o años ya que el estudio que se planea realizar es sobre la evolución de estos eventos en un largo período de tiempo.
- Las bases de datos OLAP se suelen alimentar de información procedente de los sistemas operacionales existentes, mediante un proceso de extracción, transformación y carga (ETL por sus siglas en ingles).

2.2.3.9. OLTP vs OLAP

Los sistemas OLTP se caracterizan por la creación y uso de muchos usuarios, actualizaciones o procesos de recuperación de registros atómicos. Por tal motivo, las bases de datos orientadas a dichos sistemas (OLTP) se encuentran perfectamente adaptadas para la captura masiva de dichas transacciones. Ahora bien, los sistemas OLAP son usados por analistas y gerentes quienes frecuentemente solicitan y sugieren vistas a un nivel superior de los datos. Las bases de datos orientadas a los sistemas OLAP normalmente se actualizan por lotes, a menudo de múltiples fuentes, y proporcionan un respaldo analítico poderoso para las aplicaciones de múltiples usuarios. Por tanto, las bases de datos orientadas a los sistemas OLAP se encuentran perfectamente adaptadas para el análisis.

En la tabla que a continuación se presenta (Tabla 1) podemos apreciar las diferencias más enmarcadas entre los sistemas OLTP y los sistemas OLAP.

Criterio	OLTP	OLAP
Fuente de datos	Datos operacionales, OLTP's son la fuente original de los datos.	Consolidación de datos, datos OLAP proviene de las diversas bases de datos OLTP
Finalidad de los datos	Controlar y ejecutar las tareas fundamentales del negocio	Ayudar con la planificación, resolución de problemas y toma de decisiones
Inserciones y Actualizaciones	inserciones y actualizaciones cortas y rápidas iniciadas por los usuarios finales	trabajos periódicos de larga ejecución por lotes de actualización de los datos
Consultas	Consultas estandarizadas y simples que devuelven relativamente pocos registros	Consultas complejas que implican agregaciones
Velocidad de Procesamiento	Generalmente muy rápido	En función de la cantidad de datos involucrados; lotes actualizaciones de datos y consultas complejas pueden tomar muchas horas; velocidad de las consultas se puede mejorar mediante la creación de índices
Requisitos de Espacio	Puede ser relativamente pequeña si se archiva datos históricos	Más grande debido a la existencia de estructuras de agregación y datos de historial; requiere más índices que OLTP
Base de Datos	Altamente normalizada con muchas tablas	Típicamente desnormalizada con menos tablas; uso de esquema dimensional
Seguridad y Recuperación	Copia de seguridad puntualmente; datos operativos es fundamental para manejar el negocio, pérdida de datos que pueda provocar una pérdida monetaria significativa y la responsabilidad legal	En lugar de copias de seguridad periódicas, algunos entornos pueden considerar simplemente volver a cargar los datos OLTP como un método de recuperación

TABLA 1 OLTP VS OLAP

2.2.4. Objetivos de un Sistema de Información

De acuerdo con los autores Cohen & Asín (2009) en la actualidad los sistemas de información cumplen tres objetivos básicos y fundamentales dentro de las organizaciones que se resumen en los siguientes puntos:

- Automatizan los procesos operativos: gracias a la automatización de los sistemas de información, mejor conocidos como sistemas transaccionales, los procesos operativos dentro de una organización pueden gestionar una gran cantidad de transacciones como, por ejemplo: pagos, cobros, pólizas, entradas, salidas, etc.
- Proporcionan información que sirve de apoyo en el proceso de toma de decisiones: El complemento de los sistemas transaccionales son los sistemas enfocados a la mejora de la toma de decisiones (OLAP), mejor conocidos como sistemas analíticos en línea, cuya función es la manipulación de la información con el fin de apoyar y fundamentar la toma de decisiones. Estos sistemas se agrupan a partir de dos funciones: los que se limitan al análisis de la situación y los que además del análisis de la situación hacen una recomendación sobre la decisión a tomar.
- Logran ventajas competitivas a través de su implantación y uso: Otro tipo de sistemas de acuerdo con su uso y objetivos que cumplen, son los sistemas estratégicos, que buscan desarrollar en las organizaciones ventajas competitivas, a través del uso de la tecnología de información (TI por sus siglas en ingles).

2.3. Bases de datos

De acuerdo con Elmasri y Navathe (2007) "una base de datos es una colección de datos relacionados. Con la palabra datos nos referimos a los hechos (datos) conocidos que se pueden grabar y que tienen un significado implícito. Por ejemplo, piense en los nombres, números de teléfono y direcciones de las personas que conoce. Puede tener todos estos datos grabados en un libro de direcciones indexado o los puede tener almacenados en el disco duro de un computador mediante una aplicación como Microsoft Access o Excel. Esta colección de datos relacionados con un significado implícito es una base de datos."

Según Cohen & Asín (2009) "una base de datos se define como una serie de datos organizados y relacionados entre sí, los cuales son recolectados y explotados por los sistemas de información de una empresa o negocio. También es una colección de datos con las siguientes características: conjunto de datos correlacionados; construida para un propósito específico; tiene algunos usos de interés especial para ciertas personas; representa algún aspecto del

mundo real, por tanto, los cambios en ese aspecto se deben reflejar en la base de datos; los datos se almacenan de tal forma que son independientes de los programas que los usan.”

Por datos entendemos hechos conocidos que pueden registrarse y tienen un significado implícito, datos que deben de ser persistentes y que están diseñados para satisfacer las necesidades de información de una organización, por lo que dejan de ser propiedad de un departamento y pasan a ser un recurso corporativo compartido.

2.3.1. Bases de datos Relacional

Es por excelencia el modelo usado e implementado como parte de los sistemas transaccionales. Basado en una percepción abstracta del mundo real que consta de una colección de objetos primordiales conocidos como “entidades”, cuya primordial característica en el modelo relacional erradica en su representación como una tabla acompañada de un conjunto de atributos que la describen, y un conjunto de “relaciones”, que no es más que una asociación entre una o varias entidades, entre estos objetos de acuerdo con Silberschatz (Fundamentos de bases de datos, 2002). La estructura lógica de una base de datos se puede representar gráficamente mediante un diagrama entidad-relación (E-R), que consta de un conjunto de elementos:

- Rectángulos, que representan conjuntos de entidades.
- Elipses, que representan atributos.
- Rombo, que representan relaciones entre conjuntos de entidades.
- Líneas, que unen los atributos con los conjuntos de entidades y Cada elemento se etiqueta con la entidad o relación que representa.

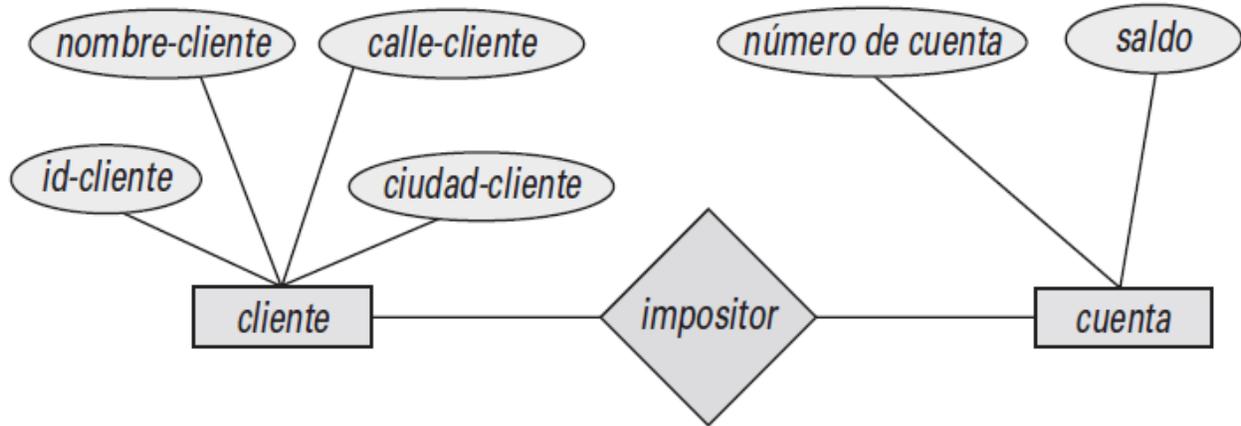


FIGURA 6 EJEMPLO DE UN DIAGRAMA ENTIDAD-RELACIÓN (SILBERSCHATZ)

Ahora bien, En el modelo relacional se utiliza un grupo de tablas para representar los datos y las relaciones entre ellos. Cada tabla está compuesta por varias columnas, y cada columna tiene un nombre único. En la Figura 6 se presenta un ejemplo de base de datos relacional consistente en seis tablas para una tienda de videos: la primera muestra los clientes de dicha tienda, la segunda los préstamos de las películas, la tercera cantidad de ejemplares por película, la cuarta las películas y en la quinta y sexta el actor principal y el director de la película respectivamente.

De acuerdo con el autor Silberschatz "el modelo relacional es un ejemplo de un modelo basado en registros. Los modelos basados en registros se denominan así porque la base de datos se estructura en registros de formato fijo de varios tipos. Cada tabla contiene registros de un tipo particular. Cada tipo de registro define un número fijo de campos, o atributos. Las columnas de la tabla corresponden a los atributos del tipo de registro".

Algunas de las características de las bases de datos relacionales son las siguientes:

- Una base de datos se compone de varias tablas o relaciones.
- No pueden existir dos tablas con el mismo nombre ni registro.
- Cada tabla es a su vez un conjunto de campos (columnas) y registros (filas).
- La relación entre una tabla padre y una tabla hijo se lleva a cabo por medio de las claves primarias y claves foráneas (o ajenas).
- Las claves primarias son la clave principal de un registro dentro de una tabla y estas deben cumplir con la integridad de datos.
- Las claves foráneas se colocan en la tabla hija, contienen el mismo valor que la clave primaria del registro padre; por medio de estas se hacen las formas relacionales.

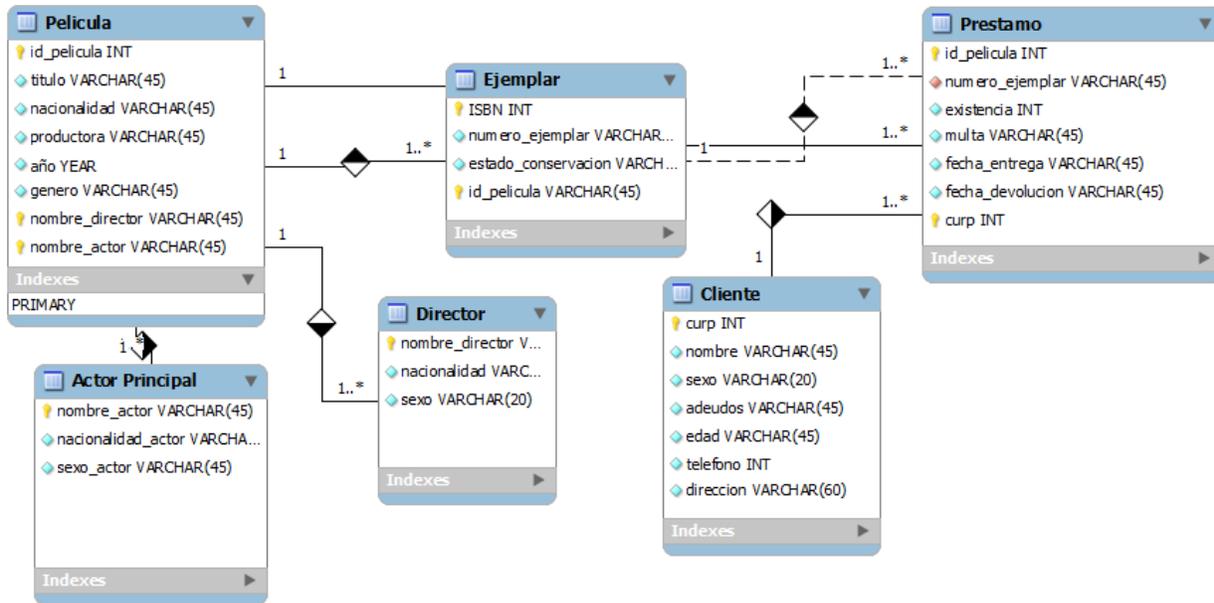


FIGURA 7 EJEMPLO DE UN DIAGRAMA DE BASE DE DATOS RELACIONAL (SILBERSCHATZ)

2.4. Almacén de Datos

Un almacén de datos según Kimball es “una copia de los datos transaccionales estructurados específicamente para consultas y análisis que ayudan a la toma de decisiones estratégicas en la entidad u organización en la que este se utiliza. Por otra parte, Inmon (Imhoff & Galemme, 1996) lo define como “una colección de datos, orientada a un determinado ámbito, integrado, no volátil y variable en el tiempo”. Adicionalmente, en muchas ocasiones, el almacén de datos es visto por otros autores como una arquitectura construida para integrar datos de múltiples fuentes heterogéneas y dar soporte a consultas estructuradas y/o ad hoc, informes analíticos y toma de decisiones.

La utilidad de un almacén de datos desde la perspectiva de un usuario es que toda la información está en un solo lugar, y no es necesario buscar en varias fuentes dispares de información o tratando de encontrar los archivos más antiguos en un desordenado sistema de correo electrónico o estructura de carpetas.

Se posee un acceso rápido al almacén de datos, encontrándose optimizado para la recuperación rápida de la información. Posee capacidad casi ilimitada sujeta solo al diseño físico. En cuanto a la historia disponible no sólo contiene información actual, sino también los datos de última semana, el mes pasado, el año pasado, y varios años atrás también. De hecho,

si nunca son eliminados los datos del almacén de datos, a menudo estos contienen mucha más información histórica que los sistemas de origen, permitiendo realizar análisis de tendencias o comparaciones complejas de forma fácil.

2.4.1. Objetivos de un Almacén de datos

En concordancia con Kimball (2002, The Data Warehouse Toolkit) y basados en la experiencia sobre dicha área, existen cierta cantidad de preocupación de carácter universal dentro de las organizaciones respecto de los almacenes de datos. Sin embargo, todas estas se traducen en citas de gestión empresarial sobre los objetivos que debe lograr cualquier almacén de datos:

Información de fácil acceso

El almacén de datos debe hacer que la información de una organización sea fácilmente accesible. El contenido del almacén de datos debe ser comprensible. Los datos deben ser intuitivos y evidentes para el usuario de negocios, no sólo el desarrollador. Comprensibilidad implica legibilidad; El contenido del almacén de datos debe estar etiquetado de manera significativa. Los usuarios de negocios quieren separar y combinar los datos en el almacén en combinaciones sin fin, un proceso comúnmente conocido como "slicing and dicing" (rebanar y cortar). Las herramientas que acceden al almacén de datos deben ser simples y fáciles de usar. También deben devolver los resultados de la consulta al usuario con un tiempo de espera mínimo.

Información coherente

El almacén de datos debe presentar la información de la organización de forma coherente. Los datos en el almacén deben ser creíbles. Los datos deben ser ensamblados cuidadosamente de una variedad de fuentes alrededor de la organización, filtrados, de calidad asegurada, y cargados solamente cuando es apto para el consumo del usuario. La información de un proceso empresarial debe coincidir con la información de otro. Si dos medidas de rendimiento tienen el mismo nombre, entonces deben significar lo mismo. Por el contrario, si dos medidas no significan lo mismo, entonces deben ser etiquetadas de manera diferente. Información consistente significa información de alta calidad. Significa que todos los datos son

contabilizados y completados. La coherencia también implica que las definiciones comunes para el contenido del almacén de datos están disponibles para los usuarios.

Adaptable y Resistente al Cambio

El almacén de datos debe ser adaptable y resistente al cambio. Simplemente no podemos evitar el cambio. Las necesidades de los usuarios, las condiciones comerciales, los datos y la tecnología están sujetos a las arenas movedizas del tiempo. El almacén de datos debe estar diseñado para manejar este cambio inevitable. Los cambios en el almacén de datos deben ser agradados, lo que significa que no invalidan los datos o aplicaciones existentes. Los datos y aplicaciones existentes no deben cambiarse o interrumpirse cuando la comunidad empresarial hace nuevas preguntas o se agregan nuevos datos al almacén. Si los datos descriptivos en el almacén se modifican, debemos contabilizar los cambios apropiadamente.

Seguro

El almacén de datos debe ser un bastión seguro que proteja nuestros activos de información. Las joyas de la corona informativa de una organización se acumulan en el almacén de datos. Como mínimo, el almacén probablemente contiene información sobre lo que estamos vendiendo a quien y a qué precio, detalles potencialmente dañinos en manos de las personas equivocadas. El almacén de datos debe controlar eficazmente el acceso a la información confidencial de la organización.

SopORTE a la Toma de Decisiones

El almacén de datos debe servir como base para mejorar la toma de decisiones. El almacén de datos debe tener los datos correctos en él para apoyar la toma de decisiones. Sólo hay una salida real de un almacén de datos: las decisiones que se toman después de que el almacén de datos ha presentado su evidencia. Estas decisiones entregan el impacto comercial y el valor atribuible al almacén. La etiqueta original que precede al almacén de datos sigue siendo la mejor descripción de lo que estamos diseñando: un sistema de apoyo a la decisión.

Solución orientada al Éxito

La comunidad empresarial debe aceptar el almacén de datos si se considera exitosa. No importa que hayamos construido una solución elegante utilizando los mejores productos y plataformas. Si la comunidad empresarial no ha abrazado el almacén de datos y ha seguido utilizándolo activamente seis meses después del entrenamiento, entonces hemos fallado en la prueba de aceptación. A diferencia de una reescritura del sistema operativo, donde los usuarios comerciales no tienen otra opción que usar el nuevo sistema, el uso del almacén de datos es a veces opcional. La aceptación del usuario empresarial tiene más que ver con la simplicidad que con cualquier otra cosa.

2.4.2. Características de un Almacén de Datos

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

- **Orientado a temas:** Los datos en una base de datos están organizados de tal manera que todos los elementos y registros relativos o relacionados al mismo evento u objeto del mundo real queden unidos entre sí. Tratándose de un almacén de datos existen áreas temáticas orientadas a los diferentes departamentos por los cuales ésta se encuentra conformada, la finalidad de este diseño se centra en el aumento de la eficacia al momento de realizar las consultas relacionadas, por supuesto, con la información de la organización.
- **Variante en el tiempo:** Los cambios producidos en los datos a lo largo del tiempo quedan registrados para que los informes que se puedan generar reflejen esas variaciones. En el caso de las bases de datos transaccionales u operacionales este factor no existe debido a su naturaleza como sistema, ya que solo almacenan información instantánea, a diferencia de los almacenes de datos en donde el factor tiempo toma un valor significativo. Por ejemplo, en caso de hacer un estudio sobre tendencias o modelos predictivos ya que los datos se encuentran estrechamente relacionados al momento de captura de estos.

- **No volátil:** La información no se modifica ni se elimina, una vez almacenado un dato, éste se convierte en información de sólo lectura, y se mantiene para futuras consultas. Para una base de datos transaccional los datos permanecen por períodos de tiempo limitados, en contraste con un almacén de datos en donde se acumulan las variaciones de los mismos datos con la finalidad de poder realizar cualquier tipo de análisis sobre estos. Es por esta razón que el almacén de datos funciona como una base de datos de solo lectura o consulta y no se modifica lo que haya ingresado eventualmente, por lo tanto, la información es permanente y cualquier actualización sobre este consiste solo en una agregación de un registro o acumulación de estos.
- **Integrado:** La base de datos contiene los datos de todos los sistemas operacionales de la organización, y dichos datos deben ser consistentes. Dentro de cualquier organización por lo general se usan diferentes sistemas transaccionales para la captura de los datos, y de igual forma existe una base de datos para cada uno de estos. En el caso de un almacén de datos todos estos sistemas deben integrarse en una base de datos centralizada que a través de distintos procesos de depuración, transformación y entrega de los datos es cargada.

2.4.3. Bodega de Datos (Data Mart)

Según Stair & Reynolds (2011, sexta edición) una bodega de datos "es un subconjunto de un almacén de datos. Las bodegas de datos aportan el concepto de almacén de datos - análisis en línea de las ventas, inventario y otros datos vitales de negocios que se han recopilado de los sistemas de procesamiento de transacciones - a pequeñas y medianas empresas y a departamentos dentro de las grandes empresas. En lugar de almacenar todos los datos empresariales en una base de datos monolítica, las bodegas de datos contienen un subconjunto de los datos para un solo aspecto del negocio de una empresa-por ejemplo, finanzas, inventario o personal. De hecho, un área específica en el mercado de datos puede contener datos más detallados que el almacén de datos."

Además, las bodegas de datos son más útiles para grupos más pequeños que desean acceder a datos detallados. Un almacén de datos contiene registros de resumen que pueden ser utilizados por toda una empresa. Debido a que las bodegas de datos normalmente contienen decenas de gigabytes de datos, a diferencia de los cientos de gigabytes en almacenes de datos, pueden implementarse en hardware menos potente con dispositivos de almacenamiento secundario más pequeños, lo que brinda ahorros significativos a una

organización. A pesar de que cualquier software de base de datos puede utilizarse para configurar un mercado de datos, algunos proveedores ofrecen software especializado diseñado y tasado específicamente para bodegas de datos.

2.4.4. Diferencias entre una Bodega de Datos y un Almacén de Datos

Para entender las diferencias entre un enfoque y otro es necesario darle un vistazo a la tabla 2. Aquí están los dos enfoques básicos diferentes: bodegas de datos dependientes de la alimentación del almacén de datos en general, y varias bodegas de datos departamentales o locales que se combinan en un almacén de datos. En el primer enfoque, se extraen datos de los sistemas transaccionales; A continuación, transformar, limpiar, integrar y mantener los datos en el almacén de datos. Entonces, ¿qué enfoque es mejor en su caso, el enfoque de arriba abajo (bottom-up) o de abajo hacia arriba (top-down)?, para responder a esta pregunta es necesario que examinemos con mayor detenimiento estos dos enfoques.

	Almacén de datos	Bodega de datos
Características	<ul style="list-style-type: none"> ◆ Empresa ◆ Unión de todos los mercados de datos ◆ Datos recibidos del área de ensayo ◆ Consultas sobre recurso de presentación ◆ Estructura de la vista corporativa de los datos ◆ Organizado en el modelo E-R 	<ul style="list-style-type: none"> ◆ Departamental ◆ Un solo proceso de negocio ◆ Star-join (hechos y dimensiones) ◆ Tecnología óptima para el acceso y análisis de datos ◆ Estructura para adaptarse a la visión departamental de los datos

TABLA 2 ALMACÉN DE DATOS VS BODEGA DE DATOS

Enfoque de arriba hacia abajo (Top-Down)

Ventajas:

- Visión empresarial de los datos
- Inherentemente estructurado
- Almacenamiento simple y centralizado de datos sobre el contenido
- Reglas centralizadas y control
- Puede ver resultados rápidos si se implementa con iteraciones

Desventajas:

- Toma más tiempo para construir incluso con un método iterativo
- Exposición alta / riesgo al fallo
- Necesita un alto nivel de habilidades inter-funcionales

Este es el enfoque general en el que se construye el almacén de datos general o empresarial. En el cual no se tiene una colección de islas fragmentadas de información. El almacén de datos es grande e integrado. Este enfoque, sin embargo, llevaría más tiempo construir y tiene un alto riesgo de fracaso. Si no tiene profesionales experimentados en su equipo, este enfoque podría ser riesgoso.

Enfoque de abajo hacia arriba (Bottom-Up)

Ventajas:

- Implementación más rápida y fácil de piezas manejables
- Rentabilidad favorable de la inversión
- Menos riesgo de fracaso
- Inherentemente incremental; Puede programar los data marts importantes primero
- Permite al equipo del proyecto aprender y crecer

Desventajas:

- Cada data mart tiene su propia visión estrecha de los datos
- Permea los datos redundantes en cada data mart
- Perpetua datos inconsistentes e irreconciliables
- Proliferación de interfaces inmanejables

En el enfoque de abajo hacia arriba, se construyen los almacenes de datos departamentales en forma individual. Se establece un esquema de prioridad para determinar qué data mart (bodega de datos) se debe construir primero. El inconveniente más grave de este enfoque es la fragmentación de los datos. Cada data mart (bodega de datos) independiente estará en desconocimiento de los requerimientos generales de toda la organización.

2.4.5. Modelo Dimensional

El modelado dimensional es “una técnica de diseño lógico que trata de presentar los datos de una manera estándar e intuitiva que permita un acceso de alta velocidad.” (Connolly, T, Begg, C, 2005).

Dimensión

Una dimensión es una entidad o una colección de entidades relacionadas, usadas por los analistas para identificar el contexto de las medidas con las que trabajan, estas determinan el contexto para las medidas.

Cuando se trata de definir las dimensiones el añadir, eliminar o cambiar propiedades particulares de las dimensiones candidatas cambia el contexto y en consecuencia el significado de la medida candidata.

Por último, se debe señalar que las dimensiones “son referenciadas por las llamadas llaves de dimensión y que estas poseen entidades, atributos, jerarquías e incluso niveles de agregación.” (Kimball, 1998).

Las dimensiones buscan determinar un contexto para el análisis de los hechos, por tanto, se tratan de grupos homogéneos de elementos, en muchas ocasiones, jerarquizados, cuyo papel es promocionar la información contenida en los hechos.

Las dimensiones las pueden compartir diferentes tablas de hechos y cubos OLAP, y una sola tabla de hechos o un solo cubo de datos puede hacer referencia a las dimensiones ya que cada tabla de dimensión tiene una clave principal simple que se corresponde con exactamente uno de los componentes de la clave compuesta de la tabla de hechos y que la identifica al mismo nivel de detalle. A continuación, se enumeran algunos de los tipos de tablas de dimensiones más comúnmente utilizados, los cuales fueron extraídos de Kimball (2008):

- **Dimensión lentamente cambiante (SCD por sus siglas en inglés):** Son dimensiones que permiten gestionar de una forma correcta los cambios que se producen en sus atributos. Existen tres métodos para el tratamiento de las SCD, los cuales son:

Añadir Fila: Si ocurre un cambio en un atributo, se crea un nuevo registro en la tabla, este nuevo registro tiene una nueva clave subrogada, de forma que una entidad de sistema operacional (por ejemplo, un cliente), puede tener varios registros en la tabla de la dimensión según se van produciendo los cambios

Sobrescribir: Se sobrescribe el valor antiguo de un atributo sin guardar el valor antiguo de dicho atributo, lo cual implica perder toda la historia del dato.

Añadir Columna: Cuando ocurre un cambio en un atributo, se guarda el valor anterior en una columna distinta, actualizando el campo con el nuevo valor (para cada campo se tiene una tupla con el valor anterior, valor actual).

- **Dimensiones Normales:** Son aquellas que agrupan diferentes atributos que están relacionados por el ámbito al que se refieren (todas las características de un cliente, los diferentes componentes de la dimensión tiempo, entre otros).
- **Dimensiones Causales:** Son aquella que incluye atributos que pueden causar cambios en los procesos de negocio (por ejemplo, la dimensión promoción en el proceso de negocio de ventas).
- **Dimensiones Heterogéneas:** Son dimensiones que agrupar conjuntos heterogéneos de atributos, que no están relacionados entre sí.
- **Dimensiones Roll-Up:** Es una dimensión que es subconjunto de otra, son necesarias para el caso en que se tengan tablas de hechos con diferente granularidad.

- Dimensiones Basura (Junk): Es una dimensión que agrupa indicadores de baja cardinalidad como pueden ser banderas (flags) o indicadores.
- Dimensiones Role-Playing: Ocurre cuando una misma dimensión interviene en una tabla de hechos varias veces (por ejemplo, la fecha en una tabla de hechos donde se registran varias fechas referidas a conceptos diferentes), es necesario reutilizar la misma dimensión, pues no tiene sentido crear tantas dimensiones como usos se hagan de ella.
- Dimensiones Degeneradas: Son dimensiones que no tienen ningún atributo y por tanto, no tienen una tabla específica de dimensión, sólo se incluye para ellas un identificador en la tabla de hechos, que identifica completamente a la dimensión (por ejemplo, un pedido de ventas), además se usan cuando interesa tener identificada la transacción (como por ejemplo para realizar minería de datos), pero los datos interesantes de este elemento los tenemos repartidos en las diferentes dimensiones (cliente, producto u otro).
- Mini dimensiones o Dimensiones Outrigger: Son un conjunto de atributos de una dimensión que se extraen de la tabla de dimensión principal pues se suelen analizar de forma diferente. El típico ejemplo de estas dimensiones son los datos sociodemográficos asociados a un cliente.

Hechos

Un hecho es una colección de medidas relacionadas con sus dimensiones y puede representar un objeto de negocio, una transacción o un evento que es utilizado por el analista de información.

Los hechos proporcionan una información cuantitativa sobre las características del negocio que se quiere analizar ya que estos son una colección de medidas relacionadas con sus dimensiones y pueden representar un objeto de negocio, una transacción o un evento que es utilizado por el analista de información. Su finalidad es proporcionar información necesaria para la gestión, facilitando el conocimiento del Negocio o Proceso a modelar, y fundamentar, entre otras, la toma de decisiones.

De modo que una tabla de hechos (conocida en el lenguaje inglés como Fact Table) se define como la tabla central de un esquema dimensional (en estrella o en copo de nieve) que contiene los valores de las medidas de negocio o dicho de otra forma los indicadores de negocio.

La tabla de hechos contiene un identificador de hecho, llaves de dimensión, medidas y varios tipos de atributos.

A continuación, se describen algunos de los tipos de tablas de hechos más comúnmente utilizados, referenciados de (Kimball, 2008):

- Tabla de hechos de instantáneas periódicas: Son tablas usadas para recoger información de forma periódica a intervalos de tiempo regulares sobre un hecho. Nos permiten tomar una instantánea (Snapshot) de la situación en un momento determinado.
- Un ejemplo puede ser la foto del stock de los materiales al final de cada día.
- Tabla de hechos de transacciones: Representan eventos que suceden en un determinado espacio-tiempo. Se caracterizan por permitir analizar los datos con el máximo detalle. Reflejan las transacciones relacionadas con nuestros procesos de negocio (ventas, compras, inventario, contabilidad, entre otros).
- Tabla de hechos sin hechos: Son tablas que no tienen medidas y representan la ocurrencia de un evento determinado. Por ejemplo, la asistencia a un curso puede ser una tabla de hechos sin métricas asociadas.
- Tabla de hechos de instantáneas acumulativas: Son tablas que representan el ciclo de vida completo de una actividad o proceso. Suelen representar valores acumulados.

Jerarquía

Se define como un conjunto de atributos para una dimensión cuya finalidad es la de realizar análisis al ser agrupados. En la figura 8 podemos observar un ejemplo de esto.

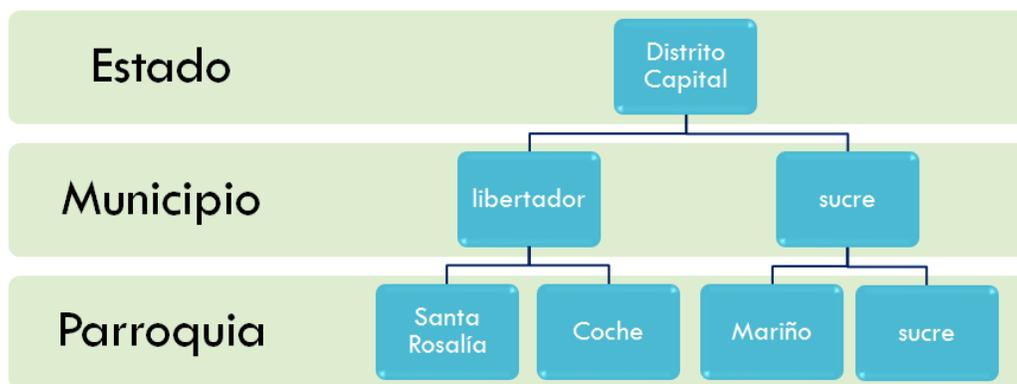


FIGURA 8 EJEMPLO DE UNA JERARQUÍA GEOGRÁFICA

Granularidad

La granularidad representa el nivel de detalle al que se desea almacenar la información sobre el negocio que se esté analizando. Por ejemplo, los datos referentes a ventas o compras realizadas por una empresa, pueden registrarse día a día, en cambio, los datos pertinentes a pagos de sueldos o cuotas de socios, podrán almacenarse a nivel de mes, por ejemplo.

Mientras mayor sea el nivel de detalle de los datos, se tendrán mayores posibilidades analíticas, ya que los mismos podrán ser resumidos o sumariados. Es decir, los datos que posean granularidad fina (nivel de detalle) podrán ser resumidos hasta obtener una granularidad media o gruesa. No sucede lo mismo en sentido contrario, ya que, por ejemplo, los datos almacenados con granularidad media podrán resumirse, pero no tendrán la facultad de ser analizados a nivel de detalle. Es decir, si la granularidad con que se guardan los registros es a nivel de día, estos datos podrán sumariarse por semana, mes, semestre y año, en cambio, si estos registros se almacenan a nivel de mes, podrán sumariarse por semestre y año, pero no lo podrán hacer por día y semana.

Esquemas

Esquema en estrella

Un esquema estrella es un modo de representar datos multidimensionales en una base de datos relacional, donde las tablas de dimensión guardan información descriptiva acerca de sus miembros y sus relaciones, mientras que las tablas de hechos almacenan datos del negocio.

Es debido destacar que las consultas multidimensionales son hechas haciendo operaciones (Joins) entre las tablas de hechos y las dimensiones. Por otra parte, una de las ventajas de utilizar un esquema como este es su simplicidad y comprensión para los usuarios finales. Muy a menudo durante la fase de diseño de un almacén de datos, esquemas de estrella se utilizan para señalar la traducción inicial de las preguntas de negocio en diagramas de base de datos lógico. (Kimball, 1996)

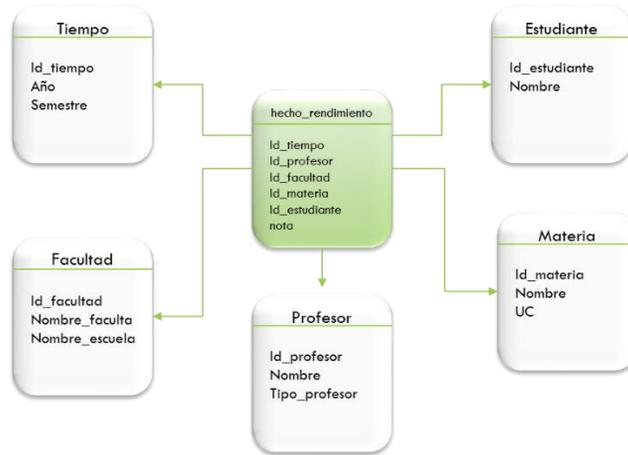


FIGURA 9 EJEMPLO DE UN ESQUEMA ESTRELLA

Esquema en copos de nieve

Un esquema en copo de nieve es una estructura algo más compleja que el esquema en estrella. Se da cuando alguna de las dimensiones se implementa con más de una tabla de datos.

La finalidad es normalizar las tablas y así reducir el espacio de almacenamiento al eliminar la redundancia de datos; pero por otra parte este esquema genera peores rendimientos al tener que crear más tablas de dimensiones y más operaciones entre las tablas lo que tiene un impacto directo sobre el rendimiento al momento de realizar la consulta. (Kimball, 1996)

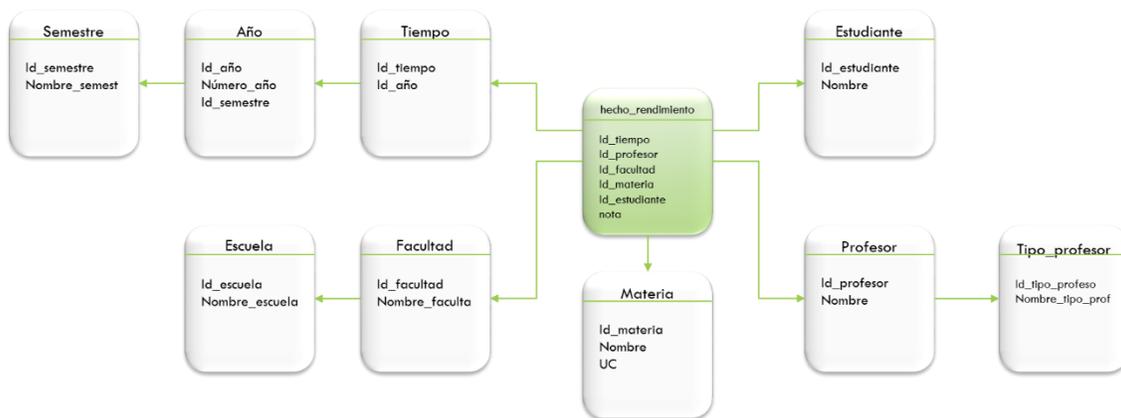


FIGURA 10 EJEMPLO DE UN ESQUEMA COPO DE NIEVE

Esquema de constelación

El esquema constelación es la conformación de muchos esquemas estrellas unidos lógicamente. Su nombre viene dado por el término astronómico de constelación que hace referencia a un cúmulo de estrellas.

Un esquema constelación permite unir distintos procesos de negocios o áreas analíticas de una organización, a partir de dimensiones o elementos de la tabla de hechos (esto último poco frecuente) en común. Esto amplía la capacidad de generar reportes que reflejen una situación más general de la organización, así como una mayor posibilidad de utilizar distintos conjuntos de data para un análisis particular.

Como se ve en la figura 11, a partir de un conjunto de dimensiones (estudiante, tiempo, facultad) permite unir dos posibles modelos estrellas, uno que mide el rendimiento estudiantil con otra que mantiene el registro de admisión de estudiantes en la universidad para análisis de todo lo que son las propiedades de estos nuevos estudiantes, es decir, cuando ingresaron, como ingresaron, de que liceo vienen y hacia que facultad van. Esta unión de modelos en un esquema mayor permite una mayor capacidad de análisis en el sentido que, por ejemplo, podría medir como es el rendimiento estudiantil de una materia dada en función de un liceo o de la forma en que ingreso el estudiante a la universidad.

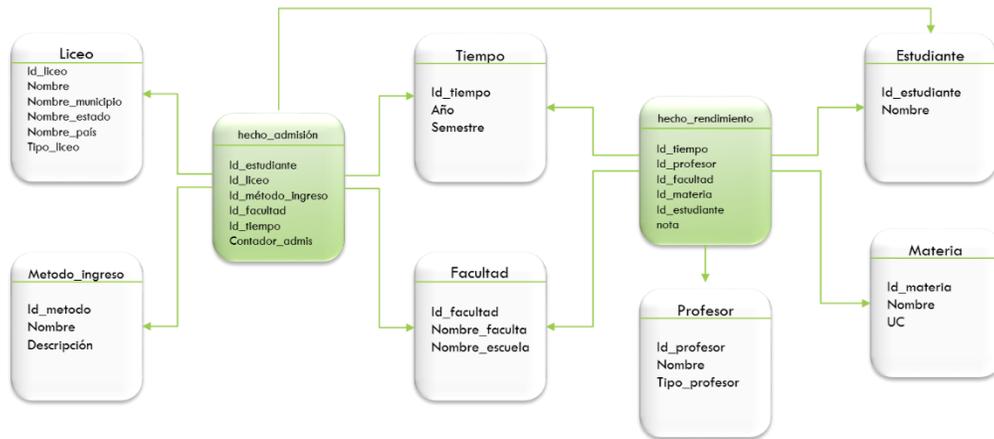


FIGURA 11 EJEMPLO DE UN ESQUEMA CONSTELACIÓN

Metodología de desarrollo para almacenes de datos

A continuación, se enumeran los pasos descritos en la metodología del modelado dimensional propuesta por Kimball la cual especifica los pasos requeridos para el diseño de un mercado de datos.

- *Selección de los procesos de negocio*

Este paso tiene por objetivo definir los procesos de negocio que van a ser analizados, el análisis de los procesos de negocio, proporcionará al usuario el conocimiento necesario para la toma de decisiones estratégicas que permitan mejorar dicho proceso de negocio. Para identificar los procesos de negocio, se pueden utilizar técnicas tradicionales de obtención de requisitos como son las mesas de trabajos o las entrevistas.

- *Selección de la granularidad*

El segundo paso consiste en describir el funcionamiento del Almacén de Datos en relación a las tareas de análisis que los usuarios pretenden realizar al interactuar con él, esto implica decidir exactamente qué es lo que va a representar cada registro de la tabla de hechos.

- *Identificación y conformación de las dimensiones*

Este paso representa los datos que serán almacenados en las dimensiones del Almacén de Datos, estos establecen en contexto para realizar pregunta acerca de los hechos, pero para esto es necesario que las dimensiones sean construidas con suficiente detalle para describir y satisfacer correctamente la granularidad. Si la dimensión aparece en más de un mercado de datos entonces dicha dimensión está conformada y debe estar sincronizada entre todos los mercados de datos que la usen, de lo contrario el almacén de datos no podrá funcionar correctamente.

- *Selección de los hechos*

Por último, el cuarto paso consiste en determinar qué hechos pueden usarse en el mercado de datos, los cuales deben ser expresado según el grano de la granularidad escogida del segundo paso. Cabe destacar que los hechos deben ser numéricos y/o aditivos y que pueden añadirse hechos adicionales a cualquier tabla de hechos en un instante posterior a la carga del almacén de datos siempre que estos nuevos hechos sean coherentes con la granularidad de la tabla.

2.4.6. Ventajas y Desventajas de los Almacenes de Datos

Ventajas

1. Proporciona información clave para la toma de decisiones dentro de cualquier organización.
2. Mejora la calidad de las decisiones tomadas.
3. Especialmente útil para la planificación a mediano y largo plazo.
4. Son sistemas relativamente sencillos de instalar si las fuentes de datos y los objetivos están claros.
5. Muy útiles para el almacenamiento de análisis y consultas de históricos.
6. Proporciona un gran poder de procesamiento de información.
7. Permite una mayor flexibilidad y rapidez en el acceso a la información.
8. Facilita la toma de decisiones en los negocios.
9. Las empresas obtienen un aumento de la productividad.
10. Proporciona una comunicación fiable entre todos los departamentos de la empresa.
11. Mejora las relaciones con los proveedores y los clientes.
12. Permite conocer qué está pasando en el negocio, es decir, estar siempre enterado de los buenos y malos resultados.
13. Transforma los datos en información y la información en conocimiento
14. Permite hacer planes de forma más efectiva.
15. Reduce los tiempos de respuesta y los costes de operación.

Desventajas

1. No es muy útil para la toma de decisiones en tiempo real debido al largo tiempo de procesamiento que puede requerir. En cualquier caso, la tendencia de los productos actuales (junto con los avances del hardware) es la de solventar este problema convirtiendo la desventaja en una ventaja.
2. Requiere de continua limpieza, transformación e integración de datos.
3. Mantenimiento.
4. En el proceso de implantación puede encontrarse dificultades ante los diferentes objetivos que pretende una organización.
5. Una vez implementado puede ser complicado añadir nuevas fuentes de datos.
6. Requieren una revisión del modelo de datos, objetos, transacciones y además del almacenamiento.
7. Tienen un diseño complejo y multidisciplinar.
8. Requieren una reestructuración de los sistemas operacionales.
9. Tienen un alto coste.
10. Requieren sistemas, aplicaciones y almacenamiento específico.

2.5. Inteligencia de Negocios

Desde una perspectiva histórica Howard Dresner, analista de largo plazo del Grupo Gartner, es a menudo acreditado con la definición de inteligencia de negocios (BI, por sus siglas en inglés).

Aunque Dresner popularizó el término, él no fue el inventor, de hecho, la inteligencia de negocio a largo plazo fue utilizada por primera vez en 1958 por Hans Peter Luhn en el diario IBM. Sin embargo, es la definición de BI de Howard Dresner hecha en 1989 la que se utiliza con mayor frecuencia hoy en día: "Conceptos y métodos para mejorar la toma de decisiones empresariales mediante el uso de sistemas de apoyo basados en hechos". Esta definición describe de una manera muy elegante que BI es mucho más que la tecnología por sí sola (un hecho que es a menudo pasado por alto por muchos de los profesionales de BI) y que BI es: un medio para un fin, no el objetivo en sí.

De lo anterior expuesto, se tiene que la inteligencia de negocio actúa como un factor estratégico para la organización, generando una potencial ventaja competitiva, que no es otra

que proporcionar información privilegiada para responder a los problemas de negocio en un determinado momento.

En definitiva, una solución de inteligencia de negocio permite responder las siguientes preguntas: ¿Qué está ocurriendo?, ¿Por qué está ocurriendo?, ¿Qué ocurrirá? y ¿Qué camino debo seguir?

2.5.1. Características de una Solución de Inteligencia de Negocio

En concordancia con cano (2007, Business Intelligence) y otros autores, toda solución de inteligencia de negocios debe contar con las siguientes características:

- **Accesibilidad a la información:** Los datos son la fuente principal de este concepto. Lo primero que deben garantizar este tipo de herramientas y técnicas será el acceso de los usuarios a los datos con independencia de la procedencia de estos.
- **Apoyo en la toma de decisiones:** Se busca ir más allá en la presentación de la información, de manera que los usuarios tengan acceso a herramientas de análisis que les permitan seleccionar y manipular sólo aquellos datos que les interesen.
- **Orientación al usuario final:** Se busca independencia entre los conocimientos técnicos de los usuarios y su capacidad para utilizar estas herramientas.
- **Visión unificada de los datos:** se pretende que los datos se encuentren en un repositorio centralizado, colocando de lado el origen de donde provengan o el tipo de dato.
- **Informes y consultas personalizadas:** orientado al desarrollo de consultas de cualquier índole y emisión de reportes que se ajustan a las peticiones de los usuarios en función de los datos almacenados en el repositorio de datos.
- **Informes analíticos a través de vistas graficas:** toda solución de inteligencia de negocios debe proveer un panel de control que facilite cualquier visualización de los indicadores sobre la organización que está siendo implementada.
- **Procesamiento de grandes cantidades de datos:** permite la comparación de data actual en contraste con la data histórica.

2.5.2. Funciones de una Solución de Inteligencia de Negocio

Tratándose de una solución de inteligencia de negocios se espera que esta provea de las siguientes funciones primordiales:

- Recopilar, ajustar y consolidar toda la información de la organización, mediante un almacén de datos, permitiendo así su explotación.
- Facilitar la extracción de información en función de los datos y el conocimiento de la información, con la utilización del software adecuado.
- Perfeccionar las consultas de alto nivel, realizando las transformaciones oportunas a cada sistema (OLTP – OLAP), y liberando los servidores operacionales.

2.5.3. Arquitectura de una Solución de Inteligencia de Negocio

Un sistema de inteligencia de negocio esta optimizado para realizar preguntas sobre los datos, a diferencia de los sistemas operaciones que están diseñados para soportar continuas operaciones CRUD acrónico en español de Crear, Obtener, Actualizar y Borrar. Estos datos pueden tener su origen en diferentes sistemas y no necesariamente deben seguir la misma convención de diseño, generalmente han sido diseñados de forma totalmente diferente.

Estas herramientas de inteligencia se basan en la utilización de un sistema de información que se forma con distintos datos extraídos de los datos de producción, con información relacionada con la empresa o sus ámbitos y con datos económicos. Esto mediante las herramientas y técnicas ETL conocidas por sus siglas en inglés como extracción, transformación y carga, lo que refiere al proceso de extraer los datos de las distintas fuentes, depurarlos y prepararlos (homogeneización de los datos) para luego cargarlos en un almacén de datos (Data Warehouse). Entendiéndose por almacén de dato según Kimball (2002) a "una copia de las transacciones de datos específicamente estructurada para la consulta y el análisis". También fue Kimball quien determinó que un Data Warehouse no era más que: "la unión de todos los Datamarts de una entidad".

Un Datamart no es más que una base de datos departamental, especializada en el almacenamiento de los datos de un área de negocio específica). Por su parte, Inmon (1992) fue uno de los primeros autores en escribir sobre el tema de los almacenes de datos, define un almacén de datos en términos de las características del repositorio de datos:

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

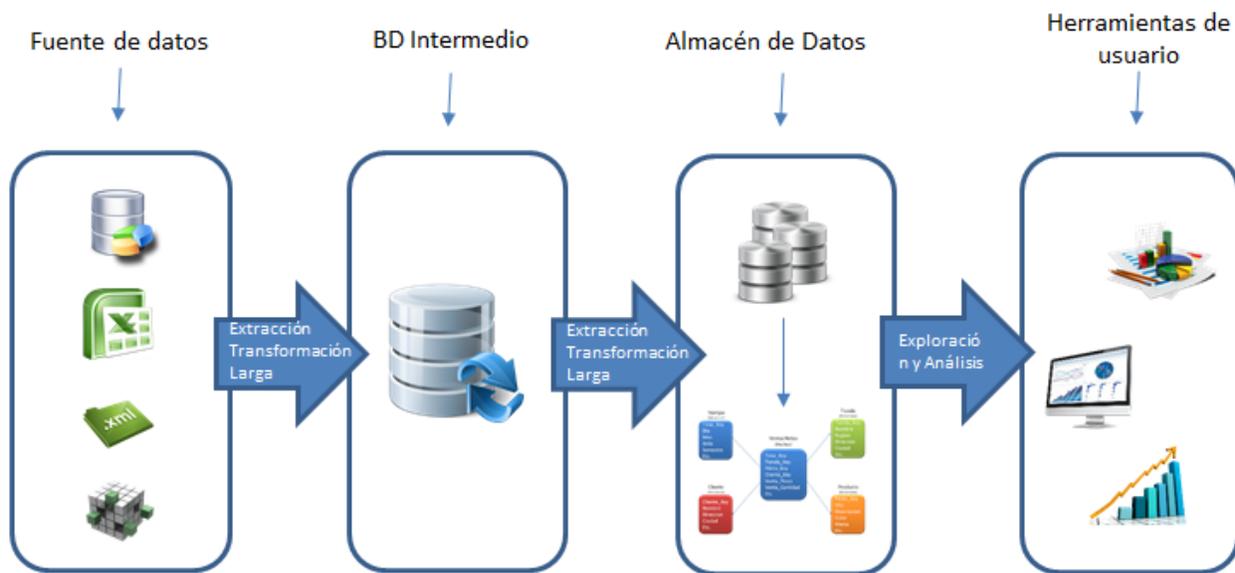


FIGURA 12 ARQUITECTURA DE UNA SOLUCIÓN DE INTELIGENCIA DE NEGOCIOS

En la Figura 12 se muestra la arquitectura de una solución de inteligencia de negocios que será explicada detalladamente a continuación:

- Fuentes de Datos: Representan los distintos repositorios de información que posee una organización, normalmente son bases de datos operacionales, archivos planos, archivos XML o incluso hojas de cálculo desarrolladas en herramientas ofimáticas como Microsoft Excel.

- **Extracción, Transformación y Carga (ETL):** Según Kimball (2008) el proceso de Extracción, Transformación y Carga (ETL) es la base de todo el proceso al momento de desarrollar una solución de inteligencia de negocio. Si el ETL se diseña adecuadamente podemos extraer los datos de los sistemas operacionales y aplicar diferentes reglas para aumentar la calidad, la consistencia y consolidar la información proveniente de nuestros sistemas.
- **Repositorio:** Una vez realizado el proceso de ETL la información transformada se almacena. Según Inmon (2002) un almacén de datos (data Warehouse) es una colección de datos integrados no volátil y variable en el tiempo, con el fin de ayudar a la toma de decisiones de la entidad que se utiliza.
- **Aplicaciones de Inteligencia de Negocios (Accesos):** Las aplicaciones de inteligencia de negocio es una parte fundamental de un proyecto de BI y su función consiste en proporcionar de una manera estructurada la posibilidad de acceder al almacén de datos. Se puede decir que estas aplicaciones son la cara visible del proyecto, y pueden estar constituidas por una amplia gama de informes y herramientas de análisis, es decir podemos tener desde informes simples hasta herramientas analíticas complejas.

2.6. Tecnologías y Herramientas para la construcción de una Solución de Inteligencia de Negocio

La selección de una u otra herramienta de inteligencia de negocios, además del costo para la organización deben tomar en cuenta múltiples aspectos a considerar, tales como los que define Cortes (2007):

- **Información requerida:** Se debe saber que información es requerida, sin complicarse, sobre todo al principio, con indicadores y modelos complejos: indicadores selectivos, sencillos, admitidos por todos los usuarios, etc. son una buena fórmula en las primeras etapas del BI.
- **A quién va dirigida:** Organización en general, gestión, dirección, dirección estratégica, etc. En este caso existen dos grandes roles de usuario de las herramientas BI:
 - **Los consumidores de la información:** Se trata de los usuarios finales de la aplicación, aquellos que consultaran los informes para la posterior toma de

decisiones. La dirección será la encargada de decidir quién tendrá acceso a qué información. En general el grupo lo conforman directivos, responsables y jefes de proyectos.

- Los productores de información: Se trata del personal del servicio técnico o de usuarios avanzados. Serán los encargados, de crear los informes o. En definitiva, son usuarios que poseen conocimientos de informática y entienden y pueden manejar la información almacenada.

- Uso de la información: Para la toma de decisiones pueden existir múltiples necesidades particulares a cada organización, tales como: analizar diferentes aspectos de la evolución de la empresa, presentar información de forma más intuitiva, comparar información en diferentes periodos de tiempo, comparar resultados con previsiones, identificar comportamientos y evoluciones excepcionales, confirmar o descubrir tendencias, necesidad de análisis predictivos, etc.
- Aspectos meramente técnicos y funcionales: tiempos de respuesta, integración, seguridad, navegación, entorno gráfico, etc.

Beneficios de utilizar Inteligencia de Negocios

- Beneficios tangibles: Como por ejemplo la reducción de costes, generación de ingresos, reducción de tiempos para las distintas actividades del negocio.
- Beneficios intangibles: El hecho de que tengamos disponible la información para la toma de decisiones hará que más usuarios utilicen dicha información para tomar decisiones y mejorar la nuestra posición competitiva.
- Beneficios estratégicos: Todos aquellos que nos facilitan la formulación de la estrategia, es decir, a qué clientes, mercados o con qué productos dirigirnos.

2.6.1. Pentaho Business Intelligence

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

lenguaje de programación Java y tiene un ambiente de implementación también basado en Java, haciendo así que Pentaho sea una solución muy flexible al cubrir una alta gama de necesidades empresariales. Las soluciones que Pentaho ofrece se componen fundamentalmente de una infraestructura de herramientas de análisis e informes integrados con un motor de flujo de negocio de los procesos de negocio.

Consta de dos presentaciones básicas, la primera en su versión Pentaho Enterprise Edition cuya plataforma cuenta con capacidades de consulta, presentación de informes, paneles de control y tableros interactivos con posibilidad de análisis sobre estos que requieren de una certificación para su uso, sin embargo, también se encuentra su versión Pentaho Community Edition en donde se presenta funcionalidades similares de pruebas pero con capacidades limitadas que solo se pueden encontrar en la primera versión mencionada.

Características

- Lenguaje: Plataforma 100% J2EE, asegurando la escalabilidad, integración y portabilidad.
 - Fuente de datos: Amplia conectividad a cualquier tipo de fuente de datos.
 - Arquitectura: Abierta y basada en estándares, además todos los componentes están expuestos vía servicios web para facilitar la integración con Arquitecturas Orientadas a Servicios (SOA).
 - Servidor: Puede correr en servidores compatibles con J2EE como JBOSS AS, WebSphere, Tomcat, WebLogic y Oracle AS.
 - Base de datos: Es compatible con una gran variedad de fuentes de datos, a través de la conectividad a base de datos de java vía JDBC. Soporta base de datos propietarias y de Software Libre, tales como: IBM DB2, Microsoft SQL Server, MySQL, Oracle, PostgreSQL, NCR, Teradata, Firebird, EnterpriseDB, entre otras.
 - Sistema operativo: No hay dependencia, es multiplataforma y su lenguaje es interpretado.
 - Lenguaje de programación: Java, JavaScript, JSP, XSL (XSLT/XPath/XSL-FO).
 - Interfaz de desarrollo: Java SWT, Eclipse, Web-based. Repositorio de datos basado en XML.
 - Seguridad: Autenticación de usuario basada en roles y la autorización de la integración de sesión única y fácil con los sistemas de seguridad externos como LDAP, Active Directory y otros.
 - Salida de datos: Salida en formatos populares como HTML, Excel, CSV, PDF y RTF.
-

Componentes de Pentaho

Data Integration: Pentaho Data Integration extrae, limpia e integra esta valiosa información y la pone en manos del usuario, lo cual resulta útil para las organizaciones que tienen información disponible en aplicaciones y base de datos separadas, proveyendo consistencia, ya que centraliza una versión de todos los recursos de información. Pentaho Data Integration anteriormente conocida como Kettle incluye cinco herramientas:

- Spoon: Es una herramienta gráfica para diseñar transformaciones ETL.
- PAN: Herramienta para ejecutar transformaciones diseñadas en el Spoon.
- CHEF: Herramienta para ejecutar trabajos complejos que automatizan los procesos de actualización de la base de datos
- Kitchen: Herramienta que ayuda a ejecutar los trabajos por lotes, permitiendo iniciar y controlar fácilmente el proceso de ETL
- Carter: Es un servidor web que permite la supervisión remota del proceso de ETL

Kettle posee un conjunto de elementos claves para llevar a cabo el proceso de ETL como lo son:

- Transformation: Es una herramienta que permite realizar una variedad de tareas con datos, moviéndolos de un lugar a otro. Su rol principal es extraer datos de diversas fuentes de datos, transformarlos de una representación a otra y cargarlos en fuentes de salida; es decir abarca los pasos de extracción (E), transformación (T) y carga (L). Consiste de un número separado de acciones llamadas steps relacionadas a través de un flujo llamado hop.
- Steps: Son las acciones o pasos que conforman a una transformación, los cuales son diseñados con una función específica.
- Hops: Es la representación gráfica del flujo de datos entre los steps, usados para transportar datos de un step a otro. También se usan en los jobs con el fin de llevar un control en el flujo de ejecución.
- Job: Es un componente que crea una secuencia de actividades que brinda un orden de ejecución, por lo tanto, son usados para crear un control de flujo. Por lo general, un Job consiste de una serie de transformaciones que se desean ejecutar en un orden.

Los módulos de la plataforma Pentaho BI son:

- Reporting: Solución proporcionada por Pentaho e integrada en su suite para el desarrollo de informes. Dentro de este existen tres opciones con diferentes enfoques y dirigidos a diferentes tipos de usuarios. Todos los informes incluyen la opción de imprimir o exportar a formato PDF, XLS, HTML y texto.
- Pentaho Report Designer: Editor basado en eclipse con prestaciones profesionales de calidad y con capacidad de personalización de informes a las necesidades de negocio destinado a desarrolladores.

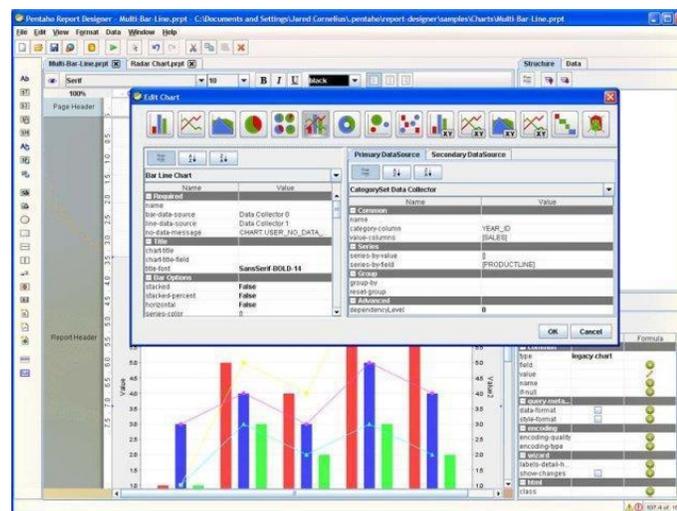


FIGURA 13 INTERFAZ DE PENTAHO REPORT DESIGNER

- Pentaho Report Design Wizard: Herramienta de diseño de informes, que facilita el trabajo y permite a los usuarios obtener resultados de forma inmediata. Está destinada a usuarios con menos conocimientos técnicos.
- Web Ad-Hoc reporting: Es el similar a la herramienta anterior, pero vía web. Extiende la capacidad de los usuarios finales para la creación de informes a partir de plantillas pre-configuradas.

- **Analysis:** Suministra a los usuarios un sistema avanzado de análisis de información. Permite al mismo navegar por los datos, ajustando la visión de los datos, los filtros de visualización, añadiendo o quitando los campos de agregación. Los datos pueden ser representados en una forma de SVG o Flash, los Dashboard widgets, están integrados con los sistemas de minería de datos y los portales web (portlets).
- **Dashboards:** Incorporar una gran variedad de tipos de gráficos, tablas y velocímetros, para brindar a los usuarios la posibilidad de visualizar de forma atractiva información crítica para el proceso de negocio. También se puede integrar con los Portlets JSP para visualizar informes, gráficos y análisis OLAP.

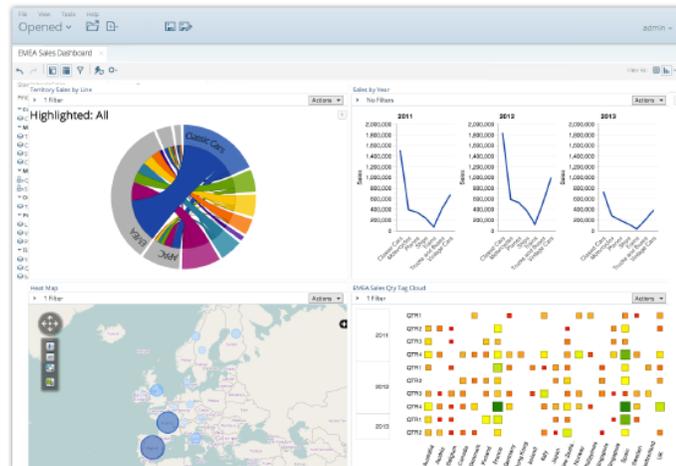


FIGURA 14 EJEMPLO DE DASHBOARDS

- **Data Mining:** Es el proceso de correr datos en algoritmos sofisticados, relevando significantes patrones y correlaciones que pueden estar escondidos. Esto puede ser usado para ayudar a entender lo mejor para el negocio y explotar el rendimiento de este en un futuro prediciendo completamente en el análisis. Se realiza con la herramienta WeKa.
- **Data Integration:** Muchas organizaciones tienen información disponible en aplicaciones y BD separadas. Pentano Data Integration abre, limpia e integra esta

valiosa información y la pone en manos del usuario. Provee consistencia, una sola versión de todos los recursos de información, que es uno de los más grandes desafíos para las organizaciones de hoy en día. Se realiza con la herramienta Kettle ETL (Pentaho Data Integration) que permite implementar los procesos ETL.

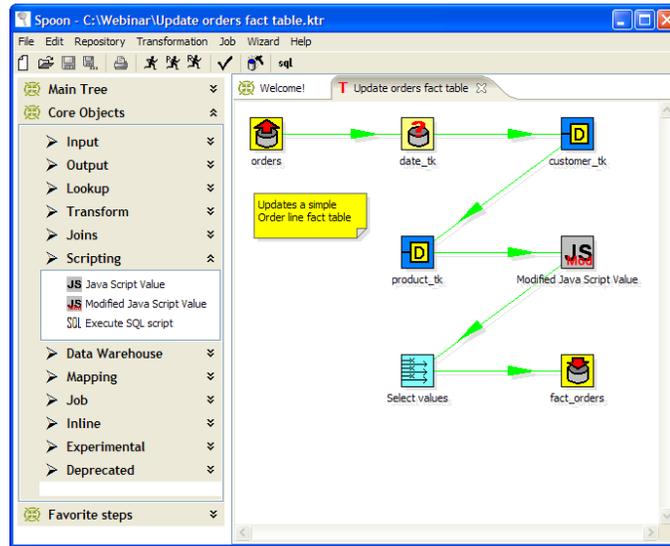


FIGURA 15 INTERFAZ DE DATA INTEGRATOR

2.6.2. Oracle Business Intelligence (OBI)

El crecimiento empresarial suele ir a la par de la generación de muchos datos, que eventualmente son necesarios para tomar decisiones relevantes pero que normalmente están distribuidos en múltiples bases de datos en los diferentes departamentos de la compañía.

Oracle BI es capaz de reunir todos estos datos y que proporcione un conjunto de herramientas para analizarlos y crear informes, cuadros de mando, alertas y otras funcionalidades que permitan dar el soporte adecuado a la toma de decisiones en forma acertada.

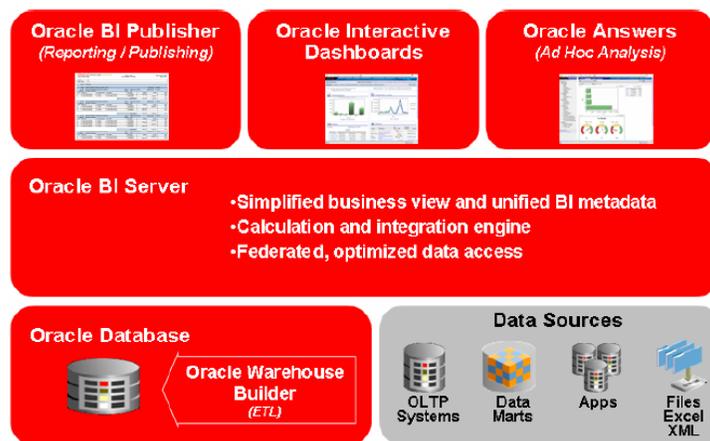


FIGURA 16 ARQUITECTURA DE ORACLE BUSINESS INTELLIGENCE

Esta aplicación cuenta con una serie de módulos disponibles los cuales son:

- BI Server

Oracle BI Server es el motor de la suite, que sirve de repositorio central de información y de cálculo para todo tipo de procesos que consumen información: cuadros de mandos, reporting, análisis OLAP, aplicaciones web, Scorecard, Essbase, etc. Gracias a su potencia, escalabilidad y flexibilidad permite integrar de manera coherente múltiples fuentes de datos, con estructuras y formatos dispares. Provee a toda la organización de acceso a información homogénea que refleja una situación única de la empresa para todos los departamentos y usuarios.

La plataforma ofrece una amplia gama de soluciones de acceso, análisis y entrega de información en un entorno web totalmente integrado. Gracias a esta flexibilidad, se pueden satisfacer las necesidades de información de los distintos agentes y departamentos dentro de la organización, todo ello manteniendo una estructura única y coherente.

- Database

Es un manejador de bases de datos relacionales de amplio uso en el mercado. Se caracteriza por ser de fácil instalación, ser completa, tener capacidad de generar respaldos, administración automatizada de memoria, soporte para distintas fuentes de datos (XML, archivos de texto, imágenes) y de poder ser instalada en distintos sistemas operativos. Dentro de la solución de Oracle bussines intelligence, Oracle Database es la base de datos encargada de registrar todos los datos de los procesos operacionales que ocurren dentro de la organización.

- Warehouse Builder

Oracle Warehouse Builder es una herramienta que permite realizar el diseño completo de un Data Warehouse y luego cargarlo en un motor de base de datos Oracle mediante el diseño de procesos ETL (Extracción Transformación y Carga por sus siglas en ingles). La aplicación incluye una potente interfaz gráfica que permite crear dimensiones, cubos, correspondencias y procesos sin tener que escribir directamente nada de código. Esta herramienta permite la integración de múltiples fuentes de datos. En este sentido soporta distintos tipos de fuentes de datos como archivos de textos, hojas de cálculo, bases de datos relacionales (como Oracle Database), bases de datos multidimensionales, entre otros.

De esta forma, se puede conectar con motores de diferentes compañías o incluso archivos planos o Excel.

Posee una interfaz gráfica para la manipulación de los objetos lo que facilita la realización de las operaciones de extracción, transformación y carga (ETC). Además, se pueden configurar de forma automatizada la carga de datos al almacén de datos, registrar cambios que ocurran en las dimensiones y generar y poblar automáticamente la dimensión tiempo dentro de nuestro esquema. En la figura 17 se observa la aplicación Oracle warehouse builder, en donde a partir de una interfaz gráfica se está generando un modelo estrella.

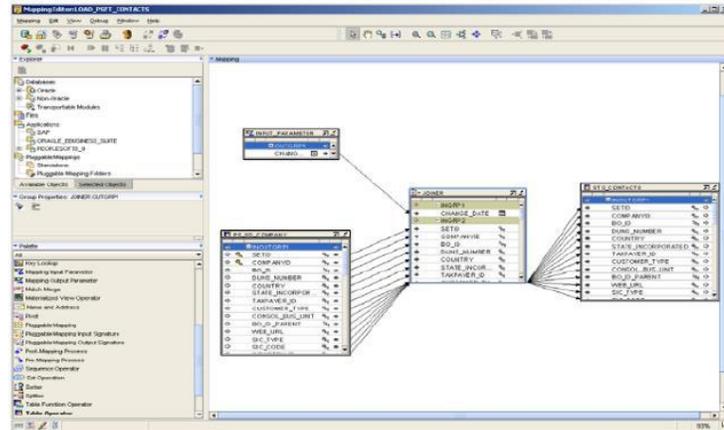


FIGURA 17 INTERFAZ DE ORACLE WAREHOUSE BUILDER

- Dashboards

Oracle BI Dashboards ofrece al usuario un acceso interactivo e intuitivo a la información más relevante de la empresa. Al ser dinámico y personalizable, el contenido ofrecido se ajusta a la perfección a los roles de cada usuario. El objetivo es que el usuario pueda tener acceso a factores clave de un sólo vistazo, ampliando fácilmente la información sobre aquellos aspectos más cruciales.

El usuario interactúa con una gran cantidad de elementos de visualización de datos, que puede configurar y adaptar fácilmente a sus necesidades, todo ello integrado en una auténtica arquitectura web. Mediante la navegación en profundidad puede ir aumentando el nivel detalle en sus consultas, llegando hasta la unidad más básica de anotación, como una factura o un registro de cliente.

Oracle BI Dashboards además puede incorporar en sus visualizaciones otras fuentes de contenidos como urls o aplicativos web, repositorios de archivos compartidos y un largo etc.

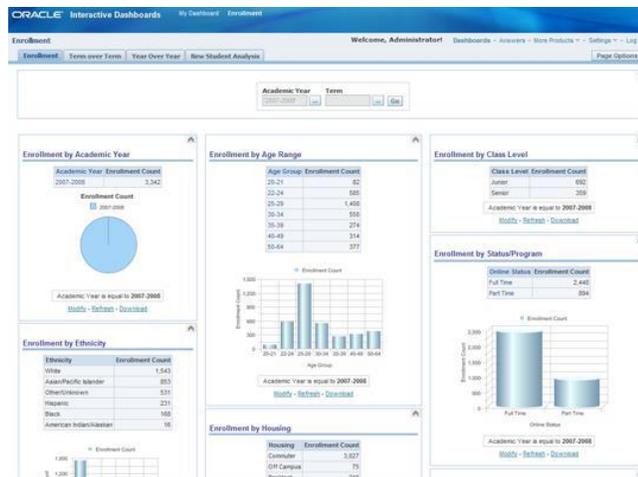


FIGURA 18 INTERFAZ INTERACTIVA DEL DASHBOARDS

- Answers

Oracle BI Answers es una solución de reporting adhoc totalmente orientada al usuario final en un entorno totalmente web. El usuario simplemente interactúa con la información según la lógica del negocio, sin necesidad de conocer las complejas estructuras de datos o de realizar consultas. Esto le permite crear de manera sencilla reports, cuadros de mandos, tablas dinámicas y demás elementos de visualización que pueden ser fácilmente almacenados, reutilizados y compartidos.

Esto permite a los usuarios alcanzar prácticamente una total autosuficiencia en necesidades de información sin necesidades de intervención de personal de IT. En un entorno totalmente seguro y controlado en el que ninguna acción puede afectar a la integridad de los datos de la organización.

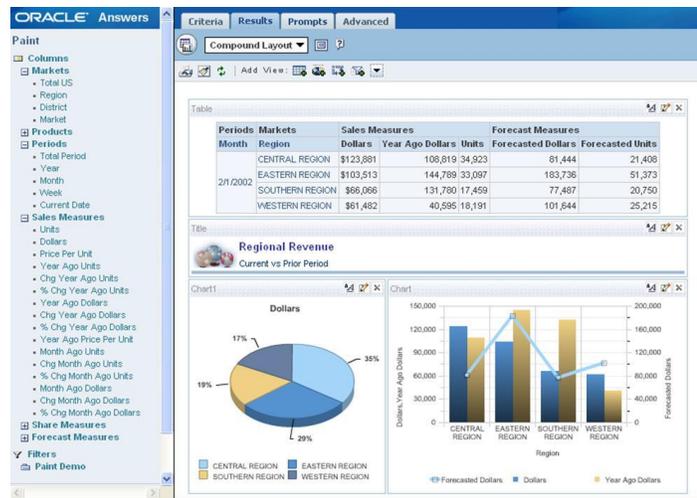


FIGURA 19 INTERFAZ DE OBI ANSWERS

- Publisher

Oracle BI Publisher es una solución de reporting distribuido, eficiente y escalable para entornos complejos.

Ofrece una arquitectura centralizada para generar y distribuir información a empleados, clientes y socios de forma segura y con el formato adecuado.

Los formatos de documento pueden ser fácilmente diseñados con herramientas como Microsoft Word o Adobe Acrobat, herramientas con las que muchos usuarios están familiarizados. Publisher permite integrar múltiples fuentes de datos en un único documento que puede ser distribuido en forma impresa, por mail o incrustado en web.

Al estar totalmente integrado en la suite se aprovecha de las estructuras comunes de elementos visuales, metadatos, cálculo y caché, aumentando todavía más la potencia de la herramienta.



FIGURA 20 EJEMPLO DE INTERFAZ BI PUBLISHER

- Delivers

Oracle BI Delivers es una solución proactiva para la monitorización de la actividad empresarial con servicios de alertas proactivas incorporado que pueden llegar al usuario por múltiples canales.

Incluye un servicio totalmente web de creación y suscripción de alertas. En bases a parámetros fijados por el usuario el sistema rastrea de manera continua toda la información disponible comprobando que los datos no rebasen las condiciones de control especificadas. En caso afirmativo, avisan automáticamente a la persona responsable informándole de la situación e incluso sugiriéndole posibles medidas correctivas.

2.6.3. Tableau

De acuerdo con su página web tableau "es una herramienta de Inteligencia de Negocios que permite visualizar grandes volúmenes de información en forma rápida, flexible y amigable. Tableau es en la actualidad la herramienta de BI líder y de mayor velocidad de crecimiento según Gartner, destacando por su facilidad de uso, potencialidad para generar visualizaciones y capacidad de manejo de grandes volúmenes de Datos". Adicionalmente, y "A diferencia de las herramientas tradicionales de Inteligencia de Negocios (BI) desarrolladas pensando en el usuario técnico del área de sistemas, Tableau está orientado a que personas de todos los

ámbitos puedan manejar información fácilmente y presentarla en forma atractiva. Así, abogados, periodistas, ingenieros, médicos, entre otros, que trabajen en una organización o en forma independiente, encontrarán en Tableau un poderoso aliado analítico”.

Es una plataforma conformada esencialmente por cinco módulos principales que proveen todas las funcionalidades necesarias y básicas para una solución de inteligencia de negocios. La empresa que fabrica dicha solución lleva su mismo nombre y fue creada en el 2003 en California, Estados Unidos. Es una de las soluciones de negocio con mayor crecimiento en el área y en sus últimos años ha sido notorio. Parte de su éxito se debe a la fácil implementación de su solución, costos reducidos de licencia, velocidades de ejecución, así como una rápida realización y puesta en producción de métodos de acceso a la información, lo que permite obtener rápidamente un retorno a la inversión. (Tableau., s.f.)

Esta solución de BI consiste esencialmente en 2 componentes principales, tableau desktop y tableau server. Aplicaciones adicionales como tableau online y tableau public, entre otras, solo ofrecen una forma para visualizar data.

- Tableau Desktop

Es la solución diseñada para instalar en un computador y ejecutar las funcionalidades de inteligencia de negocio desde dicha terminal. Se caracteriza por buscar minimizar los procesos de manipulación de data propios de un almacén de datos. Esto lo logra mediante una tecnología llamada VizQL, la cual es una tecnología diseñada en la Universidad de Stanford que logra que los clic y manipulaciones dentro de los informes se vuelvan directamente en consultas sobre la base de datos, lo cual genera una respuesta gráfica que corresponde a las necesidades analíticas. A su vez, utiliza la ubicación en memoria principal de la data (base de datos en memoria) para aumentar la velocidad de las consultas.

Como se observa en la figura 21, Tableau desktop puede generar casi todas las herramientas de visualización de datos existentes, es decir, tableros de control, gráficos de barra, torta, informes, etc. La manipulación de todos estos elementos gráficos es sencilla y se realiza por lo general a partir de unos cuantos clics sobre la interfaz.

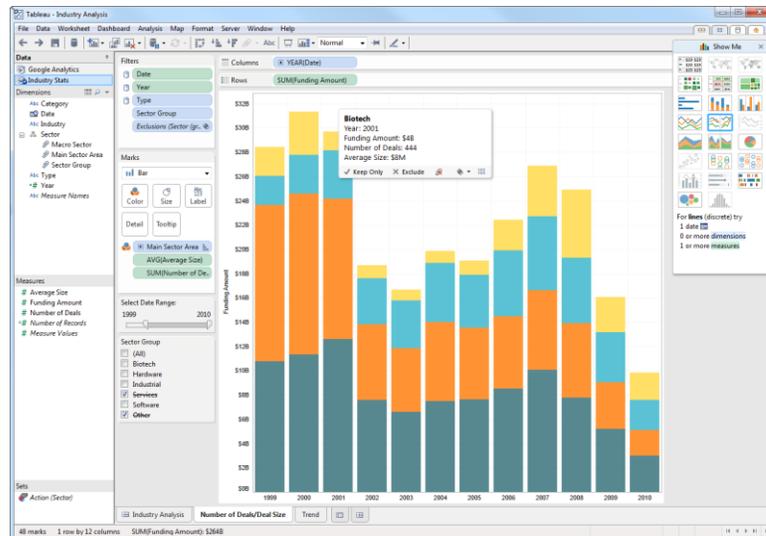


FIGURA 21 EJEMPLO DE INTERFAZ TABLEAU DESKTOP

- Tableau Server

Una vez realizado los tableros de control o cualquier otro elemento dentro de Tableau Desktop, Tableau Server nos permitirá compartir la información a través de nuestra organización. Tableau server además permite realizar análisis embebidos en el sitio donde se está visualizando, es decir, desde el punto de acceso que se utiliza, se dispone de diferentes conectores para manipular data propia de las bases de datos del lugar o manipular los indicadores establecidos. Una característica importante es la capacidad que tiene Tableau Server para establecer el diseño de la interfaz según las medidas de la pantalla del dispositivo (resolución gráfica), esto puede observarse en la ilustración 22, donde, el gráfico, menús e indicadores se adaptan según el tamaño de la pantalla. (Tableau., s.f.)

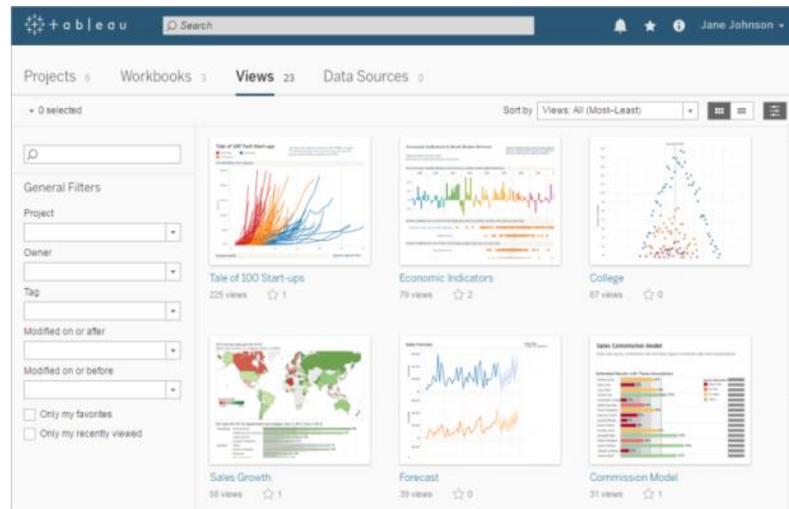


FIGURA 22 EJEMPLO DE INTERFAZ TABLEAU SERVER

2.6.4. Comparación entre las herramientas de inteligencia de Negocios

Criterio	Pentaho	Oracle	Tableau
Panel de control	Pentaho Dashboards	Oracle BI Interactive Dashboards	Tableau Desktop
ETC	Pentaho Data Integration	Oracle Warehouse Builder	-
Reportes	Pentaho Reporting	Oracle BI Publisher	Tableau Desktop
Análisis	Pentaho Analysis Services	Oracle BI Answers	Tableau Server
Licencia	Pública y Privada	Privada	Pública y Privada

TABLA 3 COMPARACIÓN ENTRE LAS HERRAMIENTAS DE INTELIGENCIA DE NEGOCIOS

2.7. El seguro

El seguro constituye la forma más perfecta y técnicamente eficaz para la cobertura de riesgos y transfiriéndolos a una organización estructurada con la técnica y operativa adecuadas para garantizar su compensación, en caso de ocurrir el evento. Se entiende por seguro como:

“El mecanismo para reducir la incertidumbre de una parte llamada el asegurado, por medio de la transferencia de ciertos riesgos a otra parte, llamada el asegurador, quien ofrece una reposición, al menos parcial, de las pérdidas económicas sufridas por el asegurado” (Pfeffer, 1956).

Se desprenden tres elementos básicos que componen el seguro y una finalidad principal. Los elementos básicos son: reducción de la incertidumbre, transferencia de ciertos riesgos (aunque no todas las causas de pérdida son trasladables al asegurador) e indemnización de una pérdida económica y su finalidad principal es transformar incertidumbre en certidumbre proporcionando seguridad al asegurado.

- Jurídico: “El seguro es un contrato sustantivo y oneroso por el que el asegurador asume el riesgo de que pueda ocurrir un acontecimiento incierto, al menos en cuanto al tiempo, obligándose a cubrir la necesidad pecuniaria sentida por la otra parte (el asegurado) a consecuencia de ese riesgo determinado en el contrato” (Garrigues, 1936).
- Económico y social: “Requiere la existencia previa de un grupo de personas u objetos sometidos al mismo peligro en el que todos sus miembros asuman el compromiso de contribuir proporcionalmente para subsanar los perjuicios económicos que sufran aquellos a los que afecten los siniestros. Esta capacidad se puede interpretar como una fórmula colectiva de protección frente al riesgo” (Albarrán, 2008)

El seguro conlleva adicionalmente otros servicios tan importantes como la ayuda al ahorro, estimulando las inversiones familiares, asistencia técnica, asistencia médica, entre otros. Adicionalmente, el servicio que presta el seguro no se refiere sólo a la entrega de unas sumas de dinero para adquirir el derecho a percibir otras, sino que hay además un conjunto de servicios integrados en él y que van desde el asesoramiento inicial en materia de riesgos (que compete principalmente al mediador de seguros, sea corredor o agente) hasta la peritación, pasando por actividades importantísimas como las de prevención de riesgos. Finalmente, se puede decir que la justificación del Seguro y base fundamental de la actividad aseguradora se encuentra en la transferencia, reparto y agrupación de los riesgos y que su esencia como negocio está en transformar riesgos (sucesos aleatorios, inciertos y de cuantía imprevisible) en primas (pagos ciertos, periódicos y presupuestables).

Entiéndase como prima el precio del seguro, el valor de la obligación del contratante o asegurado como contrapartida de las obligaciones y contraprestaciones que recibirá el asegurado, en caso de darse los supuestos establecidos en la póliza. (Albarrán, 2008)

2.7.1. Funciones del seguro

Los empresarios, en el desarrollo de su actividad productiva, se enfrentan a numerosos riesgos. El coste del seguro incrementa los costes generales pero, a cambio, proporciona

mayor estabilidad en los resultados. El seguro tiene un marcado carácter de servicio público por lo que, al definir las funciones del seguro, se han considerado sus efectos sobre la economía y la sociedad, teniendo en cuenta aspectos sociológicos, económicos y laborales. (Albarrán, 2008

Funciones económicas del seguro:

- Contribuye positivamente al desarrollo económico al eliminar riesgos y estabilizar los presupuestos económicos. Por esto, debe desarrollarse paralelamente al resto de las actividades económicas.
- El seguro es la única actividad económica que posee capacidad para generar ahorro y financiación de inversiones a largo plazo. Existen otras instituciones financieras que aportan ahorro a largo plazo pero sólo el seguro lo hace con un esquema de ahorro y financiando un tipo de inversión (global y sistemática) sustancialmente distintos a los utilizados habitualmente por otros intermediarios.

Funciones laborales del seguro:

- El seguro participa en la consecución de empleo directo e indirecto. En la actualidad muchas familias viven del seguro (empleados, corredores, peritos, liquidadores, abogados, actuarios y otros profesionales), por lo que este sector está financiando un gran número de puestos de trabajo.

2.7.2. Contrato del seguro

El contrato de seguro es aquél en virtud del cual una empresa de seguros, a cambio de una prima, asume las consecuencias de riesgos ajenos, que no se produzcan por acontecimientos que dependan enteramente de la voluntad del beneficiario, comprometiéndose a indemnizar, dentro de los límites pactados, el daño producido al tomador, al asegurado o al beneficiario, o a pagar un capital, una renta u otras prestaciones convenidas, todo subordinado a la ocurrencia de un evento denominado siniestro, cubierto por una póliza. (Decreto N° 1505, 2001)

Las disposiciones del contrato de seguro se aplicarán a los convenios mediante los cuales una persona se obliga a prestar un servicio o a pagar una cantidad de dinero en caso de que ocurra un acontecimiento futuro e incierto y que no dependa exclusivamente de la voluntad del beneficiario a cambio de una contraprestación, siempre que no exista una ley especial que los regule.

2.7.3. Clasificación de los seguros

- Seguro de personas: Es un camino fácil y prudente para preparar de antemano la ayuda apropiada a las pérdidas económicas surgidas de los peligros de la vida cotidiana, tales como la muerte, accidentes, enfermedades y la vejez. Dicho de otra manera, es un mecanismo que reduce el impacto financiero negativo causado por sucesos fortuitos que impiden concretar las expectativas.

- Seguro de automóvil: Un seguro de auto es un contrato con una Compañía de Seguros en el que ésta se compromete a pagar el monto asegurado en caso de choque, robo y daños a terceros según las coberturas incluidas.

- Seguro patrimonial: Bajo esta denominación se recogen todos los seguros cuyo fin principal es reparar la pérdida sufrida, a causa de siniestro, en el patrimonio del tomador del seguro. Son elementos esenciales de los seguros de daños: el interés asegurable, que expresa la necesidad de que el tomador del seguro tenga algún interés directo y personal de que el siniestro no se produzca, bien a título de propietario, usuario, entre otros., y el principio indemnizatorio, según el cual la indemnización no puede ser motivo de enriquecimiento para el asegurado y debe limitarse a resarcirle del daño concreto y real sufrido en su patrimonio.

- Seguro de fianza: Es un contrato por el cual una persona llamada Fiador, se obliga frente a otro acreedor, a cumplir las obligaciones asumidas por el Afianzado, si este no las cumple. El Fiador se responsabiliza por las obligaciones convenidas por el Afianzado

2.7.4. Póliza de seguro

La palabra póliza, deriva del italiano póliza, es el instrumento en el que se hace constar la totalidad de las condiciones y formalidades de un contrato. Comúnmente se refiere al contrato de seguro, aunque también puede aplicarse a otros tipos de contratos. Las pólizas, pueden ser nominativas, a la orden, o al portador. (Albarrán, 2008)

2.7.5. El Perito

La labor del perito es el engranaje esencial de una compañía de seguros, pero para conocer la verdadera dimensión del trabajo del perito, se analizan sus funciones, que se resumen en tres aspectos principales:

Aspectos técnicos:

- Valoración económica de los daños, elaborando la peritación y realizando la propuesta de indemnización a la compañía de seguros. Determinación del valor del bien asegurado, como, por ejemplo, el valor de mercado, el valor de los restos y la propuesta del importe para la indemnización, cuando se ha producido un siniestro total o una pérdida total.
- Verificación de siniestros, para la realización de informes de uso interno para la compañía de seguros con la justificación técnica de la ocurrencia del siniestro. Pueden ser informes de rehúses parciales o totales, que pueden aportarse como prueba en un juicio.
- Los informes de reconstrucción de accidentes de tráfico, a partir de huellas y vestigios, mediante cálculos físicos y matemáticos, pueden ser también un apoyo para la determinación de la culpabilidad en el juicio.
- Revisión de riesgos, para la contratación de nuevas pólizas de vehículos de segunda mano con coberturas de daños propios.
- Control de calidad de la reparación, mediante la comprobación, en primer lugar, de que la reparación se ha llevado conforme a la peritación en todas y cada una de las partidas asignadas por el perito; a continuación, que la reparación se ha realizado con las debidas garantías técnicas, de calidad y seguridad para los ocupantes del vehículo. Por último, se analizarán los defectos en la reparación, para que sean subsanados por el taller.
- Averías mecánicas: valoración y peritación de los daños mecánicos bajo la cobertura de pólizas de vehículos de renta y de pólizas de garantía de venta de vehículos usados.

Aspectos administrativos-legales:

- Implicación en la tramitación del siniestro. El perito, en contacto con el tramitador y a través del sistema de gestión de la compañía de seguros, está al día de la tramitación de los siniestros, del tipo de pólizas que comercializa la compañía de seguros, de sus coberturas y exclusiones, de los convenios entre compañías y del conocimiento de la legislación de seguros.

Aspecto negociador:

- El perito es la imagen de la compañía de seguros, ya que está en contacto con los asegurados, perjudicados, talleres, otras compañías, entre otros., con lo que su actuación está sujeta a examen continuo, y su comportamiento, a ojos del asegurado, es, por extensión, el de la compañía de seguros.
- El perito debe aportar, en todo momento, argumentos y criterios técnicos en la negociación con el taller. Ha de consensuar la peritación: debe llegar a acuerdos con el taller sobre todas y cada una de las partidas que componen una peritación.
- Realiza asesoría legal: al estar en contacto con los asegurados y el taller, entre otros. en muchas ocasiones, el perito se convierte en el asesor sobre los aspectos legales de los siniestros. (Hernández, 2012)

2.8. Introducción a la descripción del flujo asociado al proceso de solicitud y atención de la empresa de seguros.

Para poder explicar el proceso de trámite para la inspección de vehículos en los centros de inspección pertenecientes a la compañía aseguradora, se tiene en la actualidad la separación del mismo en un esquema representado por dos (2) subprocesos (Proceso de Solicitud y Proceso de Atención) el cual se puede apreciar en la Figura 23:

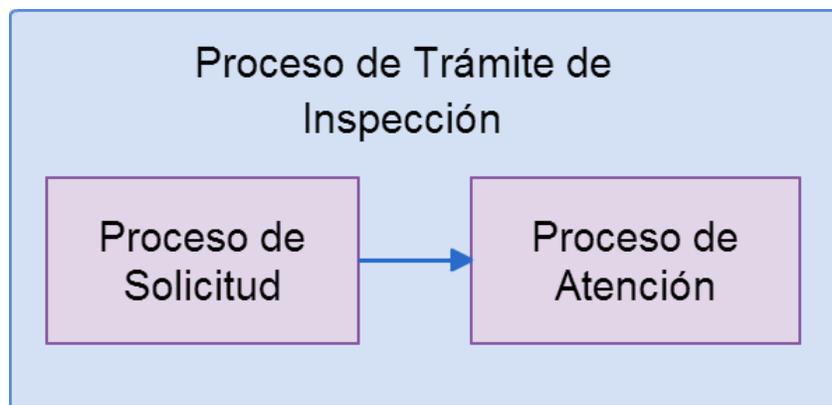


FIGURA 23 PROCESO DE TRÁMITE DE INSPECCIÓN.

El concepto del diagrama representado por la Figura 23 está asociado a que inicialmente para poder realizar la inspección de vehículos, todos los clientes deberán realizar el Proceso de Solicitud. En donde tendrán que solicitar sus números de orden ingresando al sistema con la finalidad de obtenerlo para asistir a un centro de inspección.

Una vez realizado el Proceso de Solicitud, los clientes tendrán que presentarse en el centro de inspección para el cual solicitaron el número de orden a la fecha y hora que hayan indicado. En donde serán recibidos por los empleados encargados de atender la taquilla con la finalidad de registrar la asistencia y asignar los peritos. Finalmente, los clientes que hayan sido atendidos deberán esperar a ser llamados por el perito que se les haya asignado para realizar la inspección de vehículos.

Tanto el subproceso de Solicitud como el de Atención se encuentran detallados por los flujogramas representados por las Figura 24 y la Figura 25 que se presentan en los puntos a continuación.

2.8.1. Proceso de Solicitud

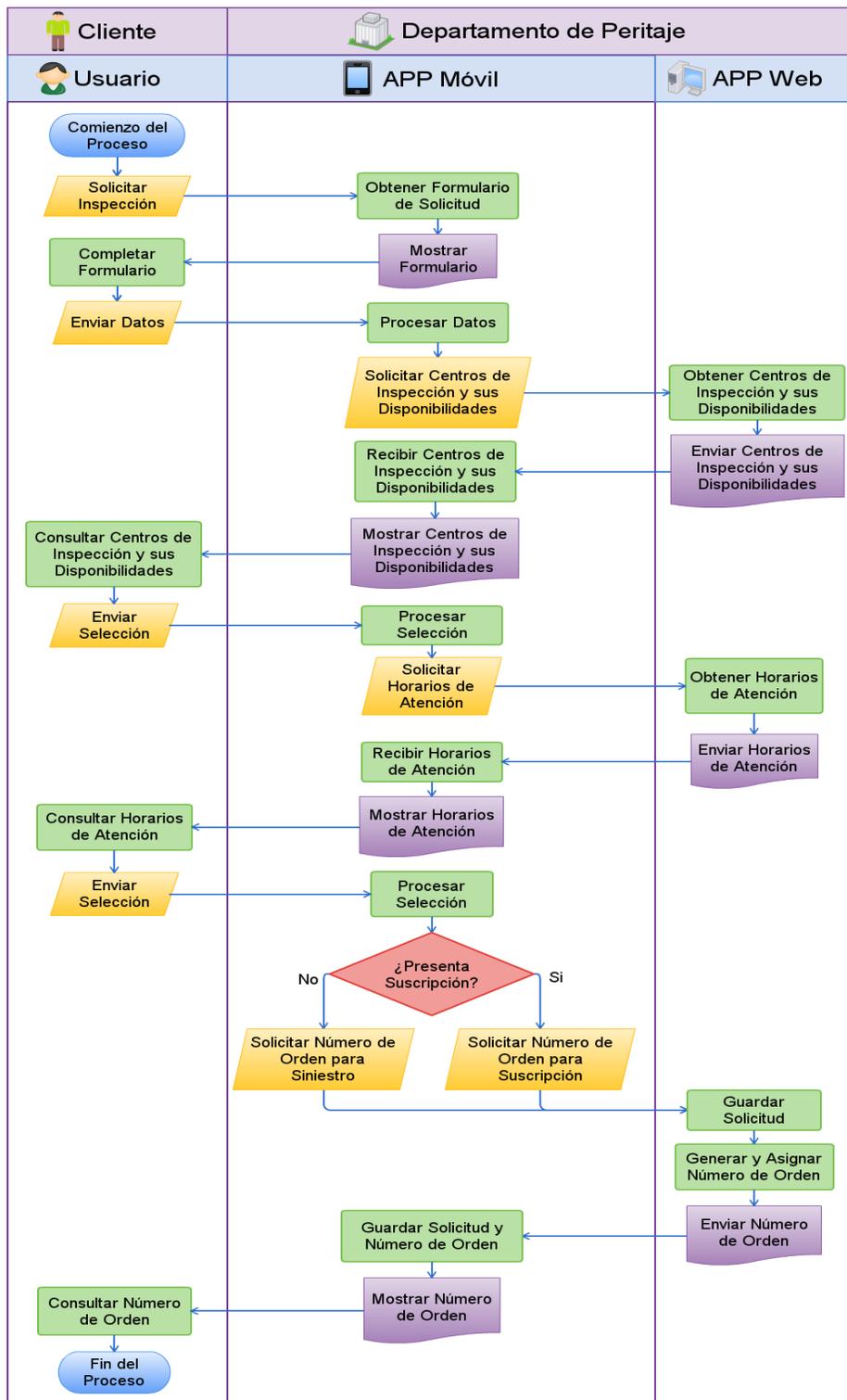


FIGURA 24 FLUJO DEL PROCESO DE SOLICITUD.

De acuerdo con la Figura 24, cuando un cliente necesita realizar una inspección a su vehículo, el mismo deberá acceder a su aplicación móvil realizando una autenticación previa. Posteriormente, podrá realizar una solicitud conectándose directamente con su dispositivo móvil a una aplicación web perteneciente a un servidor corporativo. Para ello deberá acceder por medio de la aplicación móvil al formulario de la solicitud, en donde indicará el tipo de solicitud y la fecha de asistencia junto con el estado y municipio del centro de inspección al cual desee asistir, con esto la aplicación móvil le solicitará a la aplicación web un listado de los mismos junto con sus disponibilidades y así el cliente podrá seleccionar el que le sea más conveniente. A partir de esta información la aplicación web también solicitará por medio de un servicio web los horarios de atención disponibles para el centro de inspección seleccionado por el usuario, para que el mismo pueda visualizarlos e indicar el que sea de su preferencia a fin de enviar la solicitud a la aplicación web.

Una vez recibida la solicitud, la aplicación web procede a guardarla para posteriormente generar y asignar el número de orden de acuerdo al tipo de solicitud establecido por el usuario, luego se enviará su información por medio de un servicio web a la aplicación móvil con la finalidad de registrar la solicitud y la información de su número de orden. Finalmente, el usuario podrá consultar la información de su número de orden y con el mismo podrá asistir al centro de inspección a fin de ser atendido y realizar la inspección de su vehículo.

Listado de excepciones:

- Si al momento de completar el formulario de solicitud, el cliente no está conforme con ninguno los centros de inspección que se le presentan, el mismo tendrá la opción de regresarse un paso para completar nuevamente el formulario de solicitud y obtener una nueva lista con otros centros de inspección.
- Si al momento de seleccionar un centro de inspección, el cliente no está conforme con ninguno de los horarios de atención que se le presentan, el mismo tendrá la opción de regresarse un paso para seleccionar otro centro de inspección y obtener su lista de horarios de atención. En caso contrario, también podrá volver completar nuevamente el formulario de solicitud para obtener una nueva lista con otros centros de inspección a fin de consultar sus horarios de atención.
- El cliente también tiene la opción de cancelar el procedimiento para la creación de una solicitud durante cualquiera de sus pasos.

- Una solicitud solo puede ser creada cuando hay conexión a internet, en caso contrario se le desplegará al cliente una ventana notificando la cancelación de la solicitud al momento de su creación.

De acuerdo con la Figura 25, Se tiene que cuando un cliente adquiere su número de orden por medio del proceso de Solicitud, el mismo tendrá que presentarse en el centro de inspección a la fecha y hora que indicó en un comienzo, en donde acudirá a un operador de taquilla para que pueda registrar su asistencia en la aplicación web, actualizar el estado de su número de orden como "Atendido" y asignarle un perito. Finalmente, deberá esperar hasta que sea llamado por el perito que le fue asignado para realizar la inspección de su vehículo. En caso contrario, si el cliente no se presenta en el centro de inspección, entonces el operador de la taquilla registrará su inasistencia en la aplicación web actualizando el estado de su número de orden como "No procesado". De tal forma que se le enviará una notificación junto con una encuesta de justificación, para que pueda indicar la razón de su inasistencia con la finalidad de brindar un soporte informativo que contribuya a futuras mejoras en el desarrollo del proceso.

De igual manera, luego que un número de orden actualiza su estado como "Atendido", la aplicación web enviará una notificación al cliente junto con una encuesta de satisfacción para que pueda indicar su nivel de satisfacción con respecto al servicio que le fue prestado, con la finalidad de utilizar esta información para incorporar mejoras en el proceso actual.

CAPÍTULO 3 MARCO METODOLÓGICO

Cuando se planifica el desarrollo de una solución de inteligencia de negocios es indispensable tener presente los diferentes componentes que vienen incluidos con este proceso y que incluyen la selección de metodologías, técnicas y herramientas que se ajusten a las necesidades de la organización sobre la cual se implementa pero que, adicionalmente, presente resultados no solo confiables sino además que concluyan exitosamente. Hacer caso omiso a la relevancia de estos factores puede generar que el desarrollo de dicha solución se vuelva más complicado de lo que realmente es o incluso no se logren los objetivos planteados.

Es por esta razón que es tan importante revisar diferentes metodologías para analizar cuál de ellas se ajusta mejor al esquema y arquitectura de dicha solución ya que es un requisito fundamental. De igual forma la elaboración de esta solución de inteligencia de negocios conlleva el desarrollo de un almacén de datos debido a que juega un papel esencial en la infraestructura de esta y el desarrollo abarca desde la captura de datos provenientes de la fuente hasta la integración de estos de forma tal que permitan el conjunto de análisis pautados.

Como ya lo hemos mencionado, hablar sobre el desarrollo una solución de inteligencia de negocios implica abordar temas relacionados a un almacén de datos, y en este sentido son dos las metodologías de desarrollo que por excelencia destacan en dicho plano, Ralph Kimball y Bill Inmon, en los que podemos hallar puntos de encuentro en común, pero con filosofías de trabajo opuestas que van en la búsqueda de lograr el mismo objetivo, el desarrollo de un almacén de datos.

Sin embargo, una vez revisadas dichas metodologías, debido a la naturaleza del proyecto y a la forma de adaptarse a este, se seleccionó la metodología propuesta por Ralph Kimball gracias a los cortos períodos de tiempo para ser implementada, brindar una solución completa y versátil.

3.1. Metodología de desarrollo según Ralph Kimball (Bottom-up)

El desarrollo y ejecución de un proyecto de inteligencia de negocios involucra la implementación de almacén de datos, este, puede seguir el mismo ciclo de desarrollo que

cualquier otro proyecto de desarrollo de software incluyendo los requerimientos, fase de análisis, fase de diseño, fase de construcción, pruebas e implantación de este. Sin embargo, Considerando la complejidad de construcción de un almacén de datos han surgido diferentes metodologías de desarrollo que proponen comenzar con la construcción de bodegas de datos e ir creciendo hacia un almacén o ver el almacén como un todo. Existen dos enfoques principales de metodologías para el diseño de un almacén de datos, estas son resumidas a continuación en la Tabla 4:

	Bottom-Up	Top-Down
Autor	Ralph Kimball	Bill Inmon
Enfoque	Bodega de datos	Almacén de datos
Diseño	Modelo dimensional de bodega de datos, esquema estrella	Modelo normalizado basado en la empresa
Arquitectura	Área de interés y mercado de datos	Compuesto de varios niveles de áreas de interés y mercado de datos dependientes
Data	Contiene datos atómicos y sumariados	Almacén de datos: Datos a nivel atómicos. Bodega de datos: Datos sumariados

TABLA 4 COMPARACIÓN ENTRE LAS METODOLOGÍAS DE DESARROLLO DE UN ALMACÉN DE DATOS

En el caso de este trabajo especial de grado se pretende hacer uso de la metodología con el enfoque propuesto por Ralph Kimball, esta metodología llamada Modelo Dimensional (Dimensional Modeling), está basado en lo que él denomina Ciclo de Vida Dimensional del Negocio o "Business Dimensional Lifecycle" (BDL) siendo esta una de las técnicas más utilizada por los profesionales del diseño de los almacenes de datos; el correcto desarrollo de cada una de las fases que plantea la metodología asegura el correcto proceso del desarrollo del proyecto, asegurando también la garantía de la calidad del producto, por lo que para el presente trabajo especial de grado se usó dicha metodología para llevar la gestión de todo el proyecto de inteligencia de negocios.

El diseño de un Data Warehouse propuesto por Kimball es descrito en la Figura 26, donde se muestra el diagrama que resume las distintas fases, dentro de la cual las líneas horizontales indican el flujo entre actividades mientras que el alineamiento vertical implica la dependencia entre tareas. Las fases de esta metodología serán especificadas en las secciones subsiguientes.

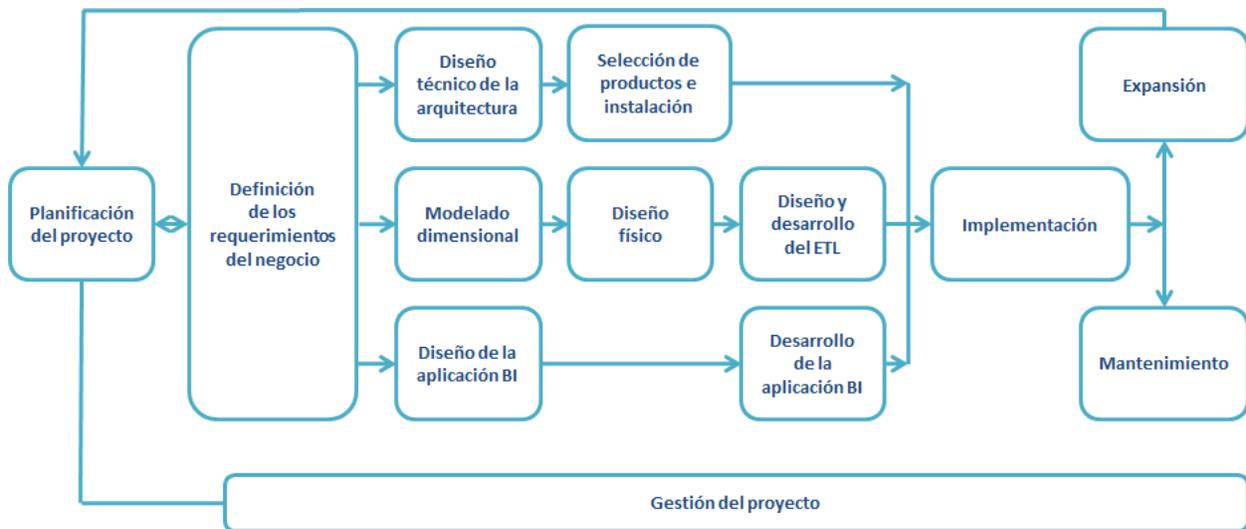


FIGURA 26 BUSINESS DIMENSIONAL LIFECYCLE DIAGRAM (KIMBALL & ROSS. 2002)

3.1.1. Planificación del proyecto

La planificación de un proyecto orientado a la construcción de un almacén de datos busca identificar la definición y el alcance que tendrá el proyecto de almacén de datos, al ser la primera etapa, está centrada en la definición del proyecto, donde, a nivel de planificación, se establece la identidad del mismo, el posible personal, desarrollo del plan de proyecto, el seguimiento y la monitorización, sin embargo, para que esto se lleve a cabo, es necesario identificar el escenario en el cual se planifica el proyecto, el cual puede ser:

- Demanda de un sector del negocio: Un ejecutivo del negocio busca obtener mejor información para la toma de decisiones.
- Demanda alta de información: Existen múltiples ejecutivos de negocios buscando obtener la mejor información para la toma de decisiones.
- Demanda incierta: El encargado de la empresa no identifica correctamente las necesidades de la empresa para con el almacén de datos.

Una vez identificado el escenario se pueden determinar los requisitos del proyecto (de ser así) y el origen de dicha demanda, en este sentido es necesario chequear si de verdad la empresa está preparada para la ejecución del proyecto, para esto se identifican cinco factores que se deben tomar en cuenta:

- Patrocinio de la gerencia del negocio: Los gerentes son influyentes dentro de la organización y determinarán el apoyo y soporte al proyecto de los demás miembros

dentro de la organización, por esta razón siempre es preferible tener varios patrocinadores y no uno solo.

- **Motivación del negocio:** Implementar un almacén de datos busca satisfacer una necesidad por parte de la organización, causado por una motivación del negocio, como por ejemplo la competencia y la visión competitiva o un mercado potencial.
- **Acompañamiento del departamento de tecnología y de negocio:** Esto es un factor clave para el éxito de un proyecto de almacén de datos y se produce gracias a la colaboración y partición de responsabilidades con los expertos en el área de la tecnología.
- **Presencia de cultura analítica:** Es importante que las decisiones de la organización se basen en hechos, más que en simples intuiciones y que estas decisiones sean determinantes y recompensadas.
- **Factibilidad:** Es preferible que la infraestructura que soporte el almacén de datos esté presente y sea robusta (responda adecuadamente a los posibles errores que se presenten), otro hecho que garantiza la factibilidad es que los datos operacionales estén filtrados y cumplan con ciertos estándares sino el proyecto tendrá retrasos respecto al cronograma planeado.

3.1.2. Definición de los Requerimientos del negocio

Esta fase indica los factores claves que guían el negocio para determinar efectivamente los requerimientos y traducirlos en consideraciones de diseño apropiadas. Antes de iniciar el proceso de levantamiento de requerimientos, se deben analizar los reportes anuales de la compañía, para determinar las decisiones y hechos estratégicos, asimismo resulta útil analizar las fortalezas y debilidades de la compañía, así como si ha existido un intento anterior de desarrollar un almacén de datos.

Es por todo lo antes expuesto que esta fase es el centro de gestión de nuestro almacén de datos ya que funciona como punto inicial de desarrollo de este, y en donde podemos definir el modelo de datos, alcance, características de dicho modelo y requerimientos de carácter analítico para los usuarios de este.

Luego de esto se debe hacer una entrevista con las personas involucradas, estas deben ser personas representativas de cada área de la organización, se deben determinar roles y responsabilidades en el equipo entrevistador. Es preferible que el mismo equipo conduzca las entrevistas a usuarios del negocio y al equipo de tecnología de la empresa. Es importante

observar el organigrama de la compañía para determinar los candidatos a entrevista. Los principales entrevistados deben ser los administradores ejecutivos del negocio, para comprender la estrategia en un alto nivel de la empresa.

Luego es importante entrevistarse con los analistas del negocio de cada área, o bien del área sobre el cual se pretende realizar dicho proyecto, quienes conocen el manejo de información que se lleva a cabo. Ya en la entrevista se deben desarrollar varios cuestionarios que serán aplicados dependiendo del rol de los entrevistados dentro de la empresa. Finalizado esto, se deben analizar, repasar los reportes y análisis reunidos en las entrevistas, lo cual comúnmente conlleva a una aproximación del descubrimiento de dimensiones para el modelo, para esto es importante que todos los requerimientos obtenidos sean documentados y comunicarlos a los usuarios para adquirir su aprobación y compromiso.

3.1.3. Diseño de la arquitectura

En esta fase se deben tener en cuenta tres factores: los requerimientos de negocio, los actuales entornos técnicos, y las directrices técnicas y estratégicas futuras planificadas por la compañía, lo que permitirá establecer el diseño de la arquitectura técnica del entorno del almacén de datos. Para hacer el diseño de la arquitectura se debe comenzar analizando los sistemas actuales, estos deben ser consistentes y manejar de forma correcta sus transacciones, pues en la metodología del desarrollo de Kimball toma como hecho que estos sistemas son confiables.

Es importante además tomar en cuenta que la infraestructura es responsable de darle soporte a la ejecución del proyecto del almacén de datos, esto incluye, por supuesto, hardware, redes, seguridad y cualquier otra función requerida para la presencia del almacén de datos.

En los sistemas de información la definición de una arquitectura permite hacer un desarrollo más confiable y eficiente. Con la definición de la arquitectura se mejora la comunicación entre las diferentes áreas del proyecto, el planeamiento del proyecto, la flexibilidad y el mantenimiento del mismo.

3.1.4. Modelo Dimensional

El modelo dimensional es una técnica de diseño lógico que busca presentar los datos de una forma intuitiva y que proporcione acceso de alto desempeño. Cada modelo dimensional se compone de una tabla con múltiples llaves foráneas, llamada tabla de hechos y un conjunto de tablas más pequeñas, llamadas tablas de dimensión.

Es importante resaltar que los requerimientos previamente establecidos le darán forma al modelo dimensional que se implemente ya que este pretende responder a las necesidades de los usuarios, y la fuente de datos definen la granularidad de dicho modelo, los indicadores que se pretenden construir, las dimensiones del esquema, jerarquías y atributos.

Los atributos de las tablas de dimensión son las fuentes de las restricciones de búsqueda necesarias para consultar una bodega de datos. Son utilizadas como título de atributo de las filas resultantes de consultas de SQL. Existen dos modelos dimensionales que predominan en las soluciones de data Warehouse como lo son el modelo estrella y el modelo copo de nieve.

También debemos entender que toda área de negocio demanda requerimientos propios asociados a las necesidades de información y por tal motivo existen modelos dimensionales departamentales (bodegas de datos) que pueden responder a estas necesidades dentro de la organización. Una buena práctica de esto suele ser la implementación de una matriz de dimensiones ya que les permite a las diferentes áreas del negocio determinar cuáles son las dimensiones que necesita en un determinado análisis.

3.1.5. Diseño de la aplicación BI

En esta fase se involucran las configuraciones de los metadatos y construcción de reportes específicos. La especificación de la aplicación final describe las plantillas de los reportes, los parámetros que manejarán el usuario y los cálculos que sea requerido por este.

3.1.6. Selección de productos e instalación

En esta fase se evalúan y seleccionan cuáles son los componentes necesarios específicos de la arquitectura (plataforma de hardware, motor del BD, herramienta de ETL, plataforma de inteligencia de negocio, entre otros). Luego de realizar la instalación de los componentes

previamente evaluados y seleccionados, estos deben ser instalados y probados para asegurar la integración apropiada de los mismos en el ambiente próximo de trabajo.

3.1.7. Diseño físico

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

3.1.8. Diseño y desarrollo del ETL

Este proceso se realiza para la creación del almacén de datos, por tanto, se debe seguir un plan para su correcto desarrollo, para esto se establecen varios pasos que conducen al desarrollo del proceso:

1. Plan de alto nivel: El proceso de diseño se inicia con un esquema simple de los componentes del plan que son conocidos: Las fuentes y los destinos de los datos. Se identifica de donde provienen los datos y las características y problemas con dichas fuentes. Con este esquema es posible comunicar la complejidad del proyecto a la gerencia y miembros del equipo de desarrollo del proyecto.
2. Herramientas ETL: Las extracciones típicamente se escriben en el lenguaje de la fuente de los datos. Existen herramientas que realizan todo el proceso de extracción, transformación y carga que buscan minimizar el tiempo requerido para estas tareas.
3. Plan detallado: El plan se inicia seleccionando las tablas en las que se va a trabajar, en cual orden y secuenciar las transformaciones para cada conjunto de datos. Se debe graficar un diagrama con estas estructuras.
4. Poblar una tabla de dimensión simple: La principal razón para iniciar el proceso con una dimensión estática y simple es la facilidad para poblar esta tabla.
5. Implementación de la lógica del cambio de una dimensión: Al cambiar los datos de una dimensión, es preferible construir la extracción de tal forma, que se extraigan únicamente los datos que han cambiado. Al determinar los cambios, se debe contar con reglas del negocio que determinen como manejar estos cambios en los atributos.

Si se determina que la modificación permite actualizar el dato, se utiliza la técnica de una dimensión cambiante.

3.1.9. Desarrollo de la aplicación BI

En esta fase se siguen las especificaciones de la aplicación, el desarrollo de la aplicación del usuario final, las herramientas de configuración y la construcción de reportes.

3.1.10. Implementación

Esta fase representa el desarrollo de la convergencia de todas las tecnologías, datos y la aplicación de presentación final que sea accesible directamente por los usuarios. Estas pueden ser de cuatro tipos:

- **Basadas en Web:** Estas aplicaciones son accedidas a través de un browser de Internet. Los usuarios podrían conectarse y ver los reportes vía intranet o Internet entrando a la aplicación
- **Herramienta independiente:** Con la herramienta se diseñan algunas plantillas de reportes que el usuario va a poder acceder a través de una interfaz. Estos reportes son muchas veces almacenados en archivos compartidos para que todas las personas la puedan acceder.
- **Herramienta de interfaz ejecutiva:** Proporciona una estructura de acceso a las plantillas de reportes a través de una serie de interfaces. Estas implementaciones permiten fácilmente la navegación en la plantilla escogida.
- **Interfaz por código:** Estas herramientas proporcionan un API que permite diseñar una interfaz. Esta es una buena posibilidad, pues se utiliza una herramienta de desarrollo gráfico y la navegación se puede ajustar mejor a las necesidades del usuario.

3.1.11. Expansión

Cuando una empresa adquiere sus sistemas de información el cambio que tendrán estos sistemas es muy poco, sin embargo, cuando se desarrolla un proyecto de almacén de datos se debe pensar en el mantenimiento posterior a la implementación, pues estas aplicaciones

tienen gran tendencia a crecer a medida que crece la información de la organización. La inversión en el mantenimiento del almacén de datos es bastante importante, sin embargo, los beneficios de contar con un almacén supera con interés los posibles costos del mismo.

Como fase posterior a la implementación, el crecimiento forma parte del proceso evolutivo de un almacén de datos en el cual se puede evaluar el éxito que este ha tenido al ser implementado, si ha brindado una mejora en el proceso de toma de decisiones y si ha traído consigo beneficios cuantificables para la organización. Esta fase puede, incluso, involucrar la ampliación del almacén de datos trayendo consigo la agregación de un nuevo proceso de negocio dentro de la organización o la solicitud de nuevos requerimientos. Esto acarrea consigo una nueva iteración del ciclo de desarrollo y por lo tanto una nueva expansión.

3.1.12. Mantenimiento

Fase donde los objetivos planteados originalmente se han cumplido pero que involucra una evaluación sobre el rendimiento del sistema y los procesos de extracción, transformación y carga, se mide la eficiencia para generar informes, la carga de los datos adecuadamente y la correspondencia de estos. Se le sigue brindando soporte a la organización sobre lo que ya se ha desarrollado y se ahonda en el adiestramiento sobre el uso de dicho proyecto.

CAPÍTULO 4

MARCO APLICATIVO

En el presente trabajo especial de grado se pretende, a través de la metodología de Ralph Kimball, realizar la implementación de una solución de inteligencia de negocios que se adapte a las necesidades solicitadas, que pueda llevarse a cabo en un período de tiempo corto y que adicionalmente brinde los resultados esperados a petición de quien lo solicita.

Para lograr esto es necesario describir el conjunto de actividades que se deben llevar a cabo hasta llegar a la solución planteada previamente, estas se encuentran divididas en un conjunto de fases contempladas en el capítulo que antecede y que van desde la planificación del proyecto, el diseño del almacén de datos y la elaboración de los procesos de extracción, transformación y carga de los datos al almacén, hasta la realización de consultas analíticas para el desarrollo y visualización de los indicadores, y así finalmente llegar a la emisión de un panel de control a través de un portal web que contendrá un tablero donde podrán verse reflejados los indicadores solicitados, los cubos de análisis para esto y los diferentes reportes que se puedan emitir relacionados a los requerimientos solicitados sobre el proceso de solicitud de citas para la inspección de vehículos.

4.1. Fases del Proyecto

4.1.1. Planificación del Proyecto

De acuerdo con la metodología seleccionada es la primera fase de diseño de la solución en donde se deben establecer todas las actividades que se pretenden desarrollar y cuáles serán las normas a seguir para esto, además permite establecer una línea de ejecución de la solución que se encuentra fundado en el ciclo de vida dimensional del negocio determinado por dicha metodología.

Según lo antes expuesto en el Capítulo I, existe la necesidad por parte de la empresa seguros Horizonte, tener una herramienta que permita generar reportes analíticos en el menor tiempo posible, para crear información de soporte a la toma de decisiones. Necesidad que se le suma al añadido de poseer un nivel de análisis mayor a través de la implementación de cubos y tableros (dashboard).

En la tabla 5 se establece el orden en que fueron ejecutadas las diferentes tareas para la creación de dicho proyecto.

Objetivo	Actividad	Duración	Fecha Origen	Fecha Fin
Analizar la lógica de negocio que manejan las organizaciones del área de seguros para vehículos, con atención sobre el proceso de solicitud de citas para la inspección de vehículos.	Reuniones, Entrevistas y Consultas	3 semanas	12/03/2018	02/04/2018
Definir los indicadores y reportes relacionados con la solicitud, asignación e inspección de vehículos en base a la identificación de los requisitos del negocio.	Definición de los Requerimientos del Negocio	3 semanas	03/04/2018	24/04/2018
Desarrollar un almacén de datos adaptable a la Solución de Inteligencia de Negocio a partir de los requisitos del negocio previamente identificados.	Diseño Técnico de la Arquitectura	1 semanas	25/04/2018	02/05/2018
	Selección de Productos e Instalación	1 semanas	03/05/2018	10/05/2018
	Modelado Dimensional	3 semanas	11/05/2018	01/06/2018
	Diseño Físico	1 semanas	02/06/2018	09/06/2018
	Diseño y Desarrollo de procesos de extracción, transformación y carga de datos.	2 semanas	10/06/2018	24/07/2018
Implementar un tablero de control, un cubo de información y un conjunto de reportes que permitan la rápida visualización de la información para apoyar la toma de decisiones.	Especificaciones de las Aplicaciones Analíticas	2 Semanas	25/07/2018	08/08/2018
	Desarrollo de Aplicaciones Analíticas	2 Semanas	09/08/2018	23/08/2018
	Implementación	3 Semanas	24/08/2018	14/09/2018

TABLA 5 ACTIVIDADES ASOCIADAS A LA IMPLEMENTACIÓN DE LA SOLUCIÓN DE INTELIGENCIA DE NEGOCIOS

4.1.2. Definición de los Requerimientos del Negocio

Con la intención de ofrecer una herramienta que de soporte a la toma de decisiones sobre cualquier proceso relacionado con la solicitud de citas vía web para el peritaje de automóviles se ha desarrollado la siguiente solución de inteligencia de negocios para cualquier organización relacionada al área de seguros que pretenda ejecutar algún tipo de control o seguimiento sobre dicho proceso. Es por esto que, para cubrir esta fase se realizaron una

serie de reuniones o entrevistas con el personal representante y dueño del proceso de solicitud de citas para la inspección de vehículos a fin de entender cómo funciona este desde el momento en que un cliente realiza la solicitud, llena el formulario, recibe un numero de orden, se perita el automóvil y finalmente responde la encuesta para, así, conocer el déficit que tienen a la hora de visualizar los datos que permiten el análisis sobre todas las operaciones relacionadas al proceso. Como parte de los resultados obtenidos a partir de dichas entrevistas se tiene la necesidad por parte de la empresa de seguros conocer lo siguiente:

- Reflejar la cantidad de inspecciones que se realizan de acuerdo con la fecha de solicitud, tipo de inspección, perito que realiza la inspección del vehículo, estatus de la solicitud realizada y ubicación geográfica del centro de inspección. Realizada y ubicación geográfica del centro de inspección. La cantidad de solicitudes de inspecciones se puede calcular con la siguiente fórmula:

$$\sum Vi$$

Donde V es el total de solicitudes hechas en diariamente e i es la fecha de solicitud.

- Promediar la cantidad de solicitudes que se realizan según la fecha de solicitud, tipo de inspección, perito que realiza la inspección del vehículo, estatus de la solicitud realizada y ubicación geográfica del centro de inspección.

El promedio de solicitudes se puede calcular aplicando la siguiente fórmula:

$$\left(\frac{V2 + V1}{2}\right)$$

Donde V1 es el total de solicitudes hechas en la fecha inicial y V2 es el total de solicitudes hechas en la fecha final.

- Obtener el porcentaje de solicitudes de inspección que se realizan en función de la fecha de solicitud, tipo de inspección, perito que realiza la inspección del vehículo, estatus de la solicitud realizada y ubicación geográfica del centro de inspección.

Se puede calcular la cantidad porcentual con la fórmula:

$$\left(\frac{V2 + V1}{\text{total}}\right) \times 100$$

En la que V1 representa el valor pasado o inicial, V2 representa el valor presente o final y total representa la sumatoria de todas las solicitudes desde el valor V1 hasta el valor V2

- Calcular la variación de la cantidad de solicitudes de inspección realizadas de acuerdo con la fecha de solicitud, tipo de inspección, perito que realiza la inspección del vehículo, estatus de la solicitud realizada y ubicación geográfica del centro de inspección.

Se puede calcular el valor de la variación con la fórmula:

$$\bullet \left(\frac{V_2 - V_1}{V_1} \right)$$

En la que "V1" representa el valor pasado o inicial y "V2" representa el valor presente o final

- Expresar el porcentaje de la variación de solicitudes de inspección de vehículos relacionado a la fecha de solicitud, tipo de inspección, perito que realiza la inspección del vehículo, estatus de la solicitud realizada y ubicación geográfica del centro de inspección.

Se puede calcular la variación porcentual con la fórmula:

$$((V_2 - V_1)/V_1) \times 100$$

En la que "V1" representa el valor pasado o inicial y "V2" representa el valor presente o final

4.1.3. Diseño de la Arquitectura

El diseño de la arquitectura para esta solución de inteligencia de negocio puede observarse en la figura 27.

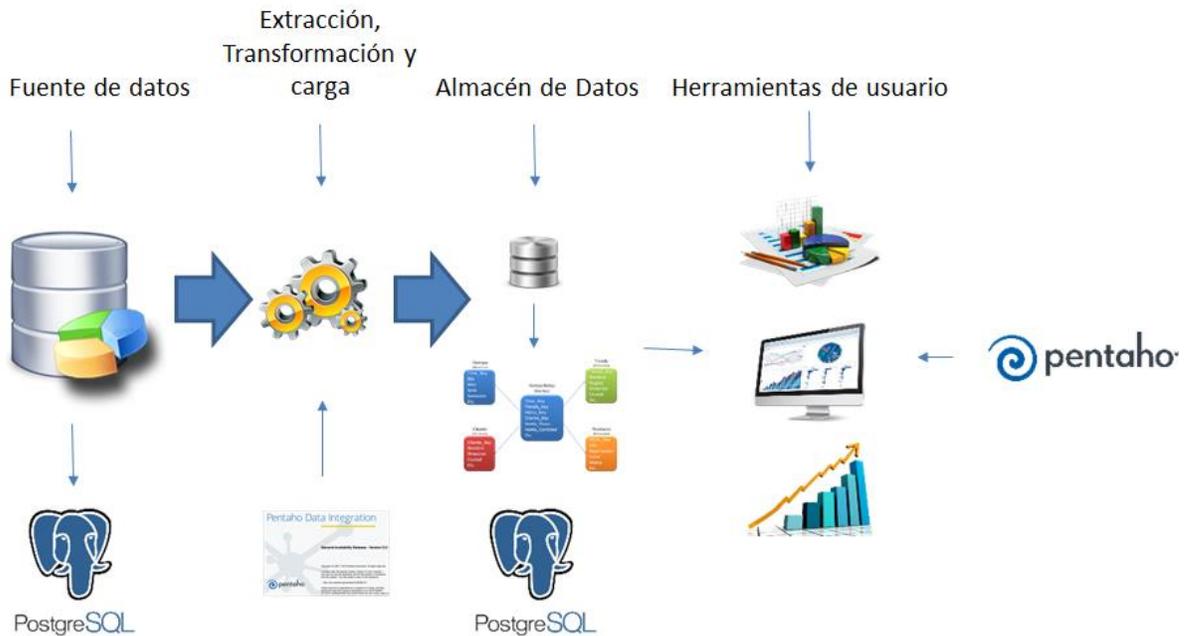


FIGURA 27 ARQUITECTURA DE SOLUCIÓN DE INTELIGENCIA DE NEGOCIOS

De acuerdo con la estructura reflejada en la imagen previa se tienen diferentes componentes, procesos y elementos que conforman dicha solución y que permiten llevar a conclusión dicho trabajo.

De izquierda a derecha, nos encontramos primero con la fuente de datos que es parte esencial de la solución ya que de ella se obtienen los datos que serán manipulados para la eventual emisión de los indicadores y reportes que se desean obtener. Forma parte esencial, de igual forma, del sistema transaccional que realiza la captura de los datos a través del proceso de negocio que se ha estudiado con anterioridad y adicionalmente se encuentra implementado en PostgreSQL como una base de datos.

Seguido de esto nos encontramos con los procesos de extracción, transformación y carga de los datos que deben ejecutarse en forma automática habitualmente ya que estos son los

encargados de cargar progresivamente nuestro almacén de datos a partir de la data proveniente de nuestra fuente a través de un conjunto de procesos y transformaciones. Para este segmento se hizo uso de la herramienta Pentaho Data Integration (PDI) 6.1 a fin de crear los diferentes procesos de extracción, transformación y carga (ETL).

El segmento final que podemos observar a la derecha de la figura 24 se compone de diferentes elementos que describiremos a continuación:

El primero de ellos es el almacén de datos que pretende recopilar toda la data histórica a partir del sistema fuente con la intención de automatizar el proceso de análisis a través del conjunto de instrumentos por el cual vienen acompañados los módulos de la herramienta Pentaho, todo esto en función de un modelo dimensional que facilita dicho análisis y que es progresiva y automáticamente llenado a partir de los procesos ETL previamente diseñados. Para el desarrollo de este almacén se hace uso de una base de datos en PostgreSQL acompañado de la herramienta pgadmin 4, versión 1.0 con la finalidad de crear todas las sentencias SQL, relaciones, claves y estructuras asociadas a la construcción de nuestro almacén de datos.

El siguiente está asociado a la forma en que es presentada la información, para esto, y como ya lo hemos mencionado, disponemos de un conjunto de herramientas que nos permiten el acceso a los datos dentro del almacén, ya diseñado, para obtener el conjunto de vistas, análisis y reportes asociados a los indicadores que se pretenden obtener como parte de la solución planteada originalmente. Para lograr esto se acude al conjunto de módulos de los cuales dispone la suite Pentaho Business Intelligence en su versión Community, dentro del cual se hizo uso de la herramienta Pentaho Report Designer para la elaboración y emisión de los diferentes reportes, Pentaho Schema Workbench para el desarrollo del esquema de datos y cubos de análisis y Pentaho BI Server para el desarrollo del panel de control.

Es importante recalcar en este punto que originalmente se había diseñado dentro de la arquitectura un área intermedia (Staging Area) que almacenaría temporalmente la data y donde se recogerían los datos que se necesitan de los sistemas de origen. Adicionalmente, se recopilarían los datos estrictamente necesarios para las cargas, y se aplicaría el mínimo de transformaciones a los mismos. Sin embargo, se decidió desistir de dicha idea ya que la fuente es única y no requiere de un paso adicional para ser cargado.

4.1.4. Selección de Productos e Instalación

La selección de software BI se basa en la comparación de los aspectos a considerar sobre cada una de las herramientas descritas previamente en el Marco Conceptual (Capítulo II). Las herramientas investigadas cuentan con soluciones para responder a las necesidades sobre el proceso de solicitud de citas para la inspección de vehículos, tales como la presentación de informes, análisis OLAP y diseño de panel de control (Dashboard).

Este proceso puede tornarse complejo ya que actualmente en el mercado la diversidad de herramientas disponibles es muy variada y cada una posee una vasta cantidad de funcionalidades que acompañan a cada una de estas. Pero, más allá de todo esto es importante que la herramienta que seleccionemos disponga de los artefactos necesarios para lograr los objetivos planteados originalmente y que nos permita recrear el repositorio centralizado de datos, diseñar y recrear el panel de control asociado a este y finalmente poder construir las consultas que atiendan a las demandas de información solicitadas a petición de la parte interesada en estos.

Como herramienta de inteligencia de negocios (BI), de todas las expuestas anteriormente en el capítulo II, Pentaho es una suite basada en elementos estándares que reúne todo los aspectos básicos de una herramienta de inteligencia de negocios (BI), asimismo se destaca por poseer una capacidad mucho mayor de personalización, pudiendo adaptarse en mejor detalle a las solicitudes que exijan los directivos de la organización en los formatos de los informes, paneles de control (Dashboard), entre otros.

Por esta razón y luego de haber realizado el correspondiente análisis sobre las diferentes herramientas disponibles en el mercado, se optó utilizar para el desarrollo de dicha solución de inteligencia de negocios las herramientas de la Suite Pentaho Business Intelligence en su versión Community.

Adicionalmente y teniendo como marco de desarrollo el uso de la Suite Pentaho Business Intelligence en su versión Community se selecciona como SMD a PostgreSQL, ya que el mismo es multiplataforma, está diseñado para ambientes de alto volumen tanto de transacciones como de usuarios y además posee una alta escalabilidad, lo que ha hecho que sea la alternativa de Software Libre idónea para competir y sustituir BD comerciales como Oracle. Además de esto, y como punto final, la base de datos del sistema transaccional del cual se extraen los datos se encuentra implementada en PostgreSQL lo cual facilita la integración y

unificación en el uso de las herramientas. En la tabla 6 se recrea el orden de uso y la funcionalidad de cada módulo usado.

Herramienta	Us0	Requisitos de Software
PostgreSQL	Creación y administración de bases de datos	Sistema Operativo: Windows o Linux Memoria RAM: min 1 GB, recomendable > 2 GB Memoria en disco duro: > 40 MB
Pentaho Data Integration	Creación y Administración de Procesos ETL	Sistema Operativo: Windows o Linux Memoria RAM: min 1 GB, recomendable > 2 GB Memoria en disco duro: 200 MB
Pentaho Schema Workbench	Creación y Administración de los esquemas de datos que serán alojados en Pentaho BI Server, estos esquemas de datos son la base para las consultas que se usan en los tableros de control y en los reportes.	Sistema Operativo: Windows o Linux Memoria RAM: min 1 GB, recomendable > 2 GB Memoria en disco duro: 200 MB
Pentaho BI Server	Creación y Administración de tableros de control, así como servidor para alojar los reportes y para la administración de los usuarios y los esquemas de datos.	Sistema Operativo: Windows o Linux Memoria RAM: min 1 GB, recomendable > 2 GB Memoria en disco duro: 500 MB
Pentaho Report Designer	Creación y Administración de Reportes, se enlaza con Pentaho BI Server para	Sistema Operativo: Windows o Linux

	hacer uso de los esquemas de datos allí alojados.	Memoria RAM: min 2 GB, recomendable > 4 GB Memoria en disco duro: 500 MB
--	---	---

TABLA 6 HERRAMIENTAS PARA EL DESARROLLO DE LA SOLUCIÓN DE INTELIGENCIA DE NEGOCIOS

4.1.5. Modelado Dimensional

En función de los requerimientos del negocio que fueron expuestos en el punto 4.1.2 se validaron las relaciones entre los criterios y la lógica del negocio de acuerdo con las estructuras de datos del sistema transaccional, para determinar qué elementos podrían considerarse como hechos medibles o dimensiones que pudieran formar parte de la solución de inteligencia de negocio, y con ello, dar inicio al diseño del modelo dimensional del almacén de datos, el cual comprende el proceso de Control y Seguimiento al proceso de solicitud de citas para la inspección de vehículos de la empresa de seguros. Ahora bien, con la finalidad de lograr el desenvolvimiento de la fase que se presenta a continuación relacionado al modelo dimensional es necesario dar seguimiento a las actividades propuestas por dicha metodología en este paso tal como se muestran en los siguientes puntos:

4.1.5.1. Definir el Proceso de Negocio

Se pretende recrear un modelo de control y seguimiento al proceso de solicitud de citas para la inspección de vehículos de una empresa de seguros en el cual se podrán observar indicadores asociados.

4.1.5.2. Identificar el Nivel de Granularidad

Con la intención de calcular el nivel de granularidad que se tiene es preciso determinar la representación de una fila en función de la tabla de hechos (sin hechos), ya que, de esta forma podemos especificar el grado asociado al detalle de la información que pretendemos presentar, esto permitirá determinar los valores de aquello que se desea medir de acuerdo con los indicadores planteados previamente. Habiendo llegado a esto podemos establecer la granularidad de la solución.

del porcentaje de la cantidad de solicitudes de inspección de vehículos pertenecientes a un municipio.

- Cantidad de solicitudes de inspección de vehículos, promedio de la cantidad de solicitudes de inspección de vehículos, porcentaje de solicitudes de inspección de vehículos, variación de la cantidad de solicitudes de inspección de vehículos, variación del porcentaje de la cantidad de solicitudes de inspección de vehículos pertenecientes a un centro de inspección de vehículos.

4.1.5.4. Definir las Dimensiones

Alineados con los requerimientos recopilados y en función de la información suministrada por los involucrados o dueños del proceso acompañado del debido análisis del sistema operacional podemos identificar en la tabla las dimensiones que formaran parte de dicha solución.

Dimensión	Nombre en el Almacén	Descripción
Tiempo	tiempo	Tabla usada para guardar datos relacionados al tiempo (año, mes, día).
Centro de Inspección	centro_de_inspeccion	Tabla que almacena datos sobre la ubicación geográfica como el estado y municipio donde se encuentra un determinado centro de inspección.
Estatus	estatus	Posee información relacionada al estado en el cual se encuentra una determinada solicitud de inspección de vehículo.
Perito	perito	Contiene información asociada a los peritos que ejecutan las inspecciones sobre los vehículos como el nombre, sexo o cédula.
Solicitud de Inspección	solicitud_de_inspeccion	Tabla en la que se almacenan datos asociados a la solicitud realizada por los clientes.
Tipo de Inspección	tipo_de_inspeccion	En esta tabla se encuentran los posibles tipos de inspección que se pueden realizar en cualquiera de los centros de inspección de la empresa de seguros bien sea para solicitar una afiliación al servicio o para reportar un siniestro ocurrido al vehículo previamente asegurado.

TABLA 7 DIMENSIONES ASOCIADAS AL MODELO DIMENSIONAL EN LA SOLUCIÓN PLANTEADA

Acompañado de esto se logró identificar un conjunto de relaciones jerárquicas a partir de las propiedades de las dimensiones previamente mencionadas y ajustadas a la lógica de negocio,

con lo cual se lograron obtener como resultado de este análisis jerarquías para la dimensión tiempo y la dimensión centro de inspección las cuales podremos considerar en la siguiente figura.

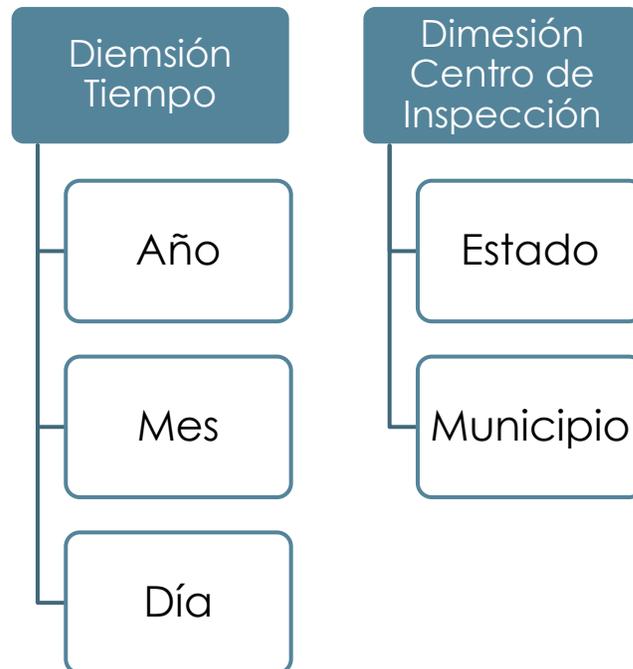


FIGURA 28 RELACIONES JERÁRQUICAS DEL MODELO DIMENSIONAL EN LA SOLUCIÓN PLANTEADA

4.1.5.5. Identificación de los Hechos y las Tablas de Hechos

Modelo dimensional de la solución propuesta

En la siguiente ilustración (Figura 26) se puede observar el diagrama de la estructura asociada al modelo dimensional de la solución.

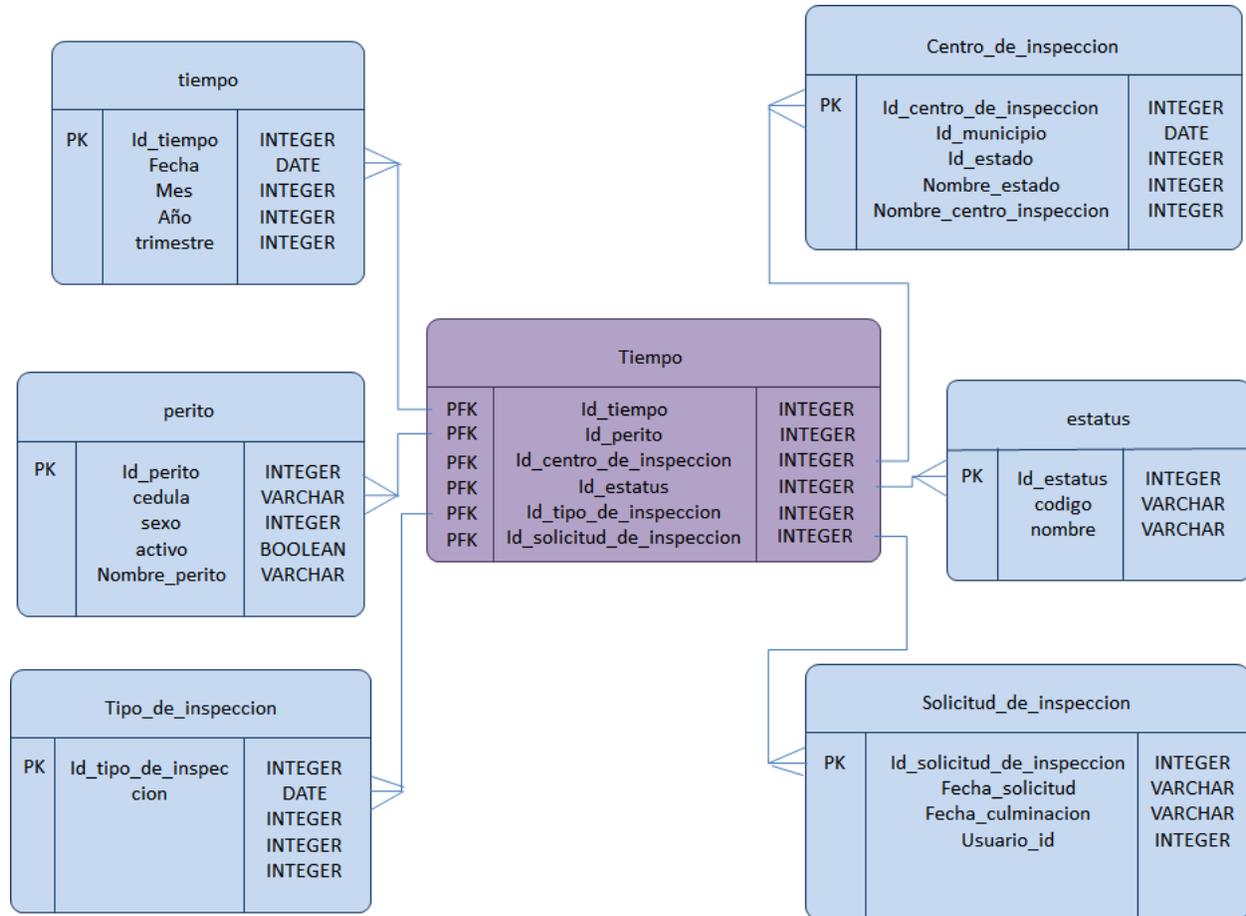


FIGURA 29 MODELO DIMENSIONAL PROPUESTO

4.1.6. Diseño Físico

La siguiente fase se centra en la selección de las estructuras necesarias para dar soporte al diseño lógico de los datos. Un elemento principal de este proceso es la definición de estándares del entorno de la base de datos. Otro elemento importante es la selección de la indexación, para lo cual se seleccionaron las claves primarias como índices.

Ya que hemos logrado definir el modelo dimensional, utilizando PostgreSQL y la herramienta pgadmin IV provista por dicha base de datos, nos concentraremos en el desarrollo y

construcción de las diferentes estructuras físicas asociadas, bien sean las dimensiones y la tabla de hechos del almacén de datos dentro del sistema manejador de base de datos por medio de sentencias SQL.

Las siguientes son las propiedades del sistema manejador de base de datos:

Servidor: localhost

Puerto: 5433

Nombre de la base de datos: dwh

Esquema: public

Usuario: PostgreSQL

A continuación, en la presente fase se muestra el diseño físico para las dimensiones y las tablas de hechos que fueron diseñadas en el modelado dimensional.

Dimensión Tiempo

Columns +						
	Name	Data type	Length	Precision	Not NULL?	Primary key?
<input checked="" type="checkbox"/>	id_tiempo	integer			<input checked="" type="checkbox"/> Yes	<input checked="" type="checkbox"/> Yes
<input checked="" type="checkbox"/>	fecha	timestamp without time zone			<input type="checkbox"/> No	<input type="checkbox"/> No
<input checked="" type="checkbox"/>	mes	double precision			<input type="checkbox"/> No	<input type="checkbox"/> No
<input checked="" type="checkbox"/>	año	double precision			<input type="checkbox"/> No	<input type="checkbox"/> No
<input checked="" type="checkbox"/>	trimestre	double precision			<input type="checkbox"/> No	<input type="checkbox"/> No

FIGURA 30 DISEÑO FÍSICO DE LA DIMENSIÓN TIEMPO

Dimensión Centro de Inspección

Columns +						
	Name	Data type	Length	Precision	Not NULL?	Primary key?
<input checked="" type="checkbox"/>	id_centro_de_inspeccion	integer			<input checked="" type="checkbox"/> Yes	<input checked="" type="checkbox"/> Yes
<input checked="" type="checkbox"/>	id_municipio	integer			<input type="checkbox"/> No	<input type="checkbox"/> No
<input checked="" type="checkbox"/>	nombre_municipio	character varying	255		<input type="checkbox"/> No	<input type="checkbox"/> No
<input checked="" type="checkbox"/>	id_estado	integer			<input type="checkbox"/> No	<input type="checkbox"/> No
<input checked="" type="checkbox"/>	nombre_estado	character varying	255		<input type="checkbox"/> No	<input type="checkbox"/> No
<input checked="" type="checkbox"/>	nombre_centro_inspeccion	character varying	255		<input type="checkbox"/> No	<input type="checkbox"/> No

FIGURA 31 DISEÑO FÍSICO DE LA DIMENSIÓN CENTRO DE INSPECCIÓN

Dimensión Estatus

Columns +						
	Name	Data type	Length	Precision	Not NULL?	Primary key?
<input checked="" type="checkbox"/>	id_estatus	integer			<input checked="" type="checkbox"/> Yes	<input checked="" type="checkbox"/> Yes
<input checked="" type="checkbox"/>	codigo	character varying	100		<input type="checkbox"/> No	<input type="checkbox"/> No
<input checked="" type="checkbox"/>	nombre	character varying	255		<input type="checkbox"/> No	<input type="checkbox"/> No

FIGURA 32 DISEÑO FÍSICO DE LA DIMENSIÓN ESTATUS

Dimensión Perito

Columns +						
	Name	Data type	Length	Precision	Not NULL?	Primary key?
<input checked="" type="checkbox"/>	id_perito	integer			<input checked="" type="checkbox"/> Yes	<input checked="" type="checkbox"/> Yes
<input checked="" type="checkbox"/>	cedula	character varying	100		<input type="checkbox"/> No	<input type="checkbox"/> No
<input checked="" type="checkbox"/>	sexo	integer			<input type="checkbox"/> No	<input type="checkbox"/> No
<input checked="" type="checkbox"/>	activo	boolean			<input type="checkbox"/> No	<input type="checkbox"/> No
<input checked="" type="checkbox"/>	nombre_perito	character varying	255		<input type="checkbox"/> No	<input type="checkbox"/> No

FIGURA 33 DISEÑO FÍSICO DE LA DIMENSIÓN PERITO

Dimensión Solicitud de Inspección

Columns +						
	Name	Data type	Length	Precision	Not NULL?	Primary key?
<input checked="" type="checkbox"/>	id_solicitud_de_inspeccion	integer			<input checked="" type="checkbox"/> Yes	<input checked="" type="checkbox"/> Yes
<input checked="" type="checkbox"/>	fecha_solicitud	character varying	100		<input type="checkbox"/> No	<input type="checkbox"/> No
<input checked="" type="checkbox"/>	fecha_culminacion	character varying	100		<input type="checkbox"/> No	<input type="checkbox"/> No
<input checked="" type="checkbox"/>	usuario_id	integer			<input type="checkbox"/> No	<input type="checkbox"/> No

FIGURA 34 DISEÑO FÍSICO DE LA DIMENSIÓN SOLICITUD DE INSPECCIÓN

Dimensión Tipo de Inspección

Columns +						
	Name	Data type	Length	Precision	Not NULL?	Primary key?
<input checked="" type="checkbox"/>	id_tipo_de_inspeccion	integer			<input checked="" type="checkbox"/> Yes	<input checked="" type="checkbox"/> Yes
<input checked="" type="checkbox"/>	codigo	character varying	50		<input type="checkbox"/> No	<input type="checkbox"/> No
<input checked="" type="checkbox"/>	inspeccion	text			<input type="checkbox"/> No	<input type="checkbox"/> No
<input checked="" type="checkbox"/>	nombre	character varying	255		<input type="checkbox"/> No	<input type="checkbox"/> No

FIGURA 35 DISEÑO FÍSICO DE LA DIMENSIÓN TIPO DE INSPECCIÓN

En la figura 36 se puede observar el modelo físico del almacén de datos que ha sido implementado en el sistema manejador de base de datos.

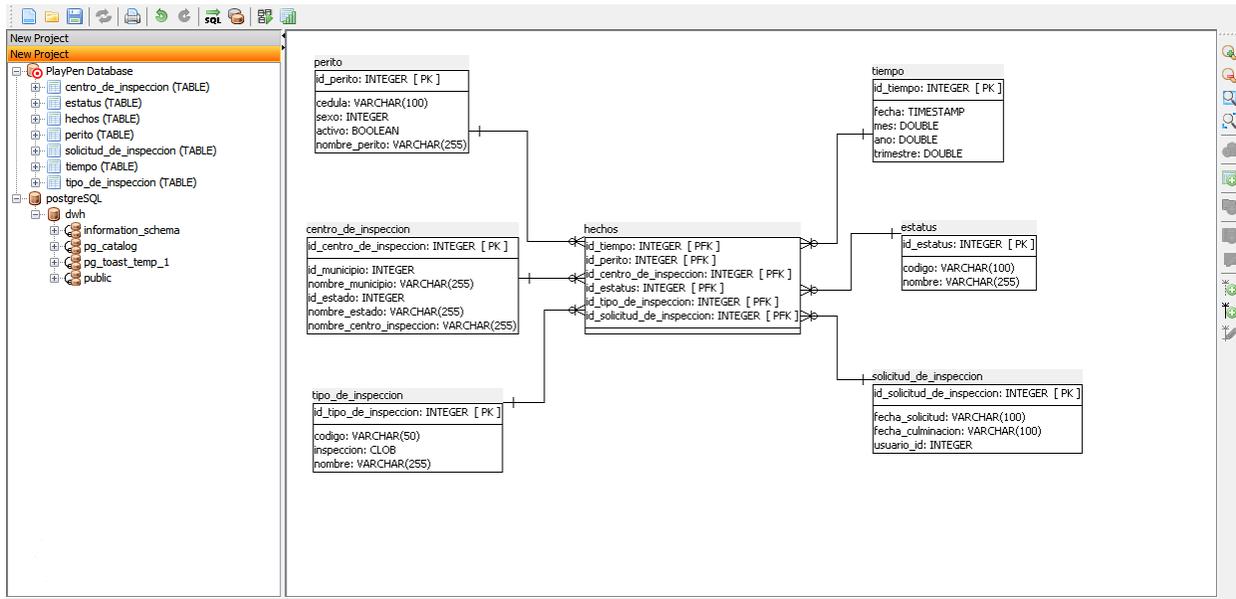


FIGURA 36 MODELO FÍSICO DEL ALMACÉN DE DATOS PROPUESTO COMO SOLUCIÓN

4.1.7. Diseño y Desarrollo de Procesos ETL

Como parte del proceso de extracción, transformación y carga (ETL por sus siglas en inglés) se realizaron una serie de Transformaciones (Transformations) y Trabajos (Jobs) a través de la herramienta Spoon de Pentaho. Esta herramienta permite almacenar los archivos de dichas transformaciones y trabajos en el sistema de archivos local o en un catálogo de PDI.

De acuerdo con esta fase se muestran las principales transformaciones (Transformations) y trabajos (Jobs) realizados para la implementación del sistema automatizado de inteligencia de negocios en el proceso de solicitud de citas para la inspección de vehículos, sin embargo, antes de abordar esto en la tabla 12 se describe la funcionalidad de los principales pasos (step) usados para realizar dichas transformaciones y trabajos.

Imagen Asociada	Descripción
 Table input	Table Input: Este paso se utiliza para leer la información desde una base de datos, utilizando una conexión y SQL
 Merge Join	Merge Join: Permite producir combinaciones de todas las filas en los flujos de entrada. Las entradas deben estar ordenadas previamente.
 Select values	Select Values: Es útil para la selección, el cambio de nombre, cambio de tipos de datos, así como la configuración de la longitud y la precisión de los campos que vienen del flujo de datos..
 Calculator	Calculator: Crea nuevos campos y ejecuta cálculos simples
 Sort rows	Sort Rows: Ordena las filas basados en el valor de algún campo (ascendente o descendente).
 Table output	Table Output: Escribe la información a la tabla de una base de datos.
 Combination lookup/update	Lookup/Update: Permite implementar las dimensiones de lento cambio para ambos tipos: tipo I y tipo II. Además no sólo se puede utilizar este paso para actualizar una tabla de dimensiones, también se puede utilizar para buscar valores en las dimensiones.

TABLA 8 PRINCIPALES PASOS UTILIZADOS EN LOS PROCESOS DE ETL

Ahora bien, una vez creado el diseño físico de nuestro almacén de datos con la herramienta PostgreSQL debemos navegar en el proceso de diseño y eventual implementación de los diferentes procesos de extracción, transformación y carga (ETL) con la finalidad de tomar los datos provenientes de la fuente, ejecutar una serie de transformaciones a estos que nos permitan obtener de ellos la información requerida y finalmente cargarlos en las dimensiones de nuestro almacén, todo esto con el objetivo de facilitar a través de este la obtención de la información ajustada y oportuna que brinde el apoyo necesario a la toma de decisiones de carácter estratégico en función del control y seguimiento del proceso de solicitud de citas para la inspección de vehículos. En este punto, es importante señalar que los datos utilizados como fuente de datos fueron recreados a partir de la implementación del modelo de datos

suministrados por la empresa de seguros y los datos que llenaron dicho modelo fueron recreados en forma aleatoria con la herramienta Data Generator que genera de forma automática los registros, por lo tanto, no es información fidedigna o fiel a la que maneja la organización.

Como parte de la respuesta para el desarrollo de estos procesos se hace uso de la herramienta Pentaho Data Integration (PDI) ya que gracias a esta se pueden recrear, como ya lo hemos mencionado, el conjunto de transformaciones que pueden enlazar la fuente de los datos y extraer de estas todo lo necesario para que, después de ser debidamente procesados, sean cargados en el eventual repositorio destino sobre el cual se levantarán los diferentes reportes de análisis que darán respuesta a los requerimientos solicitados inicialmente. Acompañado de esto también se crean los trabajos (Jobs) que dan pie a la ejecución de un conjunto de transformaciones en forma automatizada evitando la ejecución puntual de cada una de estas.

De igual forma tanto las transformaciones como los trabajos que se diseñaron se almacenan en el sistema de archivos de la aplicación, el mismo, que posee un directorio univoco el cual facilita el acceso y la ubicación dentro del sistema de archivos propio del sistema operativo donde ha sido instalada la aplicación, todo esto puede ser observado en detalle en la figura 37.

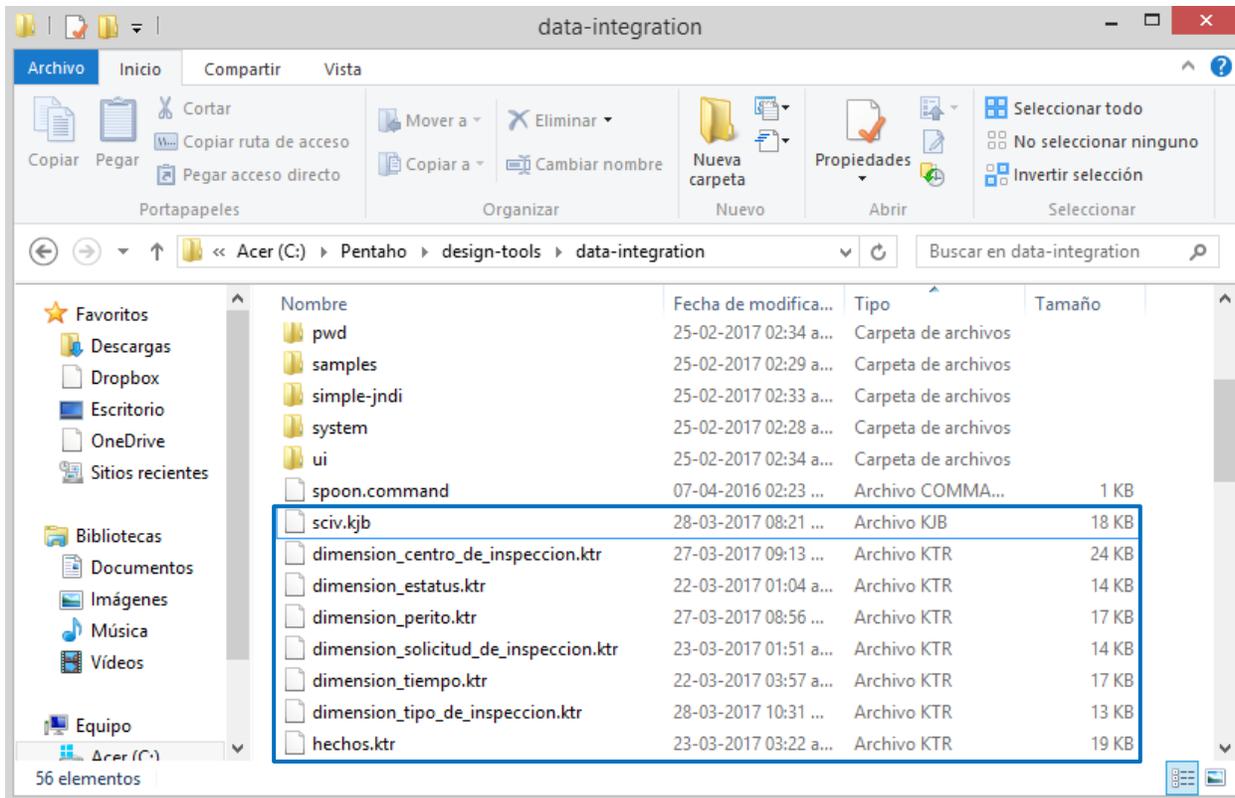


FIGURA 37 DIRECTORIO DE ARCHIVOS

A continuación, se muestran algunos de los principales procesos ETL diseñados y desarrollados para la carga de datos en las tablas que conforman el almacén de datos:



FIGURA 38 ETL PARA CARGAR LA DIMENSIÓN CENTRO DE INSPECCIÓN

En la figura 38, podemos observar que, con la finalidad de cargar todos los datos de una dimensión dentro de nuestro almacén de datos se realizan consultas provenientes de diferentes tablas fuentes entre la cuales tenemos sgt_centroinspeccion, sgt_municipio y sgt_estado, todas provenientes de la base de datos del sistema transaccional que posee la

organización. Una vez realizada la carga de los datos provenientes de las diferentes fuentes se procede a realizar la depuración de esta a través de un conjunto de transformaciones para finalmente obtener, además de un conjunto de datos estandarizados de los datos, solo aquellos campos necesarios para dar forma al registro que será introducido finalmente en la dimensión correspondiente del almacén, en este caso llamada dimensión centro de inspección.

Al igual que el proceso previo, en la figura 39 se puede apreciar la extracción de los datos provenientes de la fuente gracias a la tabla sgt_estatus de la base de datos transaccional en donde se encuentran previamente cargados todos los campos relacionados a los posibles estatus en los que se puede encontrar o que puede asumir una determinada solicitud realizada por algún cliente, es importante recalcar que los estatus no son seleccionados por el cliente que hace la solicitud ya que estos son asimilados por la petición hecha por el cliente en forma automática, por lo tanto solo queda ordenarlos de acuerdo a la línea de ejecución que pueden asimilar dichas solicitudes y finalmente ser cargados en la dimensión correspondiente dentro de nuestro almacén.



FIGURA 39 ETL PARA CARGAR LA DIMENSIÓN ESTATUS

En la siguiente imagen (figura 40) se realiza la carga de registros a partir de la tabla sgt_perito proveniente de la base de datos transaccional que eventualmente serán ordenados por la clave primaria para luego realizar un cálculo previo antes de seleccionar los campos que serán finalmente cargados en la dimensión asociada.

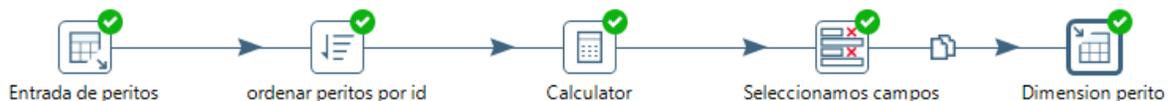


FIGURA 40 ETL PARA CARGAR LA DIMENSIÓN PERITO

En la figura 41 podemos observar de igual forma la carga de los campos relacionados a la dimensión solicitud de inspección a partir de la fuente de datos en la tabla sgt_solicitudinspeccion, cargando los datos correspondientes al repositorio final dentro del almacén de datos.



FIGURA 41 ETL PARA CARGAR LA DIMENSIÓN SOLICITUD DE INSPECCIÓN

Esta transformación crea la dimensión tiempo que, a diferencia de la suministrada como plantilla por Pentaho, toma de la fuente la fecha de solicitud de la inspección, la cual es posteriormente ordenada por fecha y después de se realizan el conjunto de cálculos que permitirán poblar dicha dimensión con los datos necesarios. En la figura 42 podemos observar el conjunto de pasos que describen dicho proceso ETL para llegar eventualmente a la creación de la dimensión en cuestión.

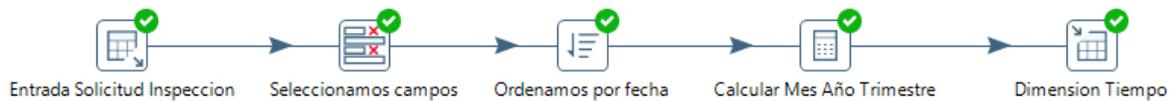


FIGURA 42 ETL PARA CARGAR LA DIMENSIÓN TIEMPO

La dimensión tipo de inspección cuenta con una particularidad única dentro del almacén ya que los datos provenientes de la fuente son cargados por completo debido a que solo existen dos posibles tipos de inspección asociados a cualquier solicitud de inspección de vehículos de los clientes, por lo tanto, en la figura 43 podemos apreciar detalladamente que solo cuenta con 2 pasos que involucran dicho proceso de construcción para esta dimensión.



FIGURA 43 ETL PARA CARGAR LA DIMENSIÓN TIPO DE INSPECCIÓN

Sumado a lo mencionado en todos los puntos anteriores, y como lo muestra la figura 44, se ha recreado una transformación para poblar la tabla de hechos, si bien esta transformación mantiene una estructura coherente con las anteriores, es un poco más compleja en el sentido

en que se realizan consultas al sistema transaccional que den coherencia a las dimensiones ya previamente cargadas, y así poder obtener todos los registros en base a sus claves primarias.

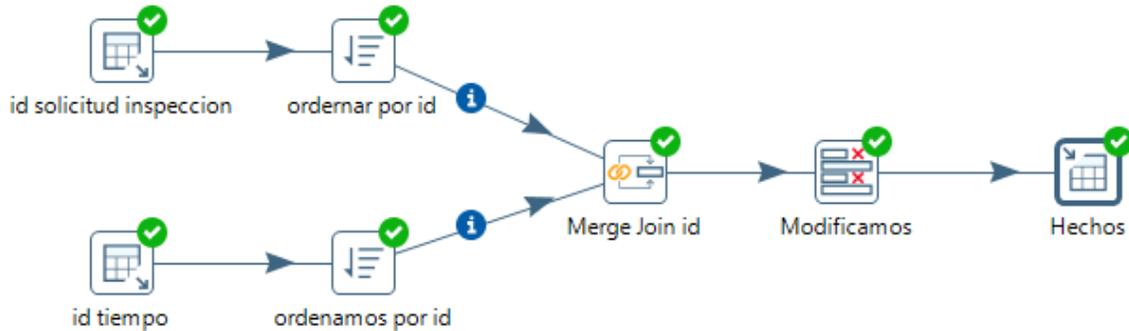


FIGURA 44 ETL PARA CARGAR LA TABLA DE HECHOS

En relación con todo lo expresado anteriormente y con la intención de no ejecutar individualmente cada transformación, se crean uno o más trabajos (Jobs) para poblar el almacén. Un trabajo (Job) es un componente de Pentaho Data Integration que permite crear una secuencia de actividades brindando un orden de ejecución, es decir que este trabajo (Job) se crea para llevar un control de flujo de las transformaciones que se deben ejecutar en un momento dado. Cada trabajo (Job) posee una estructura parecida a la reflejada en la figura 45.



FIGURA 45 TRABAJO (JOB) PARA AUTOMATIZAR TODOS LOS ETL'S

4.1.8. Especificaciones de las Aplicaciones Analíticas

Con la intención de construir exitosamente los requerimientos que solicitan los clientes es importante e indispensable que la forma en que estos son recreados lleguen a manos de quienes hacen dicha solicitud pero que además faciliten la forma en que se obtiene la

información sobre estos, que el manejo sea fácil e intuitivo a la vista del cliente, que haga cálculos y emita reportes en forma eficaz e igualmente eficiente y que además se adapten a los posibles cambios que se produzcan dentro de la organización. Es por esta razón que se ejemplifica en la figura # una muestra de los reportes y conjuntamente la estructura del panel de control a ser usado.

Como previamente lo hemos mencionado, se busca un diseño intuitivo a los ojos de quien haga uso del sistema de inteligencia de negocios propuesto, por tal motivo se concluyó recrear dicha implementación sobre un portal web ya que permite una interacción fluida y de carácter familiar para cualquier usuario que haya hecho uso de un computador, facilita el manejo del tablero donde serán emitidos los indicadores, el cubo que maneja los datos asociados a estos y reportes relacionados. Ahora bien, con la intención de que dicha aplicación cumpla con estándares básicos de implementación como seguridad, accesibilidad, maleabilidad, intuición, fácil manejo y comprensión es necesario definir algunas características con las cuales debe cumplir y que serán mencionadas a continuación.

- **Autenticación:** Con la intención de ofrecer un nivel de seguridad se pretende que cualquiera que desea hacer uso de dicho sistema de inteligencia de negocio acceda a través de un proceso de autenticación y autorización a los diferentes paneles de control que se ofrecen, además de los reportes y el manejo de la información en general gracias a la asignación de un usuario y contraseña previamente establecidos.
- **Diseño Simple:** Para lograr que el sitio web sea fácil de manejar, el conjunto de elementos dentro de este, que incluyen al panel y los reportes que se emitirán, serán presentados sin sobrecargas de información y agrupando los indicadores por cada uno de los filtros solicitados por parte de la empresa de seguros.
- **Vistas de los Reportes e Indicadores:** Debido a que nos encontramos en la búsqueda de una forma de constituir los indicadores es razonable pensar en que las vistas serán diseñadas y finalmente expresadas en función de elementos representativos de estos como lo son los gráficos de barra y gráficos de torta, primordialmente, no descartando otras formas que faciliten de igual manera el entendimiento y manejo de la información al igual que el reflejo de estos en los reportes que se podrán emitir.
- **Refinamiento de Consultas:** Los reportes que serán presentados y que se encuentran vinculados al panel de control cuentan con un conjunto de filtros que facilitan y refinan

los resultados de la información a quienes pretenden obtener deducciones a partir de estos y de esta forma obtener los respectivos análisis gracias a la emisión de los reportes al igual que los indicadores que se han desarrollado.

- **Formato Estándar:** Como previamente lo hemos mencionado, dicho sistema de inteligencia de negocios tendrá un formato estandarizado que le de cierta uniformidad en la forma de presentarse tanto los indicadores como los reportes a emitir, por lo tanto, estos contendrán encabezado, título y subtítulo.
- **Portabilidad de Información:** Bajo el ambiente de esta aplicación de inteligencia de negocios los usuarios podrán obtener los distintos reportes con la información que requieran en una amplia gama de formatos asociados, por supuesto, a todos estos.

4.1.9. Desarrollo de las Aplicaciones Analíticas

Gracias a la herramienta Schema Workbench provista por la suite de Pentaho Community se logró el desarrollo de la parte analítica del sistema a desplegar como elemento que formara parte de la solución de inteligencia de negocios originalmente propuesta, en la figura 46 se puede apreciar la generación del esquema de datos y toda la estructura asociada a este al igual que la del almacén de datos implementado como parte de la solución y cimiento de las consultas que serán usadas en la construcción de las vistas del panel de control y los reportes relacionados.

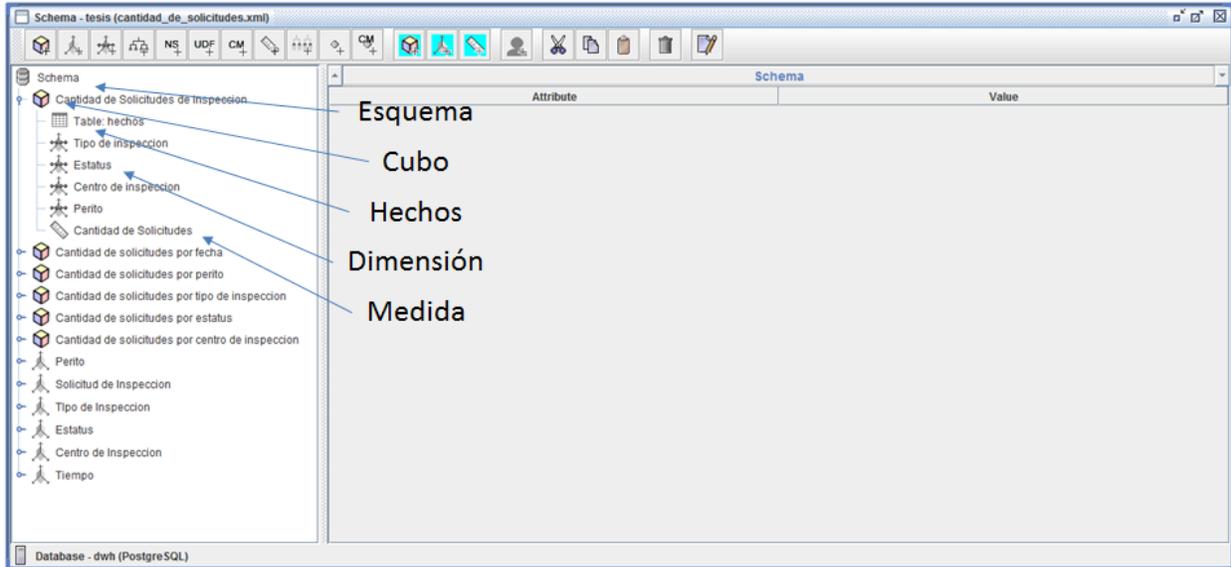


FIGURA 46 ESQUEMA DE DATOS DESARROLLADO CON PENTAHO SCHEMA WORKBENCH

4.1.10. Implementación

4.1.11. Pruebas

La implementación constituye el acoplamiento de distintos elementos como la tecnología, los datos y las aplicaciones orientadas a los usuarios finales. Una vez a bordo de la fase de implementación del Sistema de Inteligencia de Negocios para la empresa de seguros y realizado previamente todo el proceso de ETL, se procedió a la creación de los distintos cubos OLAP.

Los cubos OLAP se diseñaron con la herramienta Workbench de Pentaho y representa el conjunto de: dimensiones, tablas de hechos y las métricas que fueron descritas en la fase de Modelado Dimensional y Diseño Físico. Dicho conjunto permite mostrar los resultados de las consultas filtrando por los campos de la tabla, de manera que se puedan quitar y poner distintos criterios de búsqueda de los datos, consiguiendo un amplio abanico de posibilidades.

Para visualizar los cubos por medio de vistas dinámicas de los mismos, el usuario encargado debe levantar el servidor web de Pentaho BI-Sever. Una vez hecho esto, se podrá acceder al

Sistema de Inteligencia de Negocios propuesto en el presente T.E.G, a través del ingreso a la dirección correspondiente en un explorador web (browser).

Habiendo accedido exitosamente al sistema, el usuario puede acceder a la carpeta de Reportes donde se encuentran las vistas dinámicas de los cubos OLAP.



FIGURA 47 VISTA DINÁMICA DE LA CANTIDAD DE SOLICITUDES POR TIPO



FIGURA 48 VISTA DINÁMICA DE LA CANTIDAD DE SOLICITUDES POR ESTATUS

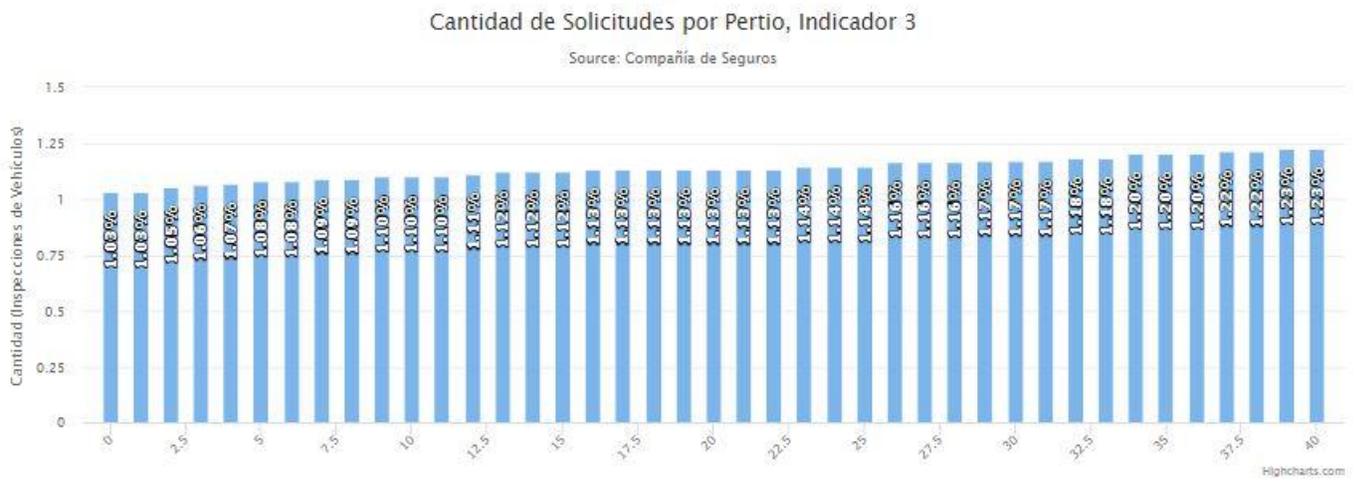


FIGURA 49 VISTA DINÁMICA DE LA CANTIDAD DE SOLICITUDES POR PERITO

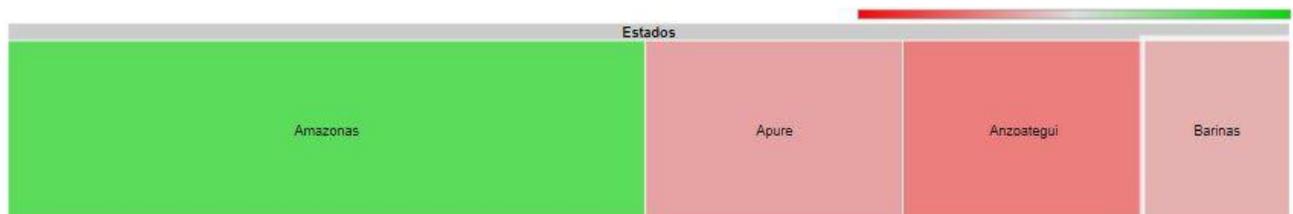


FIGURA 50 VISTA DINÁMICA DE LA CANTIDAD DE SOLICITUDES POR ESTADO Y MUNICIPIO



FIGURA 51 VISTA DINÁMICA DE LA CANTIDAD DE SOLICITUDES POR ESTADO Y CENTRO DE INSPECCIÓN

4.1.12. Pruebas

Para cumplir con la correcta implementación se requiere realizar comprobaciones de los datos que han sido procesados y las funcionalidades de las cuales dispone este a fin de contemplar el funcionamiento esperado del mismo, por esta razón se realizaron comprobaciones entre las tablas de la base de datos y la aplicación. Aunado a esto es de gran importancia conocer el nivel de aceptación para los usuarios y por tanto indispensable verificar el nivel de aprobación entre quienes harán uso de este.

4.1.12.1. Pruebas de Funcionalidad

Es importante en un desarrollo comprobar que las funciones implementadas en el sistema arrojen los resultados esperados, es por ello que se realizaron una serie de pruebas para distintas áreas de la solución.

En las figuras 52,53 y 54 se muestran que tanto la consulta realizada en la herramienta SQL como el resultado obtenido dentro del tablero coinciden en cuanto al conteo que se realizó y por lo tanto que se pretende traer antes quienes harán uso de este.



FIGURA 52 CONSULTA SQL

tipo_de_inspeccion	count
Suscripción	3905
Reclamo	2989

FIGURA 53 RESULTADO DE LA CONSULTA SQL

Fecha inicio: Fecha fin:



FIGURA 54 INDICADOR EN EL TABLERO POR TIPO DE SOLICITUD

Al igual que en el anterior caso se realizó una nueva consulta pero en esta oportunidad se ejecuta sobre los estatus de los cuales dispone el almacén de datos y se compara nuevamente con el respectivo indicador que se pretende mostrar en el dashboard.

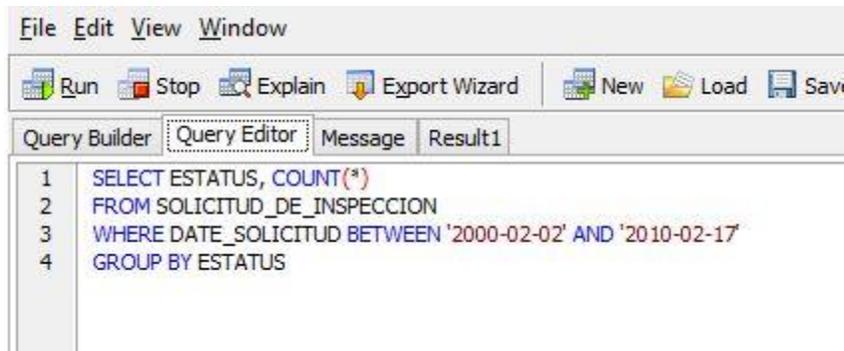


FIGURA 55 CONSULTA SQL ESTATUS

estatus	count
Rechazado	755
Aprobado	1702
Creado	1344
En Proceso	1095
Finalizado	1998

FIGURA 56 RESULTADO DE LA CONSULTA SQL



FIGURA 57 INDICADOR EN EL TABLERO POR ESTATUS DE SOLICITUD

4.1.12.2. Pruebas de Aceptación.

Con la finalidad de medir el nivel de aprobación de esta solución se llevó a cabo el método SUS (System Usability Scale o Escala de Usabilidad de Sistemas) creado por John Brooke en

1986 donde se definen diez afirmaciones, cinco son positivas y cinco negativas como se muestra en la tabla que a continuación se presenta (Tabla 9). Posterior a esto se calcula el nivel de aceptación de este sistema a través del cálculo con una fórmula.

Nº	Afirmación
1	Creo que el sistema es fácil de usar
2	Creo que necesitaría el soporte de un técnico para poder usar el sistema.
3	La estructura del sistema y sus elementos me parece adecuada.
4	Necesito aprender muchas cosas antes de manejar el sistema.
5	Encuentro el sistema consistente en sus funcionalidades.
6	Creo que el diseño y uso del color es inadecuado.
7	Pienso que los indicadores son visualmente llamativos.
8	Encuentro la información de los indicadores y reportes irrelevante para el proceso de negocio.
9	Creo que el sistema sirve de ayuda para tomar mejores decisiones en el negocio.
10	Me siento insatisfecho con el producto en general.

TABLA 9 AFIRMACIONES PARA LA PRUEBA DE USABILIDAD, FUENTE: THOMAS, N. (2015, 13 JULIO).

El nivel de acuerdo a la afirmación fue medida con cinco opciones de valoración (1-5).

1. Desacuerdo Completamente.
2. Desacuerdo.
3. Ni acuerdo ni desacuerdo.
4. Acuerdo.
5. Acuerdo Completamente.

A partir de sus resultados y mediante una fórmula se calcula el valor de aceptación de entre cero y cien puntos, en la tabla 10 se pueden ver los diferentes resultados posibles.

Evaluación	Resultado
$X \geq 80,3$	El sistema es aceptado.
$51 < X < 80,3$	El sistema es aceptado pero requiere mejoras.
$X \leq 51$	El sistema no es aceptado y requiere muchas reformas.

TABLA 10 TABLA DE RESULTADOS DE PRUEBA SUS, FUENTE: THOMAS, N. (2015, 13 JULIO).

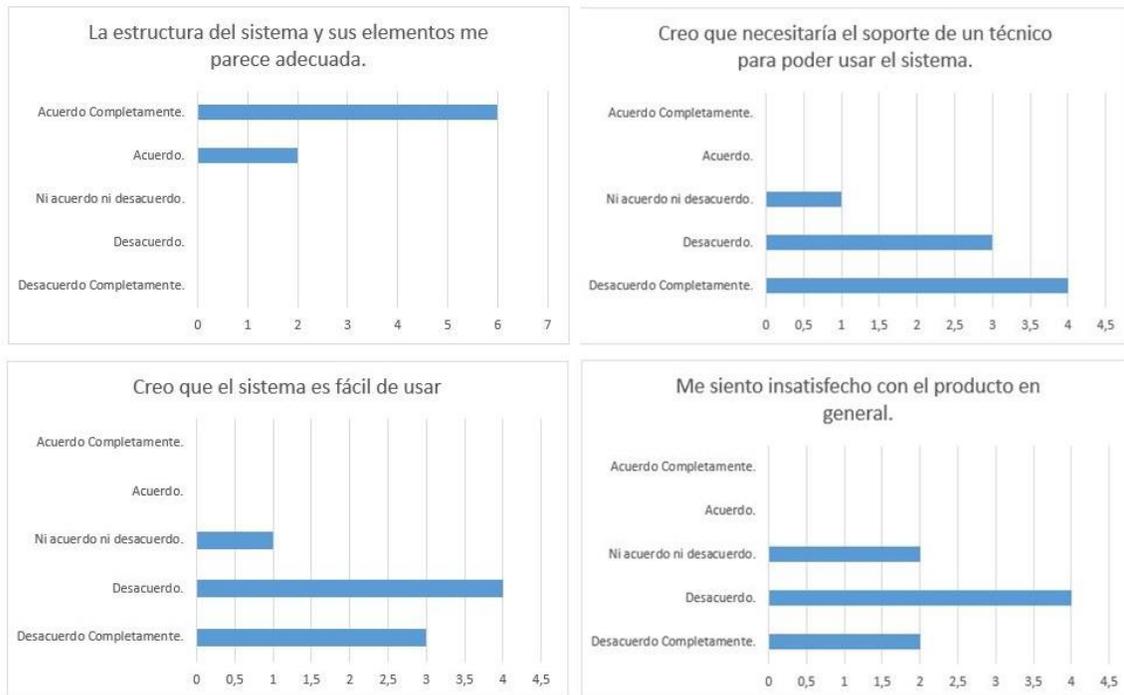
Esta prueba fue realizada sobre una selección de 3 personas:

- 1 Analista de calidad de la empresa de seguros

- 1 inspector de seguros
- 1 Analista de Sistemas de la empresa.
- 5 Personas de otros rubros.

Cada persona evaluada tuvo la oportunidad de probar el sistema en sus diferentes módulos sin mayor información que el proceso de negocio. Los resultados arrojados para cada pregunta se pueden ver reflejados en la Figura #.

Aplicando el método SUS a los resultados se obtiene que el nivel de aceptación es de 84,75 de 100 puntos, por lo que se puede concluir que la solución es aceptada por los usuarios más sin embargo se toma en cuenta que el mismo puede aún mejorar dada la cercanía del resultado al rango subsiguiente.



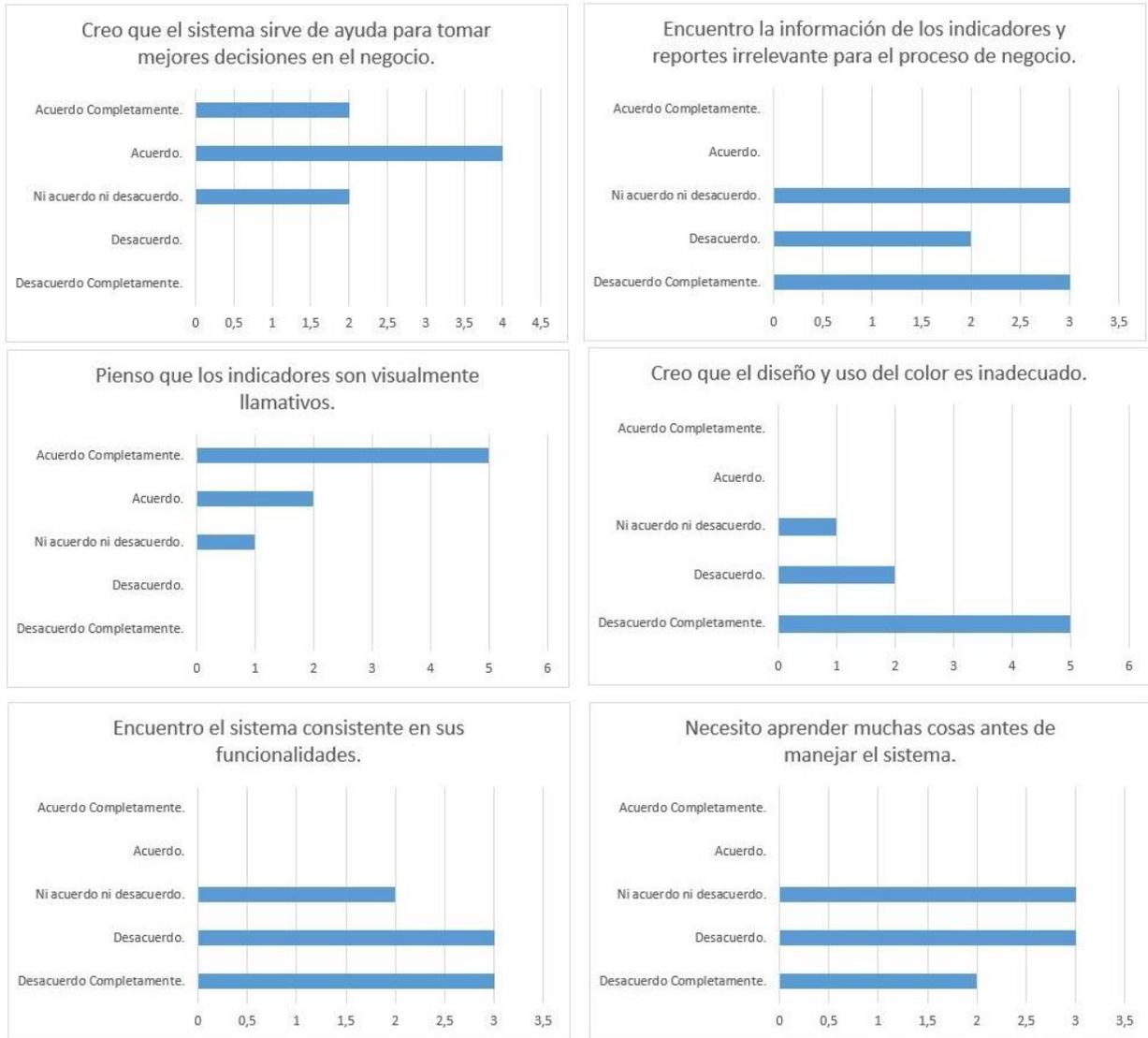


FIGURA 58 RESULTADOS DE LA PRUEBA USABILIDAD.

CONCLUSIONES Y RECOMENDACIONES

Gracias a la obtención del conjunto de resultados en el presente trabajo especial de grado, se ha logrado el objetivo general propuesto, que se fundamentó en el desarrollo de un sistema de información mediante una herramienta de inteligencia de negocio a fin de permitir, la emisión de indicadores para la ejecución de un análisis amplio sobre el proceso de solicitud de citas para la inspección de vehículos dentro de la empresa de seguros Horizonte.

El presente sistema de inteligencia de negocios fue desarrollado exitosamente siguiendo en forma detallada la metodología propuesta por Ralph Kimball, mejor conocida como Business Dimensional Lifecycle (BDL), todo esto acompañado de un estudio sobre el proceso de control y seguimiento a la gestión de solicitud de citas para la inspección de vehículos dentro de la compañía de seguros, para lograr el levantamiento de requerimientos que permitieran en análisis de estos y eventualmente la construcción y emisión de los indicadores que facilitarían el monitoreo de dicho proceso de gestión. Es importante de igual forma destacar que uno de los factores claves para conseguir los objetivos en los tiempos propuestos, fue haber llegado a cada cota definida en la que se emitían entregables dentro de las diferentes fases de la metodología.

De igual forma, la instanciación de la metodología previamente mencionada, permitió centrar esfuerzos en crear actividades, como ya lo hemos mencionado anteriormente, que estaban destinadas a la planificación del proyecto, la definición y el análisis de los requerimientos, al igual que la selección mejor ajustada de las herramientas tecnológicas para el desarrollo de este, las cuales, fueron completamente seleccionadas, bajo la filosofía de Software Libre como lo es la suite de Pentaho Business Intelligence en su versión Community (Pentaho Data Integration, Pentaho Schema Workbench, Pentaho Report Designer y Pentaho BI server) y el Sistema Manejador de Base de Datos PostgreSQL.

Es importante resaltar en primer lugar que la plataforma de la Suite Pentaho Business Intelligence en su versión Community Edition resultó apropiada para el desarrollo del sistema, ya que permitió realizar de manera no solo eficiente pero además eficaz los procesos de extracción, transformación y carga (ETL) de la data, además de esto el modelado multidimensional, el diseño y emisión de los distintos reportes, la creación de las diversas vistas dinámicas de los cubos OLAP y la visualización de los tableros o paneles de control, todos estos, los cuales, se integraron de manera automatizada en el servidor de aplicaciones propio de la plataforma. De igual manera, el sistema manejador de base de datos PostgreSQL se desempeñó cabalmente tanto en la base de datos del sistema transaccional ya implementado dentro de la organización, la simulación de dicha base de datos para el sistema operacional y la base de datos multidimensional, permitiendo así desarrollar correctamente el diseño físico de esta.

Seguido de esto, la Herramienta Pentaho BI Server se ajustó perfectamente al funcionamiento para la realización de los tableros de control, ya que la misma facilitó el diseño y la creación de los mismos, con un diseño que resultó tanto atractivo como funcional. Es por esta y muchas otras razones que hacemos recomendable el uso de la suite Pentaho Business Enterprise Edition como herramienta de desarrollo para proyectos con carácter de análisis de datos ya que posee todas las funcionalidades integradas que permiten el desarrollo de una aplicación de inteligencia de negocios y, aun cuando su versión Community posee limitaciones, se adapta perfectamente a las necesidades iniciales de requerimientos solicitados.

Finalmente y como fue mencionado al comienzo, obtuvimos como resultado una solución de inteligencia de negocio mediante la cual se integran datos del sistema transaccional en un almacén de datos aprovechando el almacenamiento eficiente de los datos que ayuda a tener acceso a grandes volúmenes de información de forma rápida y sencilla, optimizando, en consecuencia, el proceso de análisis para que de esta forma se puede obtener a través de los reportes, el tablero de control y un cubo de información la visión general y específica sobre el desempeño asociado al proceso de gestión en la solicitud de citas para la inspección de vehículos al contar con indicadores de ejecución y gestión que aportan información para el control y seguimiento que se les realiza a los mismos y de esta manera visualizar las áreas más vulnerables y apoyar la toma de decisiones para realizar las correcciones oportunas, optimizar el uso de los recursos y aumentar el desempeño en general, representando así un avance de gran valor para la organización donde fue implementado.

BIBLIOGRAFÍA

Abraham Silberschatz, Henry F. Korth, & S. Sudarshan. (2002). Fundamentos de bases de datos. (4ta. Edición). Mc Graw Hill.

Cano, J. L. (2007). Business Intelligence: Competir con Información. España: ESADE.

Connolly, T, Begg, C (2005). Sistema de Base de Datos. (4a. Ed.). Pearson Educación, Madrid.

Daniel Cohen Karen, & Enrique Asín Lares. (2009). Tecnologías de información en los negocios. (5ta. Edición). Mc Graw Hill.

Davenport, T., & Prusak, L. (1998). Working Knowledge. Boston: Harvard Business Scholl Press.

Héctor García-Molina, Jeffrey D. Ullman, & Jennifer Widom. (2002). Database Systems: The complete Book. (2da. Edición). Prentice Hall.

Inmon, B. (1996). Building the Operational Data Store. Nueva York: John Wiley and Sons.

Inmon, B. (2000). Data mart Does Not Equal Data Warehouse. EEUU: Wiley.

Inmon, B. (2002). Building the Data Warehouse. Nueva York: WILEY.

Kendall, K & Kendall, J. (2004). Análisis y Diseño de Sistemas de Información. Prentice Hall.

Kendall, K & Kendall, J. (2011). Análisis y Diseño de Sistemas. (8va. Ed.). Pearson Educación, México.

Kimball. (1996). The Data Warehouse Toolkit. EEUU: Wiley.

Kimball, R. (1998). The Data warehouse lifecycle toolkit. EEUU: Wiley.

Kimball, R., & Caserta, J. (2008). The Data Warehouse ETL Toolkit. Nueva York: Wiley.

Kimball, R., & Ross, M. (2002). *The Data Warehouse*. New York: Wiley Computer Publishing.

Laudon, K., & Laudon, J. (2012). *Sistemas de información Gerencial*. Mexico: Pearson.

Paulraj Ponniah. (2001). *DATA WAREHOUSING FUNDAMENTALS: A Comprehensive Guide for IT Professionals*. A Wiley-Interscience Publication.

Rafael Camps Paré, Carme Martín Escofet, Marc Gibert Ginestà, Dolors Costal Costa
Luis Alberto Casillas Santillán, & Oscar Pérez Mora. (2005). *Software libre: Bases de datos*.
UOC Formación de Postgrado.

Ralph Stair, & George Reynolds. *Fundamentals of Information Systems*. (6ta. Edición).
CENGAGE Learning.

Ralph Stair, & George Reynolds. (2010). *Principios de Sistemas de Información*. (9na. Edición).
CENGAGE Learning.

Ramez Elmasri, & Shamkant B. Navathe. (2007). *Fundamentos de Sistemas de Bases de Datos*. (5ta. Edición). Pearson Educación, Madrid.

Stair, R., & Reynolds, G. (2000). *Principios de Sistemas de información*. Mexico: Ciencias THOMSON.

Oracle Business Intelligence Foundation Suite: Technical Overview. (12 de abril de 2013).
<https://go.oracle.com/LP=35713?elqCampaignId=52194&src1=ad:pas:go:dq:bi&src2=wwm k160603p00071c0020&SC=sckw=WWMK160603P00071C0020&mkwid=suz2qB4TO|pcrid|105723472604|pkw|oracle%20bi|pmt|e|pdv|c|sckw=srch:oracle%20bi>

Oracle Business Intelligence 12c: Overview.
<https://www.oracle.com/es/solutions/business-analytics/business-intelligence/index.html>

Pentaho.
<http://community.pentaho.com/>

Pentaho: 2016 Dresner Big Data Analytics Market Study.

<http://www.pentaho.com/resources/2016-dresner-big-data-analytics-market-study-analyst-research>

Tableau: Análisis de autoservicio a gran escala.

<https://www.tableau.com/es-es/products/server>

Tableau.

<https://www.tableau.com/es-es/about>

Tableau: Tableau Desktop.

<https://www.tableau.com/es-es/products/desktop>

Tableau: Tableau products.

<https://www.tableau.com/es-es/products>

Enlaces adicionales:

<https://www.trustradius.com/reviews/oracle-business-analytics-2014-05-13-09-03-46>

<https://www.experfy.com/blog/pentaho-vs-tableau-comparison-visualization-dashboards>

<https://decisionesytecnologia.wordpress.com/>

<https://sites.google.com/site/tdadmonealtamirano/home/unidad4mecanicadelprocesoadministrativo/4-2-organizacion/4-2-2-proceso-de-organizacion-tipos-y-tecnicas-de-organizacion>

<http://www.gestiopolis.com/indicadores-de-gestion-que-son-y-por-que-usarlos/>

<https://es.wikipedia.org/wiki/Medici%C3%B3n>

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

https://es.wikipedia.org/wiki/Sistema_de_informaci%C3%B3n

https://es.wikipedia.org/wiki/Almac%C3%A9n_de_datos

<https://churriwifi.wordpress.com/2010/04/19/15-2-ampliacion-conceptos-del-modelado-dimensional/>

<http://datawarehouse.es.tl/Ventajas-y-desventajas.htm>

<http://www.dataprix.com/blogs/respinosamilla/qu-business-intelligence>