



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

**SOLUCIÓN DE INTELIGENCIA DE NEGOCIO PARA LA OBTENCIÓN DE  
INDICADORES DE CONTROL Y SEGUIMIENTO DE PROYECTOS DE OBRA  
CIVIL**

Trabajo Especial de Grado presentado ante la ilustre  
Universidad Central de Venezuela por  
Br. Edward Vieras.  
Para optar al título de Licenciado en Computación

**Tutor:**  
Prof. Franky Uzcátegui

Octubre, 2016

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 “Solución de Inteligencia de Negocio para la obtención de Indicadores de Control y Seguimiento de Proyectos de Obra Civil” y presentado por el bachiller: Edward Alejandro Vieras titular de la Cédula de Identidad V-19.371.014, 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 21 de octubre de 2016, a las 11:00 horas, para que el autor lo defendiera en forma pública, lo que este hizo en la Sala PB3 de la Escuela 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 \_\_\_\_ puntos.

En fe de lo cual se levanta la presente Acta, en Caracas el día 21 de octubre de 2016.

---

**Prof. Franky Uzcátegui**

**(Tutor)**

---

**Prof. Franklin Sandoval**

**(Jurado)**

---

**Prof. Mercy Ospina**

**(Jurado)**

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 "Solución de Inteligencia de Negocio para la obtención de Indicadores de Control y Seguimiento de Proyectos de Obra Civil" y presentado por el bachiller: Edward Alejandro Vieras titular de la Cédula de Identidad V-19.371.014, 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 21 de octubre de 2016, a las 11:00 horas, para que el autor lo defendiera en forma pública, lo que este hizo en la Sala PB3 de la Escuela de Computación, mediante una presentación oral de su contenido, luego de lo cual respondieron a las preguntas formuladas. Finalizada la defensa pública del Trabajo Especial de Grado, el jurado decidió aprobar con la nota de 20 puntos.

En fe de lo cual se levanta la presente Acta, en Caracas el día 21 de octubre de 2016.



Prof. Franky Uzcátegui  
(Tutor)



Prof. Franklin Sandoval  
(Jurado)



Profa. Mercy Ospina  
(Jurado)

## AGRADECIMIENTOS

Este ha sido uno de los momentos más esperado hasta ahora en mi vida, sentir la satisfacción que te deja el concluir una tarea tan titánica y emblemática, alcanzar una meta y obtener un logro más, un logro que no es solo mío, sino que también corresponde a todas aquellas personas que de manera directa o indirecta me han ayudado y han estado allí cuando lo he necesitado.

En primer lugar, quiero agradecer a mi madre María Edén Vieras, por todo su apoyo incondicional, por ser una mujer increíble y única, por siempre estar allí en las buenas y en las malas y por todos sus esfuerzos para que tuviera la oportunidad que muchos lamentablemente no tienen. En segundo lugar, a mis hermanos Frank y Cori, por ser personas admirables cada uno a su manera. En tercer lugar, a mi familia, mi tía, mi madrina, mis primas locas y mi padrino, por toda la paciencia y por demostrar siempre que la sangre es más espesa que el agua. También quiero agradecer a mi novia Walli, de la cual he aprendido muchísimo y que ha ayudado a forjar el ser humano que soy hoy en día.

Es importante mencionar el apoyo y la ayuda del Fondo de Compensación Interterritorial, en particular del Ing. Jhon Pérez, quien está a cargo de la Dirección de Seguimiento y Evaluación de Políticas Públicas del Fondo de Compensación Interterritorial, así como también al Ing. Smith Hurtado con el desarrollo y los cálculos de los indicadores. Además, quiero agradecer al Ing. Jorge Rengifo y a la Dra. Joali Moreno por el apoyo brindado.

Igualmente quiero darle un agradecimiento especial a mi tutor, el profesor Franky Uzcátegui, por ser un excelente guía, por su alto nivel de profesionalismo, por su dedicación y enseñanzas, este trabajo no sería el mismo sino fuera por él.

Por último y no menos importante, Gracias a Dios.

## **DEDICATORIA**

Es un gran honor para mí dedicarle este logro a la mujer que me ha enseñado más que nadie en la vida, a esa gran mujer que se esforzó y sacrifico tanto para que mis hermanos y yo tengamos las oportunidades para superarnos y lograr todas nuestras metas.

Eres una de las personas que más me inspira y que más admiro en la vida...

Todo lo que tengo no hubiera sido posible sin ti...

Todo lo que he alcanzado te lo agradezco a ti...

A ti que me diste todo, sin pedir nada...

A ti que dejaste todo por mi...

A ti que entregaste todo por mi...

¡Gracias Mamá!

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

**SOLUCIÓN DE INTELIGENCIA DE NEGOCIO PARA LA OBTENCIÓN DE  
INDICADORES DE CONTROL Y SEGUIMIENTO DE PROYECTOS DE OBRA  
CIVIL**

**Autor:** Br. Edward Vieras.

**Tutor:** Prof. Franky Uzcátegui.

**Fecha:** Octubre, 2016

**RESUMEN**

La planificación y ejecución de Proyectos de Obra Civil, requieren un conjunto de indicadores que permiten asegurar el éxito en el cumplimiento de los mismos. Por ello, este trabajo se enfoca en diseñar, definir y formular el desarrollo de una solución de Inteligencia de Negocio, para medir el desempeño de la gestión sobre este tipo de proyectos. Para ello se presentan conceptos, metodologías, esquemas y arquitecturas, que permiten la obtención de indicadores específicos, que permiten tomar decisiones oportunas y así garantizar el manejo eficiente y eficaz de los recursos invertidos. La solución de Inteligencia de Negocio, se fundamenta en la metodología de análisis ascendente (Ralph Kimball), para la cual se hace uso de herramientas de software libre, (PostgreSQL, Pentaho BI Server (PBI), Pentaho Schema Workbench (PSW), Pentaho Data Integration (PDI) y Pentaho Report Designer (PRD))

**Palabras Clave:** Inteligencia de Negocio, Indicadores, Metodología de Ralph Kimball, Proyectos, Obra Civil.

# ÍNDICE DE CONTENIDO

<b>INTRODUCCIÓN</b> .....	<b>i</b>
<b>CAPÍTULO 1</b> .....	<b>1</b>
<b>PROBLEMA DE INVESTIGACIÓN</b> .....	<b>1</b>
1.1. Planteamiento del problema .....	1
1.2. Objetivos.....	3
1.2.1. Objetivo General .....	3
1.2.2. Objetivos Específicos.....	3
1.3. Solución Propuesta .....	3
1.4. Justificación .....	5
1.5. Alcance .....	7
<b>CAPÍTULO 2</b> .....	<b>8</b>
<b>MARCO CONCEPTUAL</b> .....	<b>8</b>
2.1. Dato, información, conocimiento y Sistema .....	8
2.1.1. Dato .....	8
2.1.2. Información .....	8
2.1.3. Conocimiento .....	10
2.1.4. Sistema .....	11
2.2. Sistemas de Información .....	11
2.2.1. Características de los Sistemas de Información .....	13
2.2.2. Tipos de Sistemas de Información .....	13
2.2.3. Comparación entre los Sistemas OLTP y los Sistemas OLAP .....	18
2.3. Base de Datos .....	20
2.3.1. Base de Datos Relacional.....	21
2.4. Almacén de Datos (Datawarehouse). .....	22
2.4.1. Objetivos de un Almacén de Datos .....	22
2.4.2. Características de un Almacén de Datos .....	23
2.4.3. Bodega de Datos (Data Mart) .....	24
2.4.4. Diferencia entre una Bodega de Datos y un Almacén de Datos .....	25
2.4.5. Modelo Dimensional.....	26
2.4.6. Ventajas y Desventajas de los Almacenes de Datos .....	34
2.5. Inteligencia de Negocio (BI) .....	35
2.5.1. Característica de una Solución de Inteligencia de Negocio .....	35
2.5.2. Funciones de una Solución de Inteligencia de Negocio.....	36
2.5.3. Arquitectura de una Solución de Inteligencia de Negocio.....	36

2.6. Tecnologías y herramientas para la construcción de una solución de Inteligencia de Negocios .....	39
2.6.1. Pentaho BI.....	41
2.6.2. Oracle Business Intelligence (OBI) .....	48
2.6.3. Tableau .....	55
2.6.4. Comparación entre las herramientas de inteligencia de negocios.....	57
2.7. Proyecto .....	57
2.7.1. Ciclo de vida de los proyectos .....	59
2.7.2. Características de un proyecto .....	61
2.7.3. Gerencia de proyectos .....	62
2.8. Indicadores .....	65
2.8.1. Características de los Indicadores .....	66
2.8.2. Tipos de Indicadores .....	66
2.8.3. Objetivos de los Indicadores .....	67
2.8.4. Paradigmas de los indicadores .....	67
2.8.5. Utilidad de los Indicadores.....	68
2.8.6. Ventajas de los Indicadores.....	69
2.9. Obra Civil .....	69
2.9.1. Definición.....	69
2.9.2. Características de las obras civiles .....	70
2.9.3. Tipos de obras civiles.....	71
2.9.4. Proyectos de Obra Civil .....	71
<b>CAPÍTULO 3 .....</b>	<b>76</b>
<b>MARCO METODOLÓGICO .....</b>	<b>76</b>
3.1. Metodología según Ralph Kimball (Ascendente o Bottom-up).....	77
3.1.1. Planificación del proyecto.....	80
3.1.2. Definición de requerimientos del negocio. ....	81
3.1.3. Modelado dimensional.....	82
3.1.4. Diseño físico. ....	82
3.1.5. Diseño y construcción de procesos ETL.....	82
3.1.6. Diseño Técnico de la arquitectura.....	83
3.1.7. Selección e instalación de productos.....	84
3.1.8. Especificaciones de las Aplicaciones Analíticas.....	85
3.1.9. Desarrollo de Aplicaciones Analíticas .....	87
3.1.10. Implementación .....	87
3.1.11. Mantenimiento .....	88
3.1.12. Escalabilidad .....	88

3.1.13. Gestión del proyecto .....	89
<b>CAPÍTULO 4 .....</b>	<b>90</b>
<b>MARCO APLICATIVO .....</b>	<b>90</b>
4.1. Fases del Proyecto .....	90
4.1.1. Planificación del Proyecto.....	90
4.1.2. Definición de los Requerimientos del Negocio .....	92
4.1.3. Diseño Técnico de la Arquitectura.....	98
4.1.4. Selección de Productos e Instalación .....	100
4.1.5. Modelado Dimensional .....	102
4.1.6. Diseño Físico.....	107
4.1.7. Diseño y Desarrollo de Procesos ETL .....	108
4.1.8. Especificaciones de las Aplicaciones Analíticas.....	112
4.1.9. Desarrollo de las Aplicaciones Analíticas .....	114
4.1.10. Implementación.....	121
4.1.11. Gestión del proyecto .....	134
<b>CONCLUSIONES Y RECOMENDACIONES.....</b>	<b>136</b>
<b>REFERENCIAS BIBLIOGRÁFICAS Y DÍGITALES.....</b>	<b>138</b>

## ÍNDICE DE ILUSTRACIONES

Ilustración 1: Arquitectura propuesta para la solución de inteligencia de negocio. ....	5
Ilustración 2: Relación entre dato e información. (davenport & prusak, 1998) .....	9
Ilustración 3: Tipos de sistema de información. (laudon & laudon, 2008) .....	14
Ilustración 4: Comparación entre los sistemas oltp y los sistemas olap. (pinto, 2013) .....	19
Ilustración 5: Ejemplo de un diagrama de base de datos relacional. ....	21
Ilustración 6: Ejemplo de jerarquía geográfica. ....	27
Ilustración 7: Ejemplo de granularidad por tiempo. ....	28
Ilustración 8: Ejemplo de tabla de hecho y sus dimensiones.....	31
Ilustración 9: Ejemplo de un esquema estrella .....	32
Ilustración 10: Ejemplo de un esquema copo de nieve.....	33
Ilustración 11: Ejemplo de un esquema de constelación de hechos. ....	33
Ilustración 12: Arquitectura de una solución de inteligencia de negocios .....	37
Ilustración 13: Arquitectura del flujo de datos con pentaho community (mauricio, 2012)..	42
Ilustración 14: Interfaz del entorno de desarrollo de spoon. (herramienta pentaho data integration).....	43
Ilustración 15: Interfaz del entorno de desarrollo de pentaho report desing. (herramienta pentaho report desing). ....	45
Ilustración 16: Interfaz del entorno de desarrollo de schema workbench. ....	46
Ilustración 17: Ejemplo de un dashboard elaborado con pentaho dashboards. (pentaho a. H., 2005) .....	48
Ilustración 18: Arquitectura de oracle bi server. (oracle corporation, s.f.).....	49
Ilustración 19: Interfaz de oracle bi admin tool. (oracle corporation, s.f.) .....	51
Ilustración 20: Interfaz de obi answers. (bi-insider, 2011) .....	51
Ilustración 21: Interfaz de obi interactive dashboards. (bi-insider, 2011) .....	52
Ilustración 22: Ejemplo de bi publisher. (oracle corporation, s.f.) .....	53
Ilustración 23: Ambiente de trabajo de owb. (oracle corporation, s.f.) .....	54
Ilustración 24: Visualización de tableau desktop. (tableau, s.f.) .....	56
Ilustración 25: Modificación de interfaz según resolución de pantalla en tableau server. (loyola university chicago, s.f.) .....	56
Ilustración 26: Ciclo de vida de los proyectos .....	60
Ilustración 27: Puntos de vista de los proyectos. ....	63
Ilustración 28: Metodología bottom-up .....	77
Ilustración 29: Ciclo de vida dimensional del negocio. (kimball & ross, 2013) .....	79
Ilustración 30: Arquitectura de la solución de inteligencia de negocios .....	99

Ilustración 31: Relaciones jerárquicas del modelo dimensional de la solución propuesta.	105
Ilustración 32: Dimensión role-playing tiempo .....	105
Ilustración 33: Modelo dimensional de la solución propuesta. ....	107
Ilustración 34: Modelo físico del almacén de datos implementado en la base de datos.....	108
Ilustración 35: Repositorio de archivos .....	109
Ilustración 36: ETL para cargar la dimensión entidad.....	110
Ilustración 37: ETL para cargar la tabla de hechos .....	111
Ilustración 38: Trabajo (job) que ejecuta todas las transformaciones .....	111
Ilustración 39: Plantilla estándar para la creación de reportes.....	113
Ilustración 40: Vista del esquema de datos desarrollado con pentaho schema workbench	115
Ilustración 41: Publicación de un esquema de datos en el servidor pentaho bi server. ....	116
Ilustración 42: Selección del esquema de datos y creación de las consultas mdx.....	116
Ilustración 43: Diseño de la estructura del reporte .....	117
Ilustración 44: Capa community dashboard framework (cdf).....	118
Ilustración 45: Capa community chart components (ccc) .....	118
Ilustración 46: Capa community data access (cda).....	119
Ilustración 47: Indicadores a nivel nacional .....	119
Ilustración 48: Indicadores a nivel de redi.....	120
Ilustración 49: Indicadores a nivel de estado.....	120
Ilustración 50: Indicadores a nivel de entidad político-territorial.....	120
Ilustración 51: Indicadores a nivel de proyecto .....	121
Ilustración 52: Portal web para acceder al tablero de control, a los reportes y a cubo de información mondrian.....	121
Ilustración 53: Comparación entre la fuente de datos y la dimensión proyectos. ....	122
Ilustración 54: Comparación entre la fuente de datos y la dimensión entidad. ....	123
Ilustración 55: Comparación entre la fuente de datos y la dimensión ubicación geográfica .....	123
Ilustración 56: Comparación entre la fuente de datos y la dimensión estatus rendición ....	124
Ilustración 57: Comparación entre la fuente de datos y la dimensión estatus carga de proyecto. ....	124
Ilustración 58: Comparación entre la fuente de datos y la dimensión estatus de avance físico .....	125
Ilustración 59: Comparación entre la fuente de datos y la dimensión categorización.....	125
Ilustración 60: Comparación entre la fuente de datos y la dimensión categorización.....	126
Ilustración 61: Comparación entre la fuente de datos y los hechos del almacén de datos .	126
Ilustración 62: Página de inicio para ingresar a las herramientas de visualización de la información.....	128

Ilustración 63: Componentes del tablero de control .....	130
Ilustración 64: Componentes del segundo, tercer y cuarto nivel del tablero de control.....	131
Ilustración 65: Componentes del último nivel de despliegue. ....	132
Ilustración 66: Componentes de los reportes .....	133
Ilustración 67: Componentes del cubo de información .....	134

## ÍNDICE DE TABLAS

Tabla 1: Comparación entre los sistemas oltp y los sistemas olap .....	20
Tabla 2: Comparación de las herramientas de inteligencia de negocios .....	57
Tabla 3: Actividades a realizar para implementar la solución de inteligencia de negocios .	91
Tabla 4: Variables provenientes del sistema transaccional .....	93
Tabla 5: Herramientas para desarrollar la solución de inteligencia de negocio propuesta.	101
Tabla 6: Dimensiones del modelo dimensional de la solución propuesta. ....	104
Tabla 7: Hechos utilizados en la solución propuesta. ....	106

## INTRODUCCIÓN

Las obras civiles son desarrolladas para beneficio de la población de una nación, debido a que algunos de los objetivos de las mismas son la organización territorial y el aprovechamiento al máximo del territorio. Es así como carreteras, calles, autopistas, puentes, vías de ferrocarril, aeropuertos y puertos permiten la circulación de los diversos medios de transporte y el alcantarillado, los canales y las represas se encargan de administrar los recursos hídricos de un territorio. Es por ello, éstas obras civiles utilizan una variedad de conocimientos que combinan la física, el cálculo, la mecánica, el álgebra, entre otras; contribuyendo a la hora del diseño, en la construcción y el mantenimiento de las infraestructuras mencionadas anteriormente. Además, en vista de las características propias de los proyectos de obras civiles, a la inversión de los recursos y la disposición de un espacio territorial, llevar a cabo un proyecto de obra civil requiere de una buena planificación y de un constante control y seguimiento de su ejecución

Asimismo, en la República Bolivariana de Venezuela, las Entidades Político-Territoriales<sup>1</sup> son unos de los mayores planificadores y ejecutores de proyectos de Obra Civil y el Consejo Federal del Gobierno es uno de los organismos que mayor financiamiento les proporciona, por lo cual estas entidades deben regirse por las disposiciones emanadas en la Ley Orgánica del Consejo Federal de Gobierno, además de la Ley Orgánica del Ambiente, la Ley Orgánica para la Planificación y Gestión de la Ordenación del Territorio, entre otras.

Conjuntamente, estas entidades pueden planificar, administrar y ejecutar cientos e incluso miles de proyectos de obra civil en un momento determinado, y el Consejo Federal de Gobierno dispone de un sistema transaccional para que estas entidades puedan registrar los datos de los proyectos, y realizar las rendiciones de cuenta exigidas en las leyes antes mencionadas; generando así, una gran cantidad de datos de los cuales se requiere obtener información precisa a través de reportes e indicadores, que faciliten el control y seguimiento

---

<sup>1</sup> **Entidad Político-Territorial:** Se entiende por tales, a las Alcaldías y Gobernaciones de los Estados y Municipios, incluyendo al Distrito Capital, el Distrito del Alto Apure y Distrito Metropolitano.

a los proyectos y permita conocer el estado en el que estos se encuentran, de esta manera, poder tomar decisiones que permitan mejorar la planificación, ejecución y gestión de los mismo.

Lo anteriormente expuesto, conlleva al almacenamiento de una gran cantidad de datos para ser procesados y transformados en información útil, sin embargo, muchas entidades e instituciones que realizan control y seguimiento sobre los proyectos de obra civil que estas planifican y ejecutan, no poseen un sistema analítico dedicado; ocasionando eventualmente problemas para realizar análisis y monitoreos en función a los datos recolectados, los cuales afectan la correcta toma decisiones. De este modo, se hace evidente la necesidad de implementar una Solución de Inteligencia de Negocio, como una alternativa tecnológica que permita manejar la información desde la extracción, depuración y transformación de datos, hasta la exploración, distribución y visualización de la información mediante herramientas de fácil uso, que apoyen a las Entidades Político-Territoriales y a las instituciones relacionadas en el proceso de toma de decisiones para fortalecer sus procesos organizacionales respecto al control y seguimiento de estos proyectos.

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 a proyectos de obra civil, planificados por Entidades Político-Territoriales.

Dicha investigación está estructurada de la siguiente manera:

El Capítulo I, 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. A continuación, en el Capítulo II se muestra el marco conceptual, en el cual se presentan los fundamentos teóricos que servirán de base para dar soporte al desarrollo de la misma y se indicará algunas herramientas para generar soluciones de Inteligencia de Negocio existentes en el mercado actual. Posteriormente, en el Capítulo III, 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. Luego en el Capítulo IV, se especifican las actividades que se realizaron en cada una de las fases de la metodología definida para la solución planteada. Por último, se presentan las conclusiones, referencias bibliográficas y digitales consultadas y utilizadas para la elaboración de esta investigación.

En conclusión, hay que recalcar la importancia de esta investigación, por ser una alternativa tecnológica que permitirá manejar la información mediante herramientas de fácil uso en apoyo a las Entidades Político-Territoriales y a las instituciones relacionadas en el proceso de toma de decisiones para fortalecer sus procesos con respecto al control y seguimiento de sus proyectos.

## CAPÍTULO 1

### PROBLEMA DE INVESTIGACIÓN

#### 1.1. Planteamiento del problema

Hoy en día, uno de los objetivos de las organizaciones es mejorar y optimizar sus procesos de negocios, medir el éxito de los productos, o bienes y servicios, y tomar decisiones eficaces y eficientes, para así tener un mayor alcance en el mercado, con sus clientes o con la sociedad.

Usualmente las organizaciones utilizan sistemas transaccionales para registrar sus operaciones diarias, y estos sistemas presentan los datos bajo el enfoque en el que fueron contruidos, lo cual, para el proceso de obtención de información, trae inconvenientes como: largos tiempos de extracción y procesamiento, problemas al valorar un dato de acuerdo a su fuente de extracción y dificultades al momento de consolidar e interpretar la información para los usuarios de alto nivel en las organizaciones. Como consecuencia, eventualmente tienen problemas para realizar análisis y monitoreos en función a los datos recolectados lo cual afecta en la correcta toma decisiones.

No obstante, las organizaciones que administran, financian, planifican y/o ejecutan proyectos de obra civil no se escapan de esta realidad y en particular las Entidades Político-Territoriales e instituciones que realizan control y seguimiento sobre los proyectos de obra civil, ya que para estas organizaciones es muy importante conocer la evolución de los proyectos en todas sus fases; desde el inicio del proyecto hasta el cierre. Para ello utilizan diferentes mecanismos, técnicas y herramientas para realizar el control y seguimiento de estas fases, pero en muchos casos este proceso es muy lento y engorroso o no está automatizado.

Asimismo, la planificación y ejecución de los proyectos a cualquier nivel siempre ha conllevado una serie de desafíos, entre los cuales se destaca el incumplimiento en los tiempos de entrega y gastos mayores al presupuesto asignado inicialmente, entre otros. Pero, el

---

mantener un constante control y seguimiento sobre los proyectos y estudiar las dificultades que se puedan presentar, es una tarea crítica, en vista que, si no son previstos y solventados a tiempo, puede significar el fracaso del proyecto, pérdidas enormes de recursos o consecuencias relevantes para los involucrados.

En consecuencia, realizar el control y seguimiento a los proyectos de obra civil se puede volver una tarea complicada, en vista que para conocer el estado actual de los proyectos se debe prestar mucha atención a los detalles y manejar una gran cantidad de datos y así adaptarse al contexto actual de los proyectos a analizar. Posteriormente, se crean los reportes que contengan información acerca del avance de los proyectos en base a los datos recopilados, pero, si este proceso no está automatizado, los reportes pueden tardar más tiempo en llegar a la parte interesada, e incluso tienen más posibilidades de traer errores humanos.

Además, en vista de que los recursos y alcance de los objetivos de los proyectos de obra civil son limitados, la implementación de indicadores es necesaria para aumentar la eficiencia en la distribución de los recursos y llevar seguimiento a los objetivos planteados y cumplidos, las metas, las limitaciones y su correcta utilización. Este sistema analítico estaría en un ambiente centralizado para el análisis de los datos; base en la cual se podrán generar estadísticas confiables para evaluar temas de interés vinculados con la gestión, administración y control de la ejecución de los proyectos; utilizando herramientas tecnológicas que faciliten la elaboración de reportes e indicadores que apoyen el proceso de toma de decisiones.

## **1.2. Objetivos**

### **1.2.1. Objetivo General**

Desarrollar una Solución de Inteligencia de Negocio para la obtención de indicadores de control y seguimiento a Proyectos de Obra Civil.

### **1.2.2. Objetivos Específicos**

- Definir el modelo dimensional con los indicadores y reportes relacionados con la planificación, ejecución y gestión de los proyectos de obra civil en base a la identificación de los requisitos del negocio.
- Desarrollar un almacén de datos adaptable a la Solución de Inteligencia de Negocio a partir de los requisitos del negocio previamente identificados.
- Desarrollar 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.
- Realizar pruebas de calidad de los datos.

## **1.3. Solución Propuesta**

Se plantea como solución a la problemática antes mencionada, la implementación de un modelo dimensional y un almacén de datos históricos a partir del sistema transaccional que utiliza el Consejo Federal de Gobierno, que permita obtener indicadores de control y seguimiento a proyectos de obra civil, los cuales podrán ser visualizados a través de reportes, un tablero de control y de un cubo de información Mondrian (Ver punto 2.6.2) para que de esta manera los usuarios puedan obtener de forma rápida y flexible información relevante y oportuna que apoye el proceso de toma de decisiones, con el fin de evaluar la gestión de las Entidades Político-Territoriales en cuanto a la planificación y ejecución de los proyectos de obra civil.

Por ello, el tablero de control, los reportes y el cubo de información Mondrian son construidos con herramientas de Inteligencia de Negocio, de modo tal que facilite la construcción de reportes e indicadores a la medida, y su posterior publicación, para su distribución a la alta gerencia según el perfil de acceso del usuario. Igualmente, el uso de estas herramientas permite la generación de contenidos dinámicos tales como gráficos, tablas, filtros, criterios de selección y búsqueda sin necesidad de la asistencia de personal técnico. Además, cabe resaltar que los reportes realizados pueden ser exportados en distintos formatos de presentación.

La solución está conformada, como puede verse en la ilustración 1 por los siguientes componentes:

- Las fuentes 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), realizar procesos de depuración, transformación y operaciones sobre los datos y por último cargarlos en el almacén de datos.
- El Almacén de Datos<sup>2</sup>, 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 Mondrian, los cuales permitirán la visualización de los indicadores e información solicitados por la alta gerencia.

---

<sup>2</sup> Cabe destacar que el almacén de datos es una base de datos con características particulares y su principal papel es integrar datos de diversas fuentes en una única base de datos centralizada, la cual es accesible por las aplicaciones que dan soporte al proceso de toma de decisiones acelerando el proceso de análisis, aprovechando el almacenamiento eficiente y logrando el acceso de grandes volúmenes de información.

---

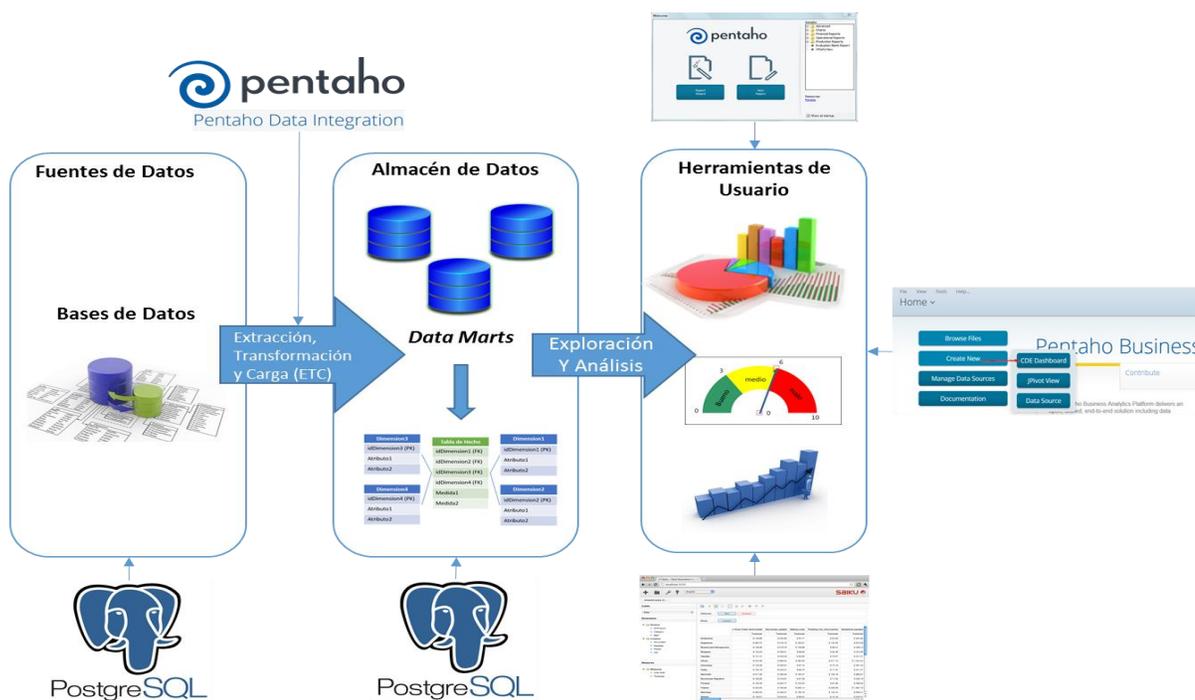


ILUSTRACIÓN 1: ARQUITECTURA PROPUESTA PARA LA SOLUCIÓN DE INTELIGENCIA DE NEGOCIO.

Se utilizará como plataforma tecnológica el Sistema Manejador de Base de Datos (SMBD) Postgresql para el almacenamiento de los datos en el almacén de datos, las herramientas de Pentaho Data Integración para la implementación de los procesos de extracción, transformación y carga de los datos desde de las fuentes de datos al almacén de datos, y las herramientas de Pentaho BI Server, Pentaho Report Designer y Pentaho Schema Workbench para la manipulación, presentación y análisis de los datos y la construcción del tablero de control, los reportes y el cubo de información Mondrian.

#### 1.4. Justificación

Actualmente, se vive en una época en la cual la información es un factor clave para alcanzar la misión, visión y objetivos de una organización, en la que los ejecutivos requieren del acceso rápido y fácil de dicha información para apoyarse para la toma de decisiones, y lograr

beneficios en un plazo de tiempo corto para mantener un desempeño funcional y óptimo dentro de la organización.

Es así, como el escaso tiempo para el análisis de la información complica el hecho de tomar decisiones adecuadas en un entorno sujeto a constantes cambios, más aún cuando las organizaciones se ven enfrentadas a situaciones, en las que se deben tomar decisiones con acierto y celeridad, que determinan las futuras acciones a llevar a cabo. Y, contar con la información histórica y bien procesada de los procesos de negocio, permitirá comparar datos en varios períodos e identificar tendencias.

En vista, que en determinado momento se pueden ejecutar cientos, e incluso miles, de proyectos a la vez, la persona responsable de llevar el control y seguimiento de los proyectos de obra civil debe monitorear constantemente la evolución de los mismos, lo cual hace que se necesite una herramienta que pueda ayudar a discernir prioridades sobre los proyectos que requieran mayor atención, se pueda tener acceso inmediato a información importante y tomar decisiones efectivas para la correcta ejecución de lo planificado; y así corregir las dificultades que se puedan presentar para optimizar la utilización de los recursos asignados.

Además, al tener indicadores de gestión se logra una mejor administración de los proyectos, ya que el responsable del control y seguimiento de los proyectos, no invertirá más tiempo del necesario analizando de manera detallada cada uno de ellos; sino que podrá ver en los indicadores aquellos proyectos que necesitan mayor atención, y además con la generación de reportes podrá estudiar detalladamente dicha información en busca de una solución al problema que se presente, estos reportes se generarán de manera inmediata, facilitando así la información del estado de los proyectos a la parte interesada, y permitiendo un análisis efectivo de la situación en un tiempo adecuado.

Por esta razón, el análisis previo de la problemática planteada en el Capítulo 1, la implementación de una Solución de Inteligencia de Negocio, es una alternativa tecnológica, para que la misma permita toma de decisiones la acertadas en todos los niveles de una organización a través del manejo de la información que conlleva, desde la extracción, depuración y transformación de datos, hasta la exploración y distribución de la información mediante herramientas de fácil uso para los usuarios, con el fin de satisfacer sus necesidades,

---

para así realizar un aporte a las Entidades Político-Territoriales e instituciones que llevan el control y seguimiento de los proyectos de obra civil.

También hay que destacar, que las Soluciones de Inteligencia de Negocio les permiten a las personas encargadas de los negocios ser más productivas al:

- Minimizar 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.
- Evitar que se manipulen manualmente los datos, hecho que puede causar que el resultado de los análisis efectuados sea incorrecto.
- Permitir 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 tableros de control.

### **1.5. Alcance**

El desarrollo de la solución de inteligencia de negocios se enfocará al diseño y despliegue de los indicadores asociados al proceso de gestión de los proyectos de obra civil, específicamente a la fase de control y seguimiento de los proyectos planificados por Entidades Político-Territorial. Con esta solución el responsable del seguimiento podrá tomar las acciones correctivas necesarias, apoyándose en la información generada por la aplicación, esta información será mostrada en el tablero de control, en los reportes propuestos y el cubo de información Mondrian, utilizando los datos del sistema transaccional como fuente para la carga del almacén de datos.

Los indicadores de gestión que se desarrollarán cubrirán las siguientes necesidades:

- Ejecución del proyecto respecto a lo planificado.
- Estado de los objetivos del proyecto con respecto a las metas.

## CAPÍTULO 2

### MARCO CONCEPTUAL

#### 2.1. Dato, información, conocimiento y Sistema

##### 2.1.1. Dato

Según (Burch & Starter, 1981) “los datos son hechos aislados y en bruto, los cuales son el elemento primordial de la información, donde su finalidad fundamental al recopilar y procesar los datos es producir precisamente esa información, la cual aumentará el conocimiento de las personas o reducirá la incertidumbre de las mismas”. Claro que, para los autores (Davenport & Prusak, 1998) “los datos son un conjunto de hechos discretos y objetivos sobre acontecimientos, a los que describe, aunque no dicen nada sobre el porqué de las cosas y por sí mismos tienen poca o ninguna relevancia o propósito”. Por ello, el dato es la base necesaria para poder obtener o generar la información, la cual es necesaria en las organizaciones para una mejor toma de decisiones.

Asimismo, los datos describen únicamente una parte de lo que pasa en la realidad y a pesar de representar hechos reales tienen poco valor más allá de su existencia y no proporcionan juicios de valor o interpretaciones. Son unidades primarias y fundamentales que constituyen la información y el conocimiento.

##### 2.1.2. Información

En ocasiones los términos dato e información se utilizan como sinónimos, lo cual es un error. Ya que, dato como fue definido con anterioridad, puede ser un número, una palabra, una imagen; los cuales son materia prima para la producción de información.

Conjuntamente, los autores (Burch & Starter, 1981) definen la información como la “adición o el procesamiento de los datos, que pueden proporcionar, conocer o entender

ciertos puntos”. Es importante determinar que la información da lugar a la toma de decisiones. También mencionan que la información no debe indicar únicamente lo que está ocurriendo sino lo que va a acontecer, con el fin de lograr contribuir a la planeación, operación, y control de las funciones en la organización.

Por otro lado, para (Davenport & Prusak, 1998) “la información son datos que han sido procesados y que aportan algún valor o significado a los mismos” (como se muestra en la Ilustración 2). Por ello, en el contexto empresarial la información es un recurso intangible, aunque, no toda la información es importante, esto dependerá del fin para el cual sea necesario utilizarla, debido a que deberá cumplir ciertos parámetros como un buen nivel de detalle, precisión, confiabilidad, oportunidad y claridad con que se presenta la misma.

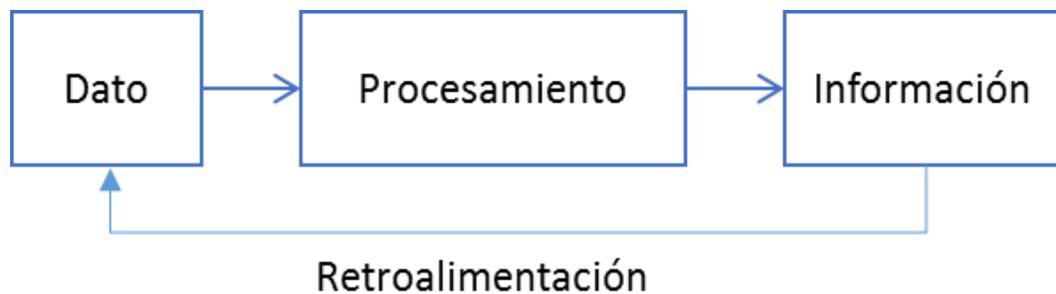


ILUSTRACIÓN 2: RELACIÓN ENTRE DATO E INFORMACIÓN. (DAVENPORT & PRUSAK, 1998)

Mientras que, para (Stair & Reynolds, 2000), “la información es un cumulo de datos, los cuales se van ordenando y en su conjunto proporcionan un valor mayor al que podrían ofrecer individualmente”. No obstante, el que datos nuevos y diferentes se añadan, significa la probabilidad de redefinir relaciones y crear otra información. Del mismo modo, (Murdick & Ross, 1990), consideran la información completamente indispensable para que una organización obtenga la vitalidad que necesita para sobrevivir. Como también creen necesario comunicarse mediante sistemas de información funcionales donde la información puede tomar la forma de señales del lenguaje, comportamiento y/o símbolos porque a medida que aumenta el tamaño y complejidad de las organizaciones, existe más posibilidad de carecer de una comunicación realmente relevante.

A diferencia de los datos, la información tiene significado (relevancia y propósito). No solo puede formar potencialmente al que la recibe, sino que está organizada para algún propósito.

### **2.1.3. Conocimiento**

Para (Davenport & Prusak, 1998) “el conocimiento es una mezcla de experiencias, información, habilidades y aptitudes que sirven como marco para la incorporación de nuevas experiencias e información que son útiles para llevar a cabo una acción”. El conocimiento es un recurso valioso. La innovación tecnológica actual de las empresas está basada en el conocimiento, lo cual es indispensable para el desarrollo de la organización. De la misma forma, (Alavi & Leidner, 2003) definen el conocimiento como la “información que el individuo posee en su mente, personalizada y subjetiva, relacionada con hechos, procedimientos, conceptos, interpretaciones, ideas, observaciones, juicios y elementos que pueden ser o no útiles, precisos o estructurables”. La información se transforma en conocimiento una vez procesada en la mente del individuo y luego nuevamente en información una vez articulado o comunicado a otras personas mediante textos, formatos electrónicos, comunicaciones orales o escritas, entre otros. En este momento, el receptor puede procesar e interiorizar la información, por lo que vuelve a transformarse en conocimiento.

Según (Laudon & Laudon, 2012), el conocimiento es un conjunto de conceptos, experiencia y perspicacia que proporcionan un marco de trabajo para crear, evaluar y utilizar información. Es el uso de los datos y de la información en interrelación con el potencial de habilidades, competencias, capacidades y motivaciones de las personas. En consecuencia, el conocimiento es esencial para la toma de decisiones de las organizaciones y la adaptación de estas para responder a situaciones nuevas.

#### **2.1.4. Sistema**

Para (Murdick & Ross, 1990), un sistema puede describirse como una serie de elementos unidos de algún modo a fin de lograr metas en común. También, (Ackoff, 1994), define un sistema como un conjunto de dos o más elementos de cualquier clase interrelacionados, por ejemplo, conceptos (como el Sistema numérico), objetos (como un sistema telefónico o el cuerpo humano), o personas (como en un sistema social).

(Radatz, 1997), explica que un sistema es un todo integrado, aunque compuesto de diversas estructuras interactuantes y especializadas con un número de objetivos que pueden variar ampliamente de un sistema a otro. Sin embargo, la definición de (Burch & Starter, 1981) indica que un sistema se puede determinar como un acoplamiento de componentes o subsistemas ideados con la finalidad de conseguir un objetivo.

Según (Blanchard, 1993), un sistema debe tener las siguientes generalidades:

- Debe tener un propósito, ser funcional y capaz de responder a una necesidad identificada; así como de alcanzar el objetivo global de manera efectiva en cuanto a costos.
- Comprende una combinación compleja de recursos en forma de seres humanos, materiales, software, facilidades, datos, dinero, entre otras.
- Estar contenido dentro de algún tipo de jerarquía.
- Puede dividirse en subsistemas y componentes relacionados donde el grado de la división depende de la complejidad y las funciones que son desempeñadas.

#### **2.2. Sistemas de Información**

Un sistema de información es un conjunto de componentes interrelacionados que reúne, procesa, almacena y distribuye información para apoyar la toma de decisiones y el control en una organización (Laudon & Laudon, 2008).

Los Sistemas de Información según (O'Brien, 2001) se componen de cinco recursos importantes: humanos, hardware, software, datos y redes. Dentro de los sistemas de información se ubican los llamados sistemas gerenciales, estos según (Antonorsi, 1999) comprenden todos los procedimientos que hacen funcionar a la organización como tal. Los sistemas son diversos, ellos pueden estar automatizados o no, hasta pueden ser mixtos. Entre los tipos de sistemas gerenciales pueden mencionarse los sistemas de planificación, de presupuesto de gasto, de materiales, de equipos, de selección de personal, entre otros.

Cuatro actividades de un sistema de información producen la información que la organización requiere para la toma de decisiones, para el control de las operaciones, el análisis de los problemas y la creación de nuevos productos y servicios. Estas actividades son las siguientes:

**Recopilación/Entrada de información:** es el proceso mediante el cual el sistema de información toma los datos que requiere para lograr los objetivos. La entrada de dato puede ser tanto manual como automática. Las manuales son las que se le proporcionan de forma directa por medio de un usuario, mientras que las automáticas son por medio de procesos de otros sistemas.

**Almacenamiento de información:** es la actividad más importante, ya que por medio de este proceso el sistema puede recordar información guardada en procesos anteriores. Esta información es guardada en estructuras llamadas archivos.

**Procesamiento de información:** es la capacidad para efectuar cálculos de acuerdo a una serie de operaciones preestablecidas. Estos datos pueden efectuarse con datos que son introducidos recientemente en el sistema o bien con datos que ya están almacenados. Estos sistemas permiten la transformación de datos fuentes en información que puede ser utilizada para la toma de decisiones.

**Salida de información:** es la capacidad que tiene un sistema de poder sacar información procesada. Los dispositivos de salida más comunes son las impresoras, terminales, cintas magnéticas, graficadores, entre otros.

---

Los sistemas de información también requieren de retroalimentación que es el producto regresado a personas indicadas dentro de la organización para ayudarles a evaluar o a corregir la etapa de alimentación. Los Sistemas de Información son elementos orientados al tratamiento y administración de datos e información para usarla, de la forma más adecuada, en una organización en función a sus objetivos.

El papel estratégico de los sistemas de información involucra el uso de la tecnología de la información para el desarrollo de productos, servicios y capacidades que dan a una empresa ventajas estratégicas sobre las fuerzas competitivas que esta enfrenta en el mercado global.

### **2.2.1. Características de los Sistemas de Información**

Algunas de las características que resultan necesarias para cualquier Sistema de Información de acuerdo con (Gil, 1997), son las siguientes:

- Disponibilidad de información cuando sea necesario y por los medios adecuados.
- Suministro de información de manera selectiva.
- Variedad en la forma de presentación de la información.
- Cierta grado de autonomía para la toma de decisiones.
- Tiempo de respuesta adecuado a las necesidades del usuario.
- Exactitud en la información suministrada.
- Generalidad, como las funciones para atender las diferentes necesidades.
- Flexibilidad, capacidad de adaptación.
- Fiabilidad, para que el sistema opere correctamente.
- Seguridad, protección contra pérdidas.

### **2.2.2. Tipos de Sistemas de Información**

En función a las diferentes actividades, áreas y necesidades dentro de una organización existen diferentes tipos de sistemas. Estos sistemas cubren el tipo de información a requerir, el cual varía en cuanto al nivel de agregación y forma de visualización.

**A nivel organizacional:** En la ilustración 10, se muestra una manera de describir los tipos de sistemas que se tienen en una organización. La organización se divide en niveles estratégicos, de administración, de conocimientos y operativos.

**Tipo de Sistemas de Información**



ILUSTRACIÓN 3: TIPOS DE SISTEMA DE INFORMACIÓN. (LAUDON & LAUDON, 2008)

**a) Sistemas de nivel operativo (sistemas de procesamiento de transacciones (TPS)):**

Son sistemas que permiten a las áreas de operaciones de una empresa, registrar todas las transacciones o eventos dentro de la organización de una forma automatizada, apoyan a los gerentes operativos al hacer seguimiento de las actividades y transacciones elementales de la institución y permite a los gerentes de operaciones realizar consultas básicas, por ejemplo, la cantidad de un producto en inventario. Como también, permite el seguimiento de las operaciones que se ejecutan en un momento determinado. Suelen ser sistemas diseñados para soportar la carga de muchas transacciones por lo general sencillas y habituales.

**b) Sistemas de nivel conocimiento (Sistemas de gestión del Conocimiento (KWS)):**

Se apoyan los trabajadores del conocimiento y los de la información de una institución. Promueven la creación y administración de conocimiento en el área operativa. Por lo general este nivel está conformado por personal profesional (ingenieros, científicos,

programadores) que generan o diseminan información creando documentación, productos y sistemas. La finalidad de estos sistemas es ayudar a la organización a integrar nuevos conocimientos y para que la organización controle el flujo de la documentación.

**c) Sistemas de nivel administrativo (Sistemas de Soporte de Decisiones (DSS) y Sistemas de Información de Gestión (MIS)):** Se diseñan para las actividades de seguimiento y control, toma de decisiones y las actividades administrativas de los gerentes de nivel medio. Los sistemas de información para el área administrativa permiten supervisar, coordinar, administrar y ayudar en la toma de decisiones, se encargan de mostrar y resumir la información proveniente de las transacciones (nivel operativo) para ser visualizada por la gerencia. Por lo general, se generan informes periódicos y se aplica agregación en los datos para mostrar de forma resumida el desempeño de la organización. Los sistemas de apoyo a la toma de decisiones ayudan a la gerencia intermedia a definir decisiones. Suelen enfocarse en problemas particulares, con cambios rápidos y que no se pueda establecer un patrón de resolución predefinido. Por lo que estos sistemas buscan utilizar data interna y externa de la organización para generar modelos, predicciones, y un conjunto de indicadores que permitan ayudar a la gerencia intermedia a definir tácticas que mejoren el desempeño de la organización.

**d) Sistemas de nivel estratégico (Sistema de Apoyo a Ejecutivos (ESS)):** Ayudan a los niveles directivos a atacar y dirigir las cuestiones estratégicas y las tendencias a largo plazo dentro y en el entorno de la institución. Su interés principal es hacer frente a los cambios que ocurren en el entorno con las capacidades con las que se cuentan. Los sistemas de apoyo a ejecutivos son sistemas enfocados a la alta gerencia de una organización. Suelen ayudar a la toma de decisiones no rutinarias (creación de una nueva fábrica, extensión del mercado a un nuevo país, etc) que suelen tener alto impacto en la organización durante años. Estos sistemas presentan una información aún más resumida que los sistemas de información gerencial y la presentación de dicha información suele darse a partir de elementos gráficos (histogramas, gráfico de torta, mapas) que ayuden a la comprensión del estado de la organización.

**A Nivel de Procesamiento de Datos:****• Procesamiento de Transacciones en Línea (OLTP - *Online Transaction Processing*):**

De acuerdo con (Classora Technologies, 2013), el nombre OLTP viene del acrónimo en inglés de Procesamiento de Transacciones en Línea; es un tipo de sistema diseñado para adquirir aplicaciones orientadas a transacciones. Los sistemas de OLTP se diseñan para responder inmediatamente a las solicitudes de usuario y están basados en bases de datos relacionales orientadas al procesamiento de transacciones.

Una transacción es una secuencia de operaciones realizadas como una sola unidad lógica de trabajo, la cual a su vez genera un proceso atómico (que debe ser válido con un *commit*, o inválido con un *rollback*), y que puede involucrar operaciones de inserción, modificación y borrado de datos. Una unidad lógica de trabajo debe exhibir cuatro propiedades, conocidas como propiedades de atomicidad, coherencia, aislamiento y durabilidad (*ACID - Atomicity, Consistency, Isolation, Durability*), para ser calificada como transacción.

Los sistemas OLTP son la versión tradicional de una base de datos: se diseñan utilizando un modelo entidad-relación, se implementan en los motores típicos de base de datos (Oracle, SQLServer, MySQL, entre otros.) y dan soporte a la mayor parte del software del mercado:

- **Optimizadas para lecturas y escrituras concurrentes:** Gracias a las propiedades ACID, el acceso a los datos está adaptado para tareas frecuentes de lectura y escritura.
- **Organizadas según la capa de aplicación:** Las tablas y los datos se estructuran según el software que los maneja: programa de gestión a medida, ERP, CRM, BPM, entre otros.
- **Adaptadas a cada empresa o departamento:** Dado que en muchas ocasiones se utiliza software no integrado, los formatos de los datos no suelen ser uniformes en los diferentes departamentos. Es común la falta de compatibilidad y la existencia de islas de datos.

- **Consultas realizadas en SQL, modificaciones en DML:** El SQL (*Standard Query Language*) es el lenguaje de consulta universal para leer bases de datos relacionales (cláusula SELECT), mientras que DML (*Data Manipulation Language*) es el estándar para realizar modificaciones (cláusulas INSERT, UPDATE y DELETE).
- **Gestión de datos históricos inexistente:** El historial de cambios suele limitarse a datos actuales o recientes. Salvo sistemas de backup, o que el software tenga una funcionalidad específica para ello, no se suelen manejar valores históricos para cada campo.

Gracias a este diseño los sistemas OLTP infieren un tiempo de procesamiento muy rápido, manteniendo la integridad de los datos en entornos multi-acceso y proporcionando una tasa de eficacia muy elevada en la gestión de transacciones por un segundo. Sin embargo, en las bases de datos OLTP no suele haber datos detallados o resumidos, siendo el esquema utilizado para almacenar bases de datos transaccionales el modelo de entidad, por lo general 3FN (3<sup>ra</sup> Forma Normal).

- **Procesamiento Analítico en Línea (OLAP – *Online Analytical Processing*):** El nombre OLAP, viene del acrónimo en inglés de Procesamiento Analítico en Línea. Es una solución utilizada en el campo de Inteligencia de Negocio cuyo objetivo implica agilizar la lectura de grandes cantidades de datos para llegar a extraer algún tipo de información útil: tendencias de ventas, patrones de comportamiento de los consumidores, elaboración de informes complejos, entre otras informaciones. El ámbito empresarial de un sistema OLAP suele ir ligado a la creación de informes en las diversas áreas de marketing, minería de datos o resúmenes de dirección.

Para representar la información que se encuentra en los diversos informes presentados en el área empresarial, los sistemas OLAP utilizan los cubos OLAP, que no es más que un vector de varias dimensiones que contienen datos resumidos en grandes bases y fuentes de datos (Sistemas OLTP, ficheros de datos, entre otros). Estos cubos se componen de hechos numéricos llamados medidas que se clasifican por dimensiones.

Desde un punto de vista relacional, el cubo puede verse como una tabla de hechos que tiene dos tipos de columnas:

- **Indicadores:** También denominados métricas, hechos o ratios, son los valores numéricos con los que se opera. Por ejemplo: n° de clientes, n° de proveedores, importe de las ventas, n° de ventas, importe de las compras, numero de compras, entre otros.
- **Dimensiones:** Son las características por las que se pueden filtrar y cruzar los indicadores. Por ejemplo: tiempo (fijando un determinado día, mes o año), geografía (fijando un determinado país, región o ciudad), proveedor, cliente, modo de pago, entre otros.

Las características de los sistemas OLAP son:

- **Optimizadas para operaciones de lectura:** Dado que la acción más común es la consulta, estas bases de datos disponen de valores agregados y resultados precalculados que les permiten responder en tiempo récord. Evitar las restricciones ACID les da agilidad.
- **Organizadas según las necesidades analíticas:** Los datos están estructurados según las áreas de negocio, y los formatos de los datos están integrados de manera uniforme en toda la organización. Se busca evitar islas de datos.
- **Asíncronas:** No siempre se actualizan en tiempo real, sino que se suelen alimentar con información procedente de las bases de datos relacionales mediante un proceso de extracción, transformación y carga, entre otros.
- **Gestión de datos históricos a largo plazo:** Una de las exigencias analíticas consiste en realizar estudios de evolución a lo largo del tiempo, esto requiere que estas bases de datos mantengan un histórico a largo plazo, normalmente no inferior a cinco años.

### 2.2.3. Comparación entre los Sistemas OLTP y los Sistemas OLAP

Los sistemas OLAP y OLTP tienen numerosas diferencias entre sí, siendo cada uno orientado a distintos objetivos, como se observa en la ilustración 12. En general, los sistemas OLAP

---

son las aplicaciones que permiten analizar el negocio e interpretar lo que ha ocurrido dentro de éste y tomar las decisiones necesarias. Además, son usados para realizar cualquier tipo de análisis, reportes, modelaje y planificación de los aspectos de negocio de la organización, es decir, permiten dar soporte a decisiones estratégicas del negocio. Por otra parte, los sistemas OLTP son las aplicaciones que ejecutan operaciones del día a día (como ventas, inventario, etc.). Estos sistemas representan una imagen de los asuntos de la organización que se actualizan constantemente, y son usados para procesar las transacciones que ocurren día a día dentro de la organización, en otras palabras, permiten dar soporte a decisiones diarias.

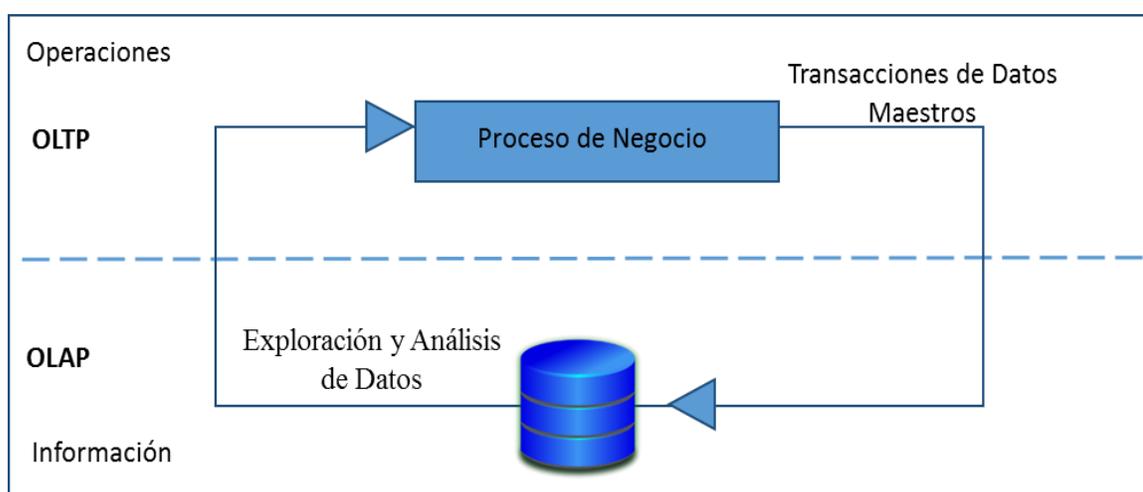


ILUSTRACIÓN 4: COMPARACIÓN ENTRE LOS SISTEMAS OLTP Y LOS SISTEMAS OLAP. (Pinto, 2013)

En la siguiente tabla (Tabla 1.) se puede apreciar de manera más resumida las principales diferencias entre los sistemas OLTP y los Sistemas OLAP.

Características	OLTP	OLAP
<b>Objetivos Principales</b>	Asistir a aplicaciones específicas y garantizar la integridad y consistencia de los datos.	Consolidar los datos ya validados y adecuados a las necesidades para la toma de decisiones.
<b>Fuente de datos</b>	Información operativa de la aplicación. Generalmente estos procesos son la fuente de datos.	Archivos y datos históricos.
<b>Orientación</b>	Orientado a la aplicación, hace cumplir las reglas del negocio.	Orientado al sujeto, se define en base a lo que el analista necesita ver.
<b>Usuarios</b>	Persona común.	Gerentes, ejecutivos, científicos de datos, comercializadores.

Modelo de datos	Bases de datos Entidad – Relación.	Datos multidimensionales.
Esquema	Esquemas normalizados. Muchas tablas y relaciones.	Esquema estrella, copo de nieve y el esquema constelación. Menos tablas no normalizadas.
Acceso y manipulación de datos	Realizan una manipulación de datos registro por registro con grandes cantidades de INSERTS, UPDATES y DELETES. Además, necesitan de rutinas de validación y transacciones a nivel de registro.	Los almacenes de datos tienen una carga y acceso masivo de datos, no se realizan UPDATES o DELETES.
Dimensión Tiempo	Almacena datos actuales. Hay una falta de soporte explícito para reconstruir la historia previa.	Datos de tiempo largo. Se almacenan históricos de los datos.
Velocidad	Rápido. Requiere algunos índices en tablas grandes.	Lenta. Dependiendo de la cantidad de datos. Requiere más índices.
Integración de datos	Los datos se encuentran típicamente no integrados, son calificados como datos primitivos o datos operacionales.	Los datos deben estar integrados. Son conocidos como datos derivados o datos DSS dado que provienen de sistemas transaccionales.

TABLA 1: COMPARACIÓN ENTRE LOS SISTEMAS OLTP Y LOS SISTEMAS OLAP

### 2.3. Base de Datos

(Molina, 2008) indica que una base de datos es una colección de datos relacionados lógicamente que existen por un periodo de tiempo determinado y que es manejada por un sistema manejador de base de datos. Las bases de datos permiten servir como repositorio de datos para el almacenamiento, procesamiento y extracción de datos. En una base de datos, aparte de los datos que son almacenados representado las operaciones de las organizaciones, también hay una serie de datos que describen estos datos, como el tipo de dato, ubicación física, relación con otros datos, etc. A estos datos se le conoce como metadato. Los datos que pueden ser almacenados en una base de datos dependerán del sistema manejador de base de datos, más a nivel conceptual pueden soportar cualquier de datos (audio, texto, imágenes, etc.). Dentro de las bases de datos, destaca el concepto de base de dato relacional, la cual es un tipo de base de datos que se genera a partir de un modelo dimensional.

### 2.3.1. Base de Datos Relacional

Es un modelo ampliamente utilizado e implementado en el desarrollo de sistemas transaccionales. El modelo relacional se sustenta en el concepto matemático de relación, donde la data es representada en forma de “tablas”, donde cada instancia de la tabla se interpreta como un conjunto de valores ordenados. A su vez cada instancia de una tabla puede estar interconectada a otra instancia en cualquier otra tabla (Ilustración 3).

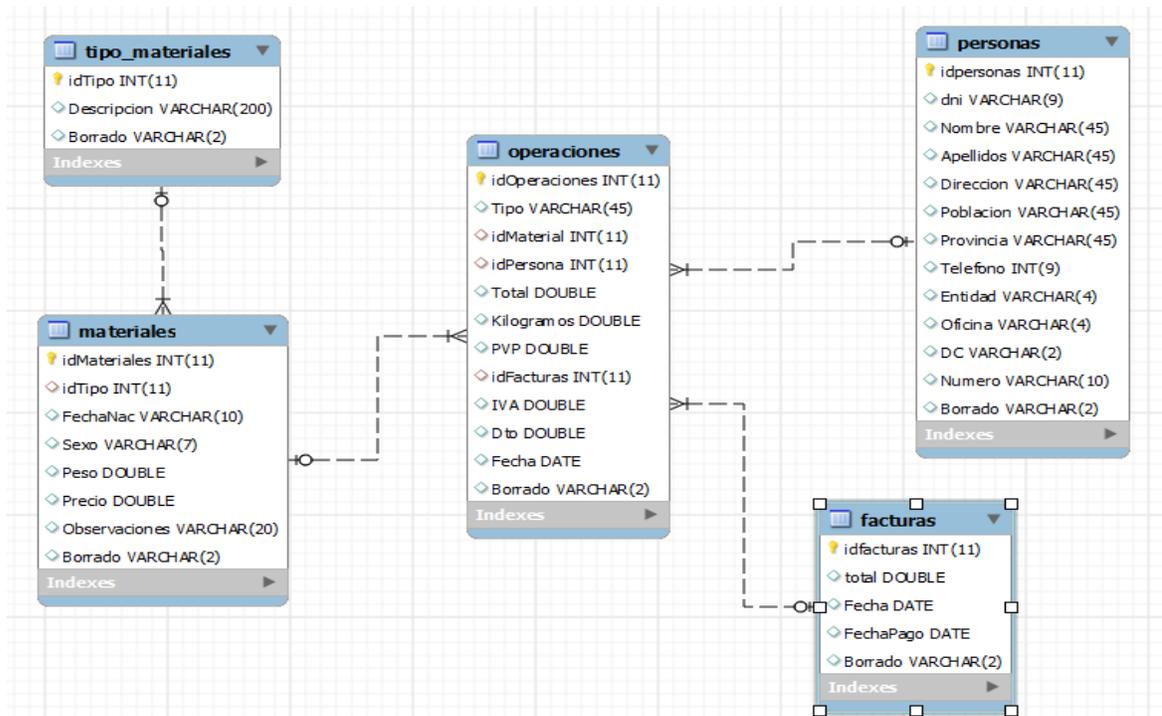


ILUSTRACIÓN 5: EJEMPLO DE UN DIAGRAMA DE BASE DE DATOS RELACIONAL.

La normalización, específicamente la tercera forma normal definida por (Codd, 1971) busca generar atomicidad en los datos y por tanto eliminar la redundancia en el esquema. Esta forma normal se obtiene al evitar dependencias entre los datos de una misma tabla, teniendo un solo atributo clave en la tabla y eliminando dependencias transitivas entre atributos.

Cuando se implementa un esquema relacional para sustentar las operaciones organizacionales, la tercera forma normal suele ser implementada para buscar la integridad en los datos y por tanto asegurar la veracidad de los mismos.

Ciertas características aplican en la base de datos relacional:

- Está compuesta por varias relaciones y varias tablas.
- No existen dos tablas con el mismo nombre.
- Cada tabla está compuesta por un conjunto de filas y columnas.
- La relación entre una tabla padre y una tabla hijo se lleva a cabo por medio de una clave primaria y una clave foránea.
- La clave primaria es la clave principal de un registro dentro de una tabla cumpliendo con la integridad de datos.

#### **2.4. Almacén de Datos (Datawarehouse).**

La definición de Almacenes de Datos – conocido en el medio como *Data Warehouse* - (AD), más extendida, es la propuesta por Bill Inmon: “Un Almacén de Datos es una colección de datos orientados al dominio, integrados, no volátiles y variables en el tiempo, organizados para dar apoyo al proceso de toma de decisiones” (Inmon B. , 2002). Esta definición, incluye el objetivo (ayuda a la toma de decisiones) y las principales características (orientados al dominio, integrados, no volátiles y variables en el tiempo).

Según (Kimball, 1996), otro conocido autor de los almacenes de datos; define un Almacén de Datos como: “Una copia de las transformaciones de datos específicamente estructurada para la consulta y el análisis”. También fue Kimball quien determinó que un *Data Warehouse* no es más que: “La unión de todos los *Data Marts* (subconjunto del *Data Warehouse*, usado normalmente para el análisis parcial de los datos asociados a un área en particular del negocio) de una organización, siendo una copia de los datos transaccionales estructurados de una forma especial para el análisis”.

##### **2.4.1. Objetivos de un Almacén de Datos**

Los objetivos principales que tiene un Almacén de Datos son:

- **Lograr que la información de la organización sea accesible:** permitiendo que el contenido del Almacén de Datos sea entendible, manejable, rápido de acceder y navegable, facilitándole al personal obtener esa información para poder lograr un objetivo en específico.
- **Lograr que la información de la organización sea consistente:** permitiendo que la información que se trabaja en la organización tenga el mismo nombre y el mismo significado, lógica, coherencia y solidez entre ellos.
- **Proporcionar información adaptable y elástica:** es decir, que el Almacén de Datos está hecho para cambios continuos en la información, cuando se le hacen consultas o se agregan datos, tanto los datos como la tecnología existente no cambian en absoluto.
- **Proporcionar un seguro que protege los valores de la información:** se le brinda al usuario el derecho de poder manipular los datos o hacer uso de ellos dependiendo de una permisología que se le asigna, logrando brindar una seguridad en la información ya que no todos los usuarios pueden hacer y deshacer información.
- **Fundamentar la toma de decisiones:** el Almacén de Datos tiene los datos correctos, como el volumen de datos necesarios para apoyar la toma de decisiones de una organización.

#### 2.4.2. Características de un Almacén de Datos

Según (Inmon B. , 1996), un almacén de datos, debe cumplir con cuatro características:

- **Integrados:** En las organizaciones generalmente se utilizan distintos sistemas operacionales, cada uno de ellos con su base de datos diseñada para soportar los procesos específicos de un área del negocio. En la construcción de un Almacén de Datos, todos

estos sistemas operacionales deben ser integrados en una única base de datos, este proceso de integración implica costosas tareas de limpieza, transformación y derivación de datos.

- **Orientado a Temas:** Un Almacén de Datos se clasifica por áreas temáticas en base a los aspectos que son de interés para la empresa. Es por esto que el diseño se orienta a realizar consultas eficientes en relación a la información de las actividades básicas de la organización, como por ejemplo ventas, compras o reclamos. Estos temas pueden verse como un conjunto de indicadores o medidas que son de interés para la empresa.
- **Variables en el Tiempo:** Generalmente, en los sistemas operacionales los datos no tienen una dimensión temporal explícita ya que en estos sistemas se almacena información actual (último año, último periodo, entre otros). En un Almacén de Datos, sin embargo, el tiempo adquiere un valor importante. Cuando se analizan los datos para descubrir tendencias, es importante conocer “la variación de los datos en el tiempo”, por este motivo, en un Almacén de Datos los datos deben estar ligados siempre a un instante específico de tiempo o a un intervalo.
- **No Volátil:** En los sistemas operacionales, los datos son almacenados por cortos periodos de tiempo, por ejemplo, un año, ya que son de interés para la empresa durante ese periodo. En el análisis de datos, sin embargo, es frecuente buscar tendencias del negocio haciendo comparaciones entre los datos de diferentes periodos de tiempo. El Almacén de Datos, existe para ser consultado y no para ser modificado. La información es por tanto permanente y la actualización del Almacén de Datos consiste exclusivamente en la incorporación de datos correspondientes al último periodo de tiempo.

### 2.4.3. Bodega de Datos (Data Mart)

Según (Inmon B. , 2000), una Bodega de Datos (Data Mart) es una estructura departamental de suministro de datos desde el Almacén de Datos donde los datos están desnormalizados, basado en la información que necesite un departamento.

En un Almacén de Datos se almacena toda la información de la empresa u organización, y en una Bodega de Datos se almacena un subconjunto de datos, con el propósito de ayudar a un área específica dentro de un negocio para tomar mejores decisiones.

Las Bodegas de Datos, suelen ser cargadas con datos de sólo unas pocas fuentes. Las fuentes pueden ser sistemas internos de funcionamiento, almacenes de datos centrales, o de datos externos.

#### **2.4.3.1. Características de las Bodegas de Datos**

- Trata con una cantidad de usuarios mucho más limitada.
- Trabaja con un área específica de una empresa u organización.
- Tiene un propósito u objetivo específico.
- Tiene una función de apoyo con respecto al Almacén de Datos.
- Tiene una mayor rapidez de consulta.
- Facilidad para la historización de los datos.
- Validación directa de los datos.

#### **2.4.3.2. Ventajas de una Bodega de Datos**

- Fácil acceso a los datos que se necesitan frecuentemente.
- Crea vista colectiva para grupo de usuarios.
- Mejora el tiempo de respuesta del usuario final.
- Costo inferior al de la aplicación de un completo Almacén de Datos.
- Los usuarios potenciales son claramente más identificables que en un Almacén de Datos completo.

#### **2.4.4. Diferencia entre una Bodega de Datos y un Almacén de Datos**

Un Almacén de Datos, a diferencia de una Bodega de Datos, se ocupa de varios temas y es típicamente controlado e implementado por una unidad central organizacional, se entiende

como central lo que es conocido como el Almacén de Datos de la empresa. Por lo general un Almacén de Datos reúne información de múltiples sistemas.

Ninguna de las definiciones básicas de una Bodega de Datos limita el tamaño como la complejidad del apoyo y la toma de decisiones que puede contener. Sin embargo, las Bodegas de Datos son más pequeñas y mucho menos complejas que los almacenes de datos, por lo que en general son más fáciles de construir y mantener.

#### **2.4.5. Modelo Dimensional**

Según (Kimball, 1996), el modelo dimensional de un Almacén de Datos, es una técnica de diseño lógico la cual tiene por fin presentar los datos de manera estándar e intuitiva permitiendo a su vez un acceso de alto rendimiento a esos datos; usada para la construcción de los almacenes de datos.

El modelo dimensional, que es el producto generado a partir del proceso de modelado dimensional, es clave para la generación de sistemas que ayuden a la toma de decisiones, porque permiten facilitar el entendimiento de la data y mejorar las prestaciones del sistema. Esencialmente, los modelos relacionales que son los utilizados en los sistemas transaccionales, si bien permiten el registro y carga de una gran cantidad de transacciones, carecen de la capacidad de realizar consultas o una navegación compleja sobre la data. Esto ocurre porque el modelo dimensional se estructura en función de valores o requerimientos claves para la organización y partir de esta necesidad generar un modelo que permita el análisis para la toma de decisiones.

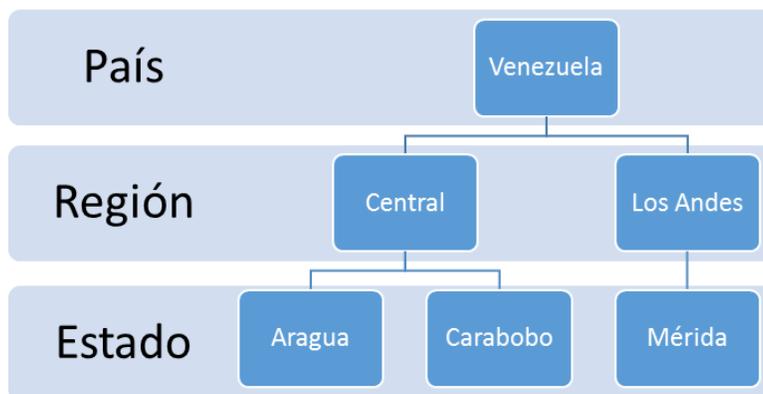
Para realizar un análisis de un requerimiento o área de la organización, el modelo dimensional consta esencialmente de dos elementos. Por un lado, valores o medidas que sirven como indicadores (hecho) y el conjunto de parámetros que definen la existencia de ese hecho en particular (dimensión). Además, existen otros elementos que intervienen en el modelado dimensional.

**Hecho:** Es una representación de una medida dentro de la organización. Los hechos dentro del modelo dimensional permiten evaluar el desempeño de la organización en un proceso específico. Por lo general los hechos suelen ser numéricos y se puede aplicar sobre ellos alguna operación de agregación (suma, promedio, etc.), (Kimball & Ross, 2013).

**Dimensión:** Según (Kimball & Ross, 2013) “las dimensiones provén el punto de entrada a la data”. Una dimensión es una característica que permite analizar a los hechos. En este sentido, la dimensión le va a agregar valor al hecho, situándolo en un contexto que lo define, por lo que, un hecho, suele estar relacionado con una o varias dimensiones. Ya que los datos dentro de la dimensión buscan describir un hecho. Estos son tipo cualitativo. Siendo una característica resaltante el hecho de que dichos atributos pueden organizarse para dar distintos niveles de detalle a la hora de realizar un análisis particular. A esta organización se le llama jerarquía.

**Jerarquía:** Una jerarquía, como se observa en la Ilustración 4, define la forma en como un conjunto de atributos de una dimensión pueden ser agrupados para realizar análisis.

ILUSTRACIÓN 6: EJEMPLO DE JERARQUÍA GEOGRÁFICA.



**Granularidad:** (Inmon B. , 2002) define la granularidad como “el nivel de detalle o sumario de las unidades de data [...] Entre más detalle existe, menor es el nivel de granularidad”. La granularidad, como se muestra en la Ilustración 5, representa el nivel de detalle al que se desea almacenar la información sobre el Almacén de Datos que se esté analizando.

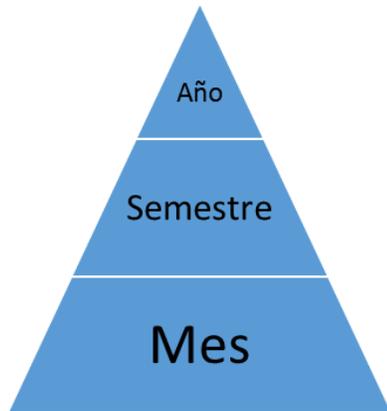


ILUSTRACIÓN 7: EJEMPLO DE GRANULARIDAD POR TIEMPO.

**Agregación:** “La agregación es una medida, es decir un dato contable” (William, 2007). Por ejemplo, si se quiere tener el total del salario de una empresa, lo que se hace es hacerle un proceso de agregación, definiendo como operación la suma de los sueldos de los empleados, logrando que cada vez que entre más información en el atributo sueldo, se va a ir mostrando sumariado.

#### 2.4.5.1. Fases del modelado dimensional.

Todo diseño de un Almacén de Datos debe comenzar con un modelo dimensional, debido a que, con este se identifican cuáles son los hechos que se desean medir y desde cuáles perspectivas se desean agrupar los mismos. Para desarrollar un modelo dimensional, (Kimball & Ross, 2013) sugiere cuatro fases para su construcción. Dichas fases son presentadas a continuación.

- a) **Seleccionar el proceso de negocio:** Dentro de una organización, existen múltiples tareas o procesos que se interrelacionan entre sí. El primer paso en el modelado dimensional consiste en identificar cual proceso de negocio se quiere modelar. El proceso define la granularidad, hechos y dimensiones. Un aspecto importante en la selección del proceso de negocio, es establecer los requerimientos analíticos. En este sentido el levantamiento de requerimientos y por consecuencia la creación del modelo dimensional, pueden darse a partir de distintos escenarios, desde los cuales es necesario la construcción de un modelo dimensional como base para la creación

de herramientas que permitan realizar análisis y ayudar en la toma de decisiones. Existen tres escenarios posibles desde los cuales se establecen los requerimientos de un modelo dimensional:

- **Listado de requerimientos:** se parte la construcción del modelo dimensional a partir de una lista de requerimientos de información suministradas por el dueño del proceso. En este sentido, el modelo de datos va a estar ajustado a extraer, calcular y cargar los datos requeridos para cumplir con los requerimientos preestablecidos.
- **Análisis del proceso:** en este escenario el modelo dimensional busca a partir del entendimiento del proceso de negocio establecer y resolver las necesidades analíticas. Esto requiere que el diseñador comprenda el negocio, los factores que influyen en el mismo, etc. Por lo que el mismo requiere comprender los sistemas transaccionales que operan en el negocio, así como utilizar cualquier otro material que le suministra información del proceso.
- **Modelado de datos transaccional:** en función de los modelos transaccionales físicos que operan en la organización se realiza el modelo dimensional. No existen requerimientos a priori, por lo que se busca es crear un modelo dimensional capaz de satisfacer cualquier requerimiento posible del negocio.

En cualquiera de los tres escenarios, los modelos transaccionales e incluso el proceso de negocio pueden ser modificados para poder satisfacer algún requerimiento.

- b) **Definir la granularidad:** Los sistemas operacionales al registrar todas las transacciones de una organización, van a tener el nivel más bajo de granularidad, por ejemplo, una venta, el registro de un cliente, etc. En el modelado dimensional, es importante el término de granularidad, porque esto va a definir los indicadores o métricas dentro de la tabla de hechos. Un bajo nivel de granularidad permite

mayor flexibilidad en la creación de reportes y manejo de la data para generar nuevas agregaciones y análisis.

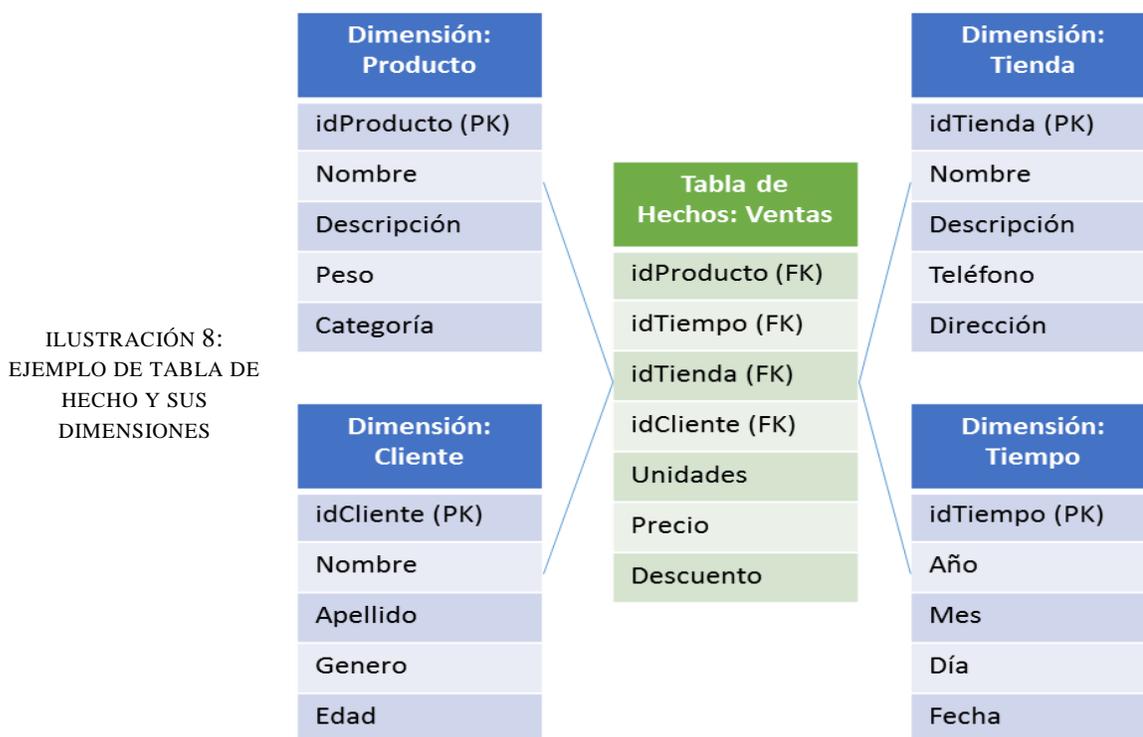
- c) **Identificar dimensiones:** Como se describió antes (2.3.5), las dimensiones caracterizan los hechos dentro del modelo dimensional. Por lo que identificar los elementos que le dan sentido a la ocurrencia o existencia de un dato que sirve para la toma de decisiones es fundamental para que el modelo dimensional tenga la capacidad de manipular los hechos en función de distintas variables
  
- d) **Identificar hechos:** Como último paso, a partir de la granularidad y las dimensiones, se establecen los hechos, que no son más que las medidas con las cuales los analistas se apoyaran para tomar decisiones.

#### 2.4.5.2. Esquemas

Antes de conocer los diferentes esquemas que puede tener un modelo dimensional, es importante mencionar dos conceptos:

**Tabla de Hechos:** “En el modelo dimensional, la *tabla de hechos* es la tabla principal con las mediciones de rendimiento numéricas que son caracterizadas por una clave compuesta, donde cada elemento de la misma es una clave foránea que corresponde a una tabla de dimensiones” (Kimball & Caserta, 2008). Se entiende como tabla de hechos, la tabla principal en un modelo numérico, donde las mediciones de una organización o empresa se almacenan en ellas. Cada medida se toma de la intersección de las dimensiones que la definen, estas dimensiones son las que van a estar alrededor de ella, es decir, relacionada directamente con ella. Una fila de la tabla de hechos corresponde a la medida. Todas las medidas deben estar basadas en un mismo nivel de granularidad (grano). Los hechos más útiles en una tabla de hechos son tanto los numéricos como los aditivos. La tabla de hechos representa la relación de muchos a muchos, entre las distintas dimensiones que se relacionan con ella.

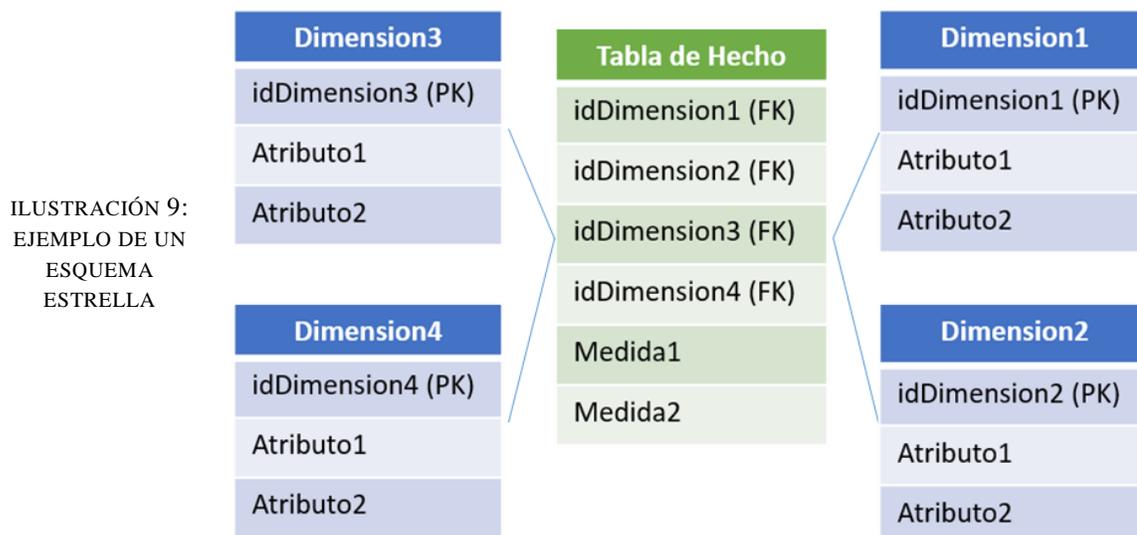
**Tabla Dimensión:** “La tabla de dimensión es una tabla en el modelo dimensional con una clave primaria simple y columnas de atributos descriptivos”. (Kimball & Caserta, 2008). Estas tablas de dimensiones acompañan a la tabla de hecho para definir los parámetros de los que dependen los hechos registrados en ella, para eso se tiene que entender que una tabla de dimensión son ciertos elementos que contiene atributos (campos), que se utilizan para agrupar o restringir los datos que se almacenan en una tabla de hechos que son consultadas en un entorno de Bodega de Datos o Almacén de Datos. La ilustración 6, muestra un ejemplo de una tabla de Hechos con sus dimensiones.



Un modelo dimensional consta básicamente de una o varias tablas de hechos asociadas a varias dimensiones formando con ello un esquema que puede tener la siguiente estructura:

**Esquema Estrella:** “El esquema de estrella es la representación genérica de un modelo dimensional en una base de datos relacional en donde una tabla de hechos con una clave compuesta es unida a un número de tablas de dimensiones, cada una con una clave primaria simple” (Kimball & Caserta, 2008).

Es el esquema de modelado más simple e implementa un diseño lógico relacional de base de datos, como se observa en la Ilustración 7, que resulta en que las tablas de hechos representan la Tercera Forma Normal (3FN) y las dimensiones representan la Segunda Forma Normal (2FN).



El motivo por dejar de mantener las tablas en el modelo relacional y permitir el almacenamiento de información redundante, es optimizar el tiempo de respuesta de base de datos y dar información a un usuario en menos tiempo posible.

- **Esquema Copo de Nieve:** “El esquema de copo de nieve es una dimensión normalizada donde una tabla de dimensión simple es descompuesta en una estructura de árbol con potencialmente muchos niveles” (Kimball & Ross, 2013). Es una variedad más compleja del esquema estrella. El refinamiento está orientado a facilitar el mantenimiento de dimensiones. Lo que distingue a la arquitectura en copo de nieve de la estrella, como se muestra en la ilustración 8, es que las tablas de dimensiones en este modelo representan relaciones normalizadas y forma parte de un modelo relacional de base de datos. Con varios usos del esquema de copo de nieve, el más común es cuando las tablas dimensiones están muy grandes o complejos y es muy difícil representar los datos en esquema estrella.

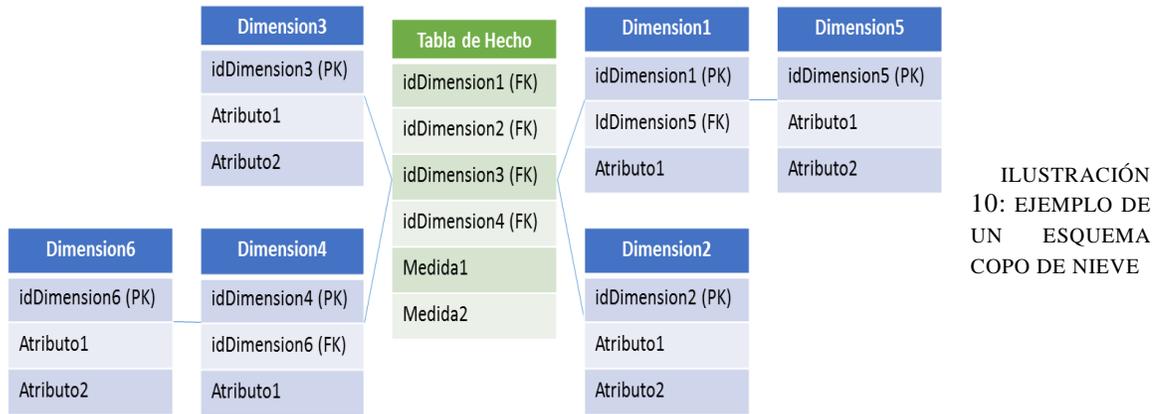


ILUSTRACIÓN 10: EJEMPLO DE UN ESQUEMA COPO DE NIEVE

- Esquema de Constelación de Hechos:** Es una combinación de un esquema de estrella y un esquema de copo de nieve. Los esquemas de constelación son esquemas de copo de nieve en los que sólo algunas de las tablas de dimensiones se han desnormalizado, como se ve reflejado en la Ilustración 9. El objetivo de los esquemas de constelación es aprovechar las ventajas de los esquemas de estrella y de copo de nieve. Las jerarquías de los esquemas de estrella están desnormalizadas, mientras que las jerarquías de los esquemas de copo de nieve están normalizadas.

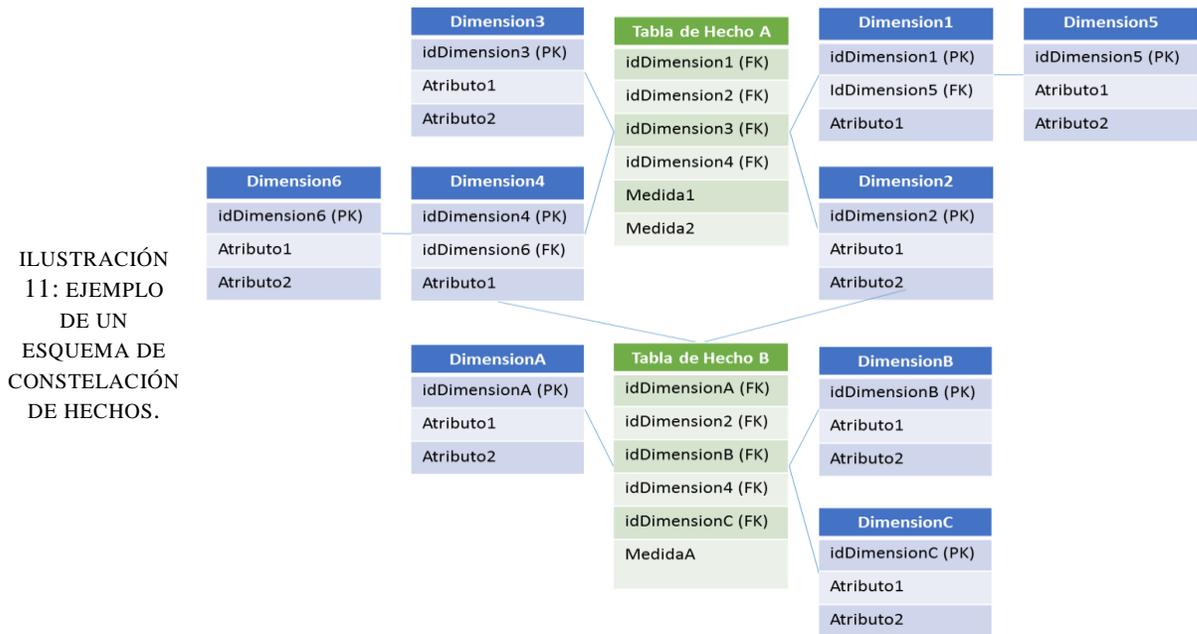


ILUSTRACIÓN 11: EJEMPLO DE UN ESQUEMA DE CONSTELACIÓN DE HECHOS.

### 2.4.6. Ventajas y Desventajas de los Almacenes de Datos

#### ▪ Ventajas de los Almacenes de Datos

Algunas de las ventajas que poseen los almacenes de datos son las siguientes:

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

#### ▪ Desventajas de los Almacenes de Datos

Algunas de las desventajas que presentan los almacenes de datos son las siguientes:

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

## **2.5. Inteligencia de Negocio (BI)**

Howard Dresner (1989), popularizó el acrónimo de “BI” (“Business Intelligence” o Inteligencia de Negocios), para indicar el conjunto de conceptos y métodos para mejorar la toma de decisiones en los negocios, utilizando sistemas de apoyo basados en hechos.

Espiñera, Sheldon y Asociados (2008), en la actualidad el concepto de Inteligencia de Negocios incluye una amplia categoría de metodologías aplicaciones y tecnologías que permiten reunir, acceder, transformar y analizar los datos, transacciones e información no estructurada (interna y externa), con el propósito de ayudar a los usuarios de una compañía a tomar mejores decisiones de negocio. Esto se puede lograr, bien sea mediante la explotación directa (consulta, reportes, etc.) o haciendo uso del análisis y conversión en conocimiento.

Según The Data Warehouse Institute (2006), lo define como la combinación de tecnología, herramientas y procesos que me permiten transformar los datos almacenados en información, esta información en conocimiento y este conocimiento dirigido a un plan o una estrategia comercial. La inteligencia de negocios debe ser parte de la estrategia empresarial, esta le permite optimizar la utilización de recursos, monitorear el cumplimiento de los objetivos de la empresa y la capacidad de tomar buenas decisiones para así obtener mejores resultados.

La inteligencia de Negocio permite conocer el estado de los procesos de la organización y resolver incógnitas de gran importancia como lo que está sucediendo en un determinado momento, él porque está ocurriendo, que debería hacer el equipo para mejorar la situación, y que camino debe seguir el negocio.

### **2.5.1. Característica de una Solución de Inteligencia de Negocio**

Toda Solución de Inteligencia de Negocio, según (Cano J. L., 2007) debe cumplir con las siguientes características:

- **Apoyo a la toma de decisiones:** Básicamente consiste en organizar y presentar los datos relevantes para que sirvan como soporte a la hora de tomar decisiones. Esto implica la utilización de tecnologías, técnicas de análisis y todo lo que sea necesario con el fin de obtener solamente aquella información relevante y útil.
- **Visión unificada de los datos:** Todos los datos deben estar localizados en un único repositorio de datos; sin importar el tipo de datos o la fuente de donde provengan, para así lograr dar percepción de que los datos están centralizados.
- **Creación personalizada de informes y consultas:** Permite el desarrollo de consultas y reportes a la medida sobre información contenida en los almacenes de datos.
- **Presentación de información analítica por medio de vistas de gráficas interactivas:** A través de tableros de control estratégicos se facilita la visualización de los indicadores de negocio.
- **Capacidad de procesamiento de grandes volúmenes de datos:** Las soluciones de BI permiten realizar consultas comparando los datos actuales con los históricos.

### 2.5.2. Funciones de una Solución de Inteligencia de Negocio

- Permiten recolectar, estandarizar y consolidar toda la información de la organización, mediante un almacén de datos, permitiendo así su explotación sin esfuerzo.
- Facilita la extracción de información de los datos y el conocimiento de la información, con la utilización del software adecuado.
- Permiten el perfeccionamiento de 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

La arquitectura de una solución de una solución de Inteligencia de negocios, como se muestra en la Ilustración 10, parte de las fuentes de datos usadas por las aplicaciones de negocio en

---

una organización, sobre los que suele ser necesario aplicar una transformación estructural para optimizar su proceso analítico. Para ello se realiza una fase de extracción, transformación y carga (ETC o ETL) de datos. Esta etapa suele apoyarse (opcional) en un área intermedia de almacenamiento que actúa como un área de paso en donde son descargados los datos, crudos y sin transformaciones, esto se hace con el propósito de ocupar el mínimo tiempo posible los servidores de origen que entregan los datos. En esta área se efectúan las tareas de transformación de los datos para que la información resultante, ya unificada, depurada y consolidada, sea cargada en un almacén de datos, que se caracterizan por poseer la estructura óptima para el análisis de los datos de la organización.

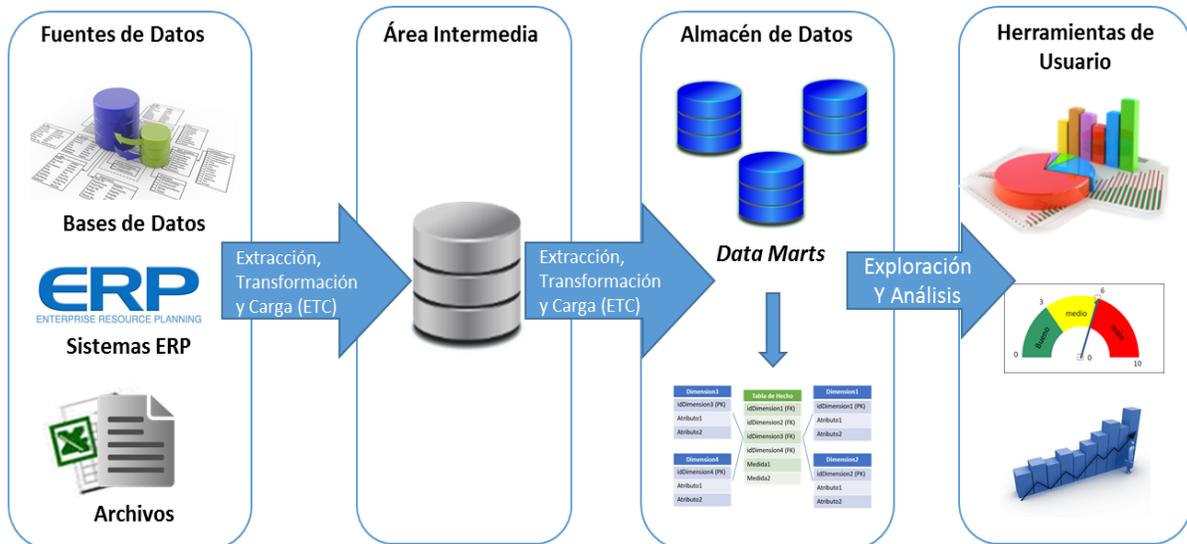


ILUSTRACIÓN 12: ARQUITECTURA DE UNA SOLUCIÓN DE INTELIGENCIA DE NEGOCIOS

### 2.5.3.1. Fuente de Datos

Básicamente son los datos generados por los sistemas origen de una organización, los cuales suelen ser operacionales o transaccionales (bases de datos, ERP (Enterprise Resource Planning), CRM (Customer Relationship Management), ficheros de texto, hojas de cálculo, entre otros).

### 2.5.3.2. Procesos de Extracción, Transformación y Carga (ETC o ETL)

Comúnmente estos procesos son conocidos como ETL por sus siglas en inglés (*Extract, Transformation and Load*) y son los procesos en los cuales se invierte más tiempo para la construcción de un almacén de datos, ya que pasan por una serie de pasos como: filtrado, limpieza, transformación y redefinición de datos. Los procesos de ETC o ETL se dividen en tres pasos principalmente:

- **Extracción:** Consiste en obtener los datos del sistema origen, en algunos casos pueden extraerse los datos completos o solo una parte de ellos, esto de acuerdo a los requerimientos de la solución planteada. En este subproceso también se realizan y aplican filtros a los datos con el fin de eliminar datos redundantes; todo esto se puede realizar mediante herramientas especializadas o incluso por medio de lenguajes de programación. Un requerimiento de importancia en este subproceso es que este cause el menor impacto posible en el/los sistemas origen. Esto se debe a que frecuentemente la gran cantidad de datos que se deben extraer puede ralentizar e incluso colapsar el sistema origen. Por esta razón todas las operaciones de extracción suelen realizarse en horarios o días donde este impacto sea mínimo.
- **Transformación:** Consiste en aplicar transformaciones a los datos fuentes, estas transformaciones pueden ser de tipo como: creación de datos a raíz de otros datos, cambiar el formato del dato, entre otros. Las transformaciones que se realicen se establecen bajo las reglas establecidas en la organización para el negocio.
- **Carga:** Este es el último paso en el proceso de ETC y en ella ya los datos están preparados para ser cargados en el sistema destino. Dependiendo de los requerimientos de la organización, este proceso puede abarcar una amplia variedad de acciones diferentes. En algunas bases de datos se sobrescribe la información antigua con la información nueva. En cambio, en el Almacén de Datos se mantiene un historial de los registros de manera que se pueda hacer una auditoria de los mismos, teniendo almacenada toda la historia de un valor a lo largo del tiempo.

---

Para cargar los datos en un Almacén de Datos existen dos formas que son:

- **Carga Inicial:** Se refiere a la primera carga de datos que se realiza al Almacén de Datos. Por lo general, esta tarea consume un tiempo bastante considerable, ya que se deben insertar registros que han sido generados aproximadamente, y en casos ideales, durante más de cinco años.
- **Carga de Mantenimiento:** Se refiere a la carga que mueve pequeños volúmenes de datos de forma periódica. Su frecuencia está dada en función de la granularidad del almacén y de los requerimientos del negocio. El objetivo de esta tarea es añadir al Almacén de Datos aquellos datos nuevos que se fueron generando desde la última actualización.

### 2.5.3.3. Almacén de Datos

El almacén de datos (definido previamente) son repositorios que se caracterizan por poseer la estructura óptima para el análisis de los datos de la organización diseñados para facilitar la confección de los informes y la realización de análisis.

### 2.5.3.4. Herramientas de Usuario

Es un tipo de software que permite a los usuarios finales la visualización de la información, generada a través de los procesos definidos y ya automatizados, de manera rápida y en tiempo real.

## 2.6. Tecnologías y herramientas para la construcción de una solución de Inteligencia de Negocios

Una herramienta para la construcción de una Solución de Inteligencia de Negocio es un tipo de software diseñado para colaborar en los procesos de negocio de las organizaciones. Según (Martins, 2014), específicamente se trata de herramientas que asisten en el análisis y la presentación de los datos.

Estos conjuntos de herramientas tienen en común las siguientes características:

- **Accesibilidad a la Información:** Los datos son la fuente principal de este concepto. Lo primero que deben garantizar este tipo de herramientas y técnicas será el acceso de los usuarios a los datos, con independencia de la procedencia de éstos.
- **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.

Las organizaciones han descubierto la necesidad de usar sus datos para apoyar la toma de decisiones, realizar una optimización de procesos y realizar reportes operacionales. A raíz de esto, existen diferentes tipos de componentes dentro de las herramientas de Inteligencia de Negocio, tales como:

- **Tableros de Control:** También conocidos como *Business Intelligence Dashboards*, o *Dashboards Ejecutivos*. Son resúmenes visuales de información del negocio, que muestran de una mirada la comprensión del global de las condiciones del negocio mediante métricas e Indicadores Clave de Desempeño (KPIs - *Key Performance Indicators*). Esta es una Herramienta de Inteligencia de Negocios muy popular desde hace unos pocos años.
- **OLAP:** Es la capacidad de algunos sistemas de soporte de decisiones gerenciales que permiten examinar de manera interactiva grandes volúmenes de información desde varias perspectivas.
- **Aplicaciones de Reportes:** Genera vistas de datos agregadas para mantener a la gerencia informada sobre el estado de su negocio.
- **Minería de Datos:** Extracción de información de las bases de datos acerca del consumidor, mediante la utilización de aplicaciones que pueden aislar e identificar patrones o tendencias del consumidor en un alto volumen de datos.

En el mercado se tiene una cierta cantidad de herramientas BI (Business intelligence/ Inteligencia de Negocios), solo que se hace difícil o compleja la elección de ellos (Oracle BI, Pentaho BI, Report Portal BI Solution, Business Objects, IBM Cognos, entre otros) ya que tienen una gran gama de funcionalidades que brinda cada herramienta.

Para la elección de la herramienta a utilizar, hay que revisar los objetivos a lograr o las necesidades de la empresa u organización. Independientemente de la herramienta que se seleccione, se tiene que tener un repositorio de datos, visualización de tableros de control y la posibilidad de construir ciertas consultas.

### **2.6.1. Pentaho BI**

De acuerdo con (Pentaho, 2012), es una plataforma de Inteligencia de Negocio desarrollada bajo la filosofía del software libre, la cual está orientada a soluciones, centrada en procesos e incluye todos los principales componentes requeridos para implementar soluciones basadas en proceso, para la gestión y toma de decisiones empresariales.

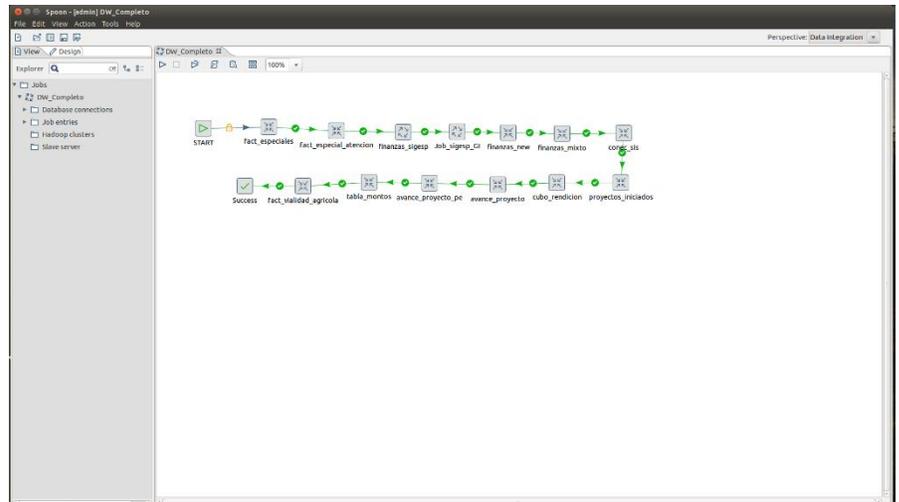
Es una plataforma compuesta por diferentes programas que satisfacen los requerimientos fundamentales para una Solución de Inteligencia de Negocio. Ofrece soluciones para la gestión y análisis de la información, incluyendo el análisis multidimensional OLAP, presentación de informes, minería de datos, creación de tableros de control para el usuario, entre otros.

La plataforma de Pentaho consta de dos (2) presentaciones:

- **Pentaho Enterprise Edition (EE):** Es una plataforma que tiene capacidades de consulta y presentación de informes completos, análisis interactivo de datos y tableros de control, y un portal centralizado que reúne todo esto. Esta edición posee características adicionales que no se encuentran en la edición de la comunidad. Es un software certificado, listo para ambientes de producción, posee módulos exclusivos y facilidades de uso; además incluye soporte técnico oficial.



ILUSTRACIÓN 14: INTERFAZ DEL ENTORNO DE DESARROLLO DE SPOON. (HERRAMIENTA PENTAHO DATA INTEGRATION)



- **Kitchen:** Es el programa que ejecuta los trabajos diseñados en *Spoon*, permitiendo ejecutarlos desde línea de comandos o en modo *batch*.
- **Pan:** Es un programa que ejecuta las transformaciones diseñadas en *Spoon*, permitiendo programarlas y ejecutarlas en modo *batch*.

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.

- **Pentaho Reporting**

Según (Gravitar, 2014), es un conjunto de herramientas de código abierto que permite crear informes relacionales y análisis de una amplia gama de fuentes de datos. Pentaho Reporting es capaz de crear PDF, Excel, HTML, Texto, Rich-Text-Archivo y XML y CSV como salidas de sus datos. Las expresiones Excel-fórmula/OpenFormula proporcionan ayuda para crear informes más dinámicos y de la forma que se desee. También permite programar tareas y ejecutar informes de forma automática con una determinada periodicidad. Así mismo incorpora la librería JPivot, gracias a la cual se pueden ver tablas OLAP a través de un browser y realizar las aplicaciones típicas de análisis OLAP.

En la actualidad existen tres productos con diversos enfoques y dirigidos a diferentes tipos de usuarios, los cuales son:

- **Pentaho Report Designer (Diseñador de Informes):** Es un editor basado en eclipse con capacidad de personalización de informes de acuerdo a las necesidades de negocio para ejecutarlos directamente o publicarlos en la plataforma BI con el fin de que puedan ser ejecutados por los usuarios. Está destinado a desarrolladores. Incluye asistentes para facilitar la configuración de propiedades. Está estructurado de forma que los desarrolladores puedan acceder fácilmente a los datos que serán utilizados para crear el informe que necesiten, en la Ilustración 15 se puede observar la interfaz de esta herramienta.

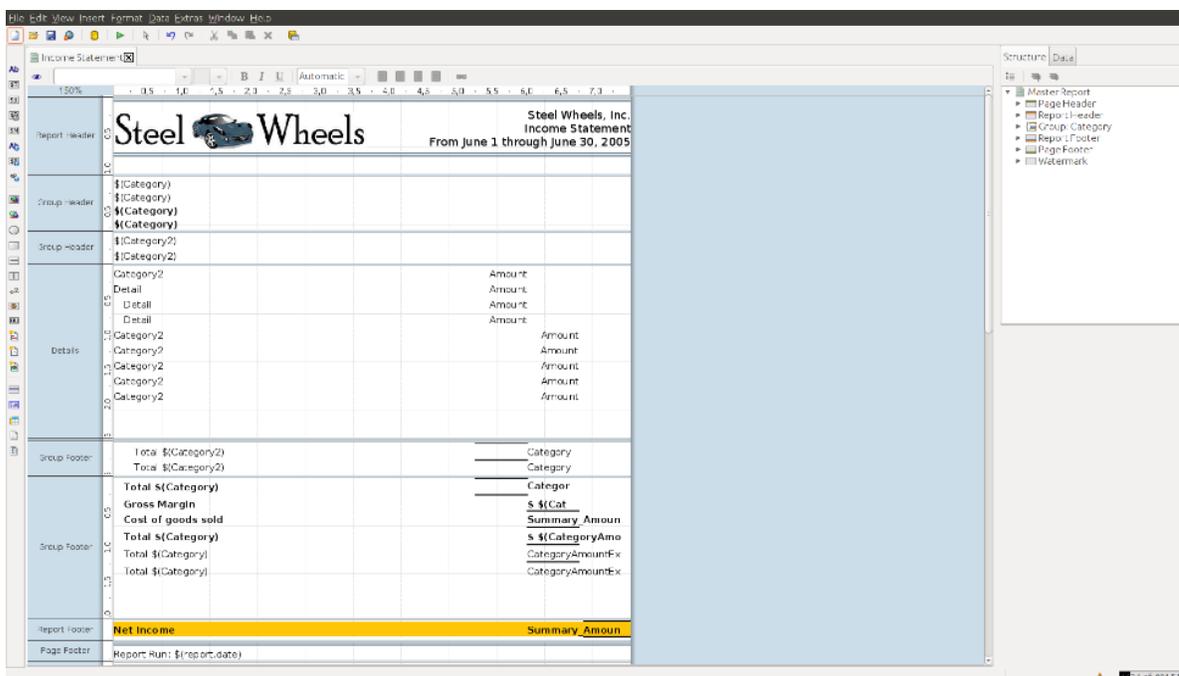


ILUSTRACIÓN 15: INTERFAZ DEL ENTORNO DE DESARROLLO DE PENTAHO REPORT DESING. (HERRAMIENTA PENTAHO REPORT DESING).

- **Pentaho Report Design Wizard (Asistente de Diseño de Informe):** Es una herramienta que se utiliza para crear definiciones de informes mediante el uso de un editor gráfico. Está destinada a usuarios con pocos conocimientos técnicos, ya que permite reducir la brecha entre no tener una herramienta de creación de informes y tener un diseñador de informes en toda su magnitud.
- **Web Ad-Hoc Reporting:** Es 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 y siguiendo un asistente de creación. Además, permite visualizar los distintos tipos de reportes desarrollados con cualquiera de las herramientas de Pentaho reporting. No existe una herramienta para realizar tableros de control Ad-hoc en la versión Community.

▪ **Pentaho Analysis Services (Mondrian)**

De acuerdo con (Sandra & Lemus, 2009), es un servidor que permite realizar procesamiento analítico en línea (OLAP). Soporta el lenguaje de consulta MDX (expresiones multidimensionales) y lenguaje XML para el análisis y especificaciones. Funciona sobre las bases de datos estándares típicas del mercado: Oracle, DB2, SQL-Server, MySQL, PostgreSQL, entre otros. Suministra a los usuarios un sistema para el análisis de información. A través del uso de las tablas, el usuario puede navegar por los datos, ajustando la visión de los mismos, los filtros de visualización, añadiendo o quitando los campos de agregación. Los datos pueden ser representados en una forma de SVG o Flash, dashboards widgets, o también integrados con los sistemas de minería de datos y los portales web (portlets).

Pentaho Analysis se compone de un servidor y de varias herramientas de cliente como:

- **Schema Workbench:** Es la herramienta gráfica que permite la construcción de los esquemas multidimensionales (cubos OLAP) para ser usados por el motor de Mondrian, y además permite publicarlos al servidor BI para que puedan ser utilizados en los análisis por los usuarios de la plataforma, la Ilustración 16 muestra la interfaz de esta herramienta.

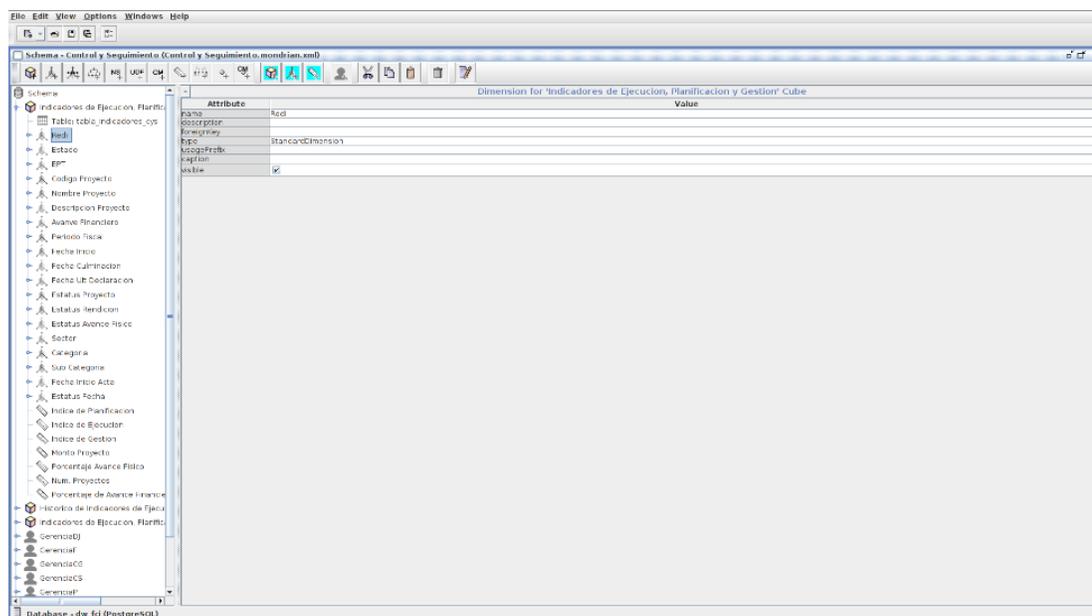


ILUSTRACIÓN 16: INTERFAZ DEL ENTORNO DE DESARROLLO DE SCHEMA WORKBENCH.

- **Desing Studio (PDS):** Es una herramienta de la suite Pentaho que se utiliza para describir y mantener Xactions. Una Xaction o secuencia de acción es un conjunto predefinido de acciones. Las Xactions son ejecutados por Pentaho BI Server, ya sea cuando se inicializa, a través de la orden de un usuario, un evento calendarizado, entre otros. A través de PDS se pueden crear Xactions de forma visual y simple. Además, PDS cuenta con selectores para muchas de sus opciones, lo cual facilita mucho el trabajo a la hora de editar o crear una Xaction.
- **Pentaho Administration Console:** Es una herramienta cliente, donde se pueden crear nuevas vistas de análisis y reportes.
- **Diseñador de Agregación:** Esta herramienta simplifica la creación y despliegue de tablas de agregación que mejoran el rendimiento de los cubos OLAP.

#### ▪ **Pentaho Dashboards**

Es una aplicación que se utiliza para presentar el contenido de BI de alto nivel a los usuarios finales. El contenido del Dashboard, es casi siempre de carácter gráfico; lo que proporciona cierto grado de interactividad que permite al usuario navegar por contenidos más detallados. (Gravitar, 2014).

Todos los componentes del módulo Pentaho Reporting y Pentaho Análisis pueden formar parte de un Dashboard. Los dashboards permiten a las organizaciones:

- **Monitorear** actividades y procesos de negocios críticos utilizando métricas de rendimiento del negocio que disparen alertas cuando se avisten problemas potenciales.
- **Analizar** la causa raíz de los problemas por medio de la explotación de la información relevante y oportuna desde múltiples perspectivas y varios niveles de detalle.
- **Gestionar** personas y procesos para mejorar las decisiones, optimizar el rendimiento y enrumbar a la organización en la dirección correcta.

Los dashboards son tres aplicaciones en una: una aplicación de monitoreo; una aplicación de análisis; y una aplicación de gestión. En estos dashboards los usuarios navegan en tres capas de vistas de información: una vista grafica sumariizada; una vista multidimensional; y una vista detallada u operacional. En la Ilustración 17 se muestra un ejemplo de un Dashboard elaborado con esta herramienta.

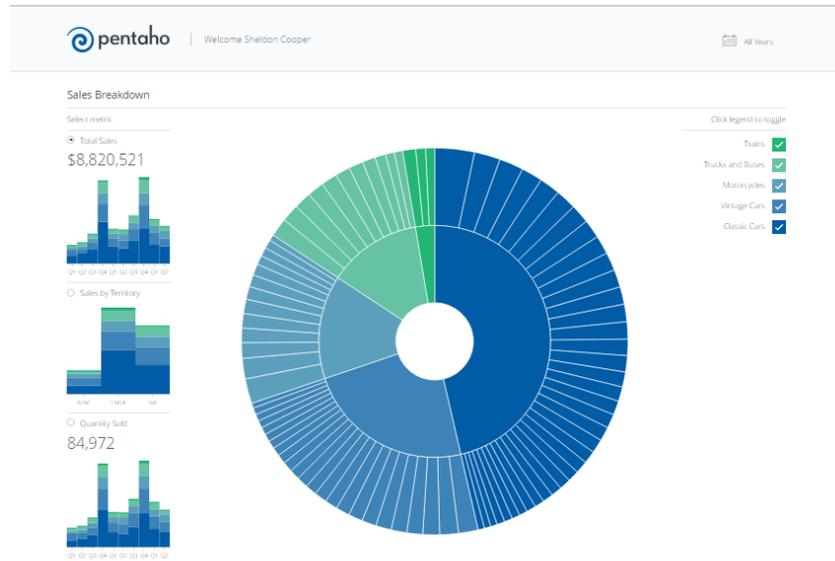


ILUSTRACIÓN 17: EJEMPLO DE UN DASHBOARD ELABORADO CON PENTAHO DASHBOARDS. (PENTAHO A. H., 2005)

### 2.6.2. Oracle Business Intelligence (OBI)

Es una plataforma de Oracle con soluciones de BI y almacenamiento de datos, que permite visualizar los datos y convertirlos en una base para tomar decisiones correctas en la organización (Tecnología de Gestión y Comunicación S.A., 2015). OBI es una plataforma para la creación de Soluciones de BI que ofrece una infraestructura unificada e integrada que incluye un conjunto de productos que abarca:

- Construcción de Almacenes de Datos.
- Consulta y Análisis.
- Creación de Reportes.
- Creación de tableros de control.
- Integración de Microsoft Office, entre otros.

Esta plataforma posee tres ediciones:

- Oracle BI Standard Edition
- Oracle BI Standard Edition ONE
- Oracle BI Enterprise Edition

A continuación, se definen los componentes de OBI Standard Edition One:

▪ **Oracle BI Server**

Es la base de la plataforma, es un servidor de análisis y consulta que se apoya en un modelo lógico de información común. Es capaz de integrar distintas fuentes de datos heterogéneas en una sola visión. Puede generar simultáneamente SQL optimizado frente a múltiples fuentes de datos, sean éstos archivos ASCII, multidimensionales o relacionales. Este servidor permite realizar análisis operacionales sobre los tableros de control o las consultas ad-hoc frente a un solo nivel de presentación que representa todos los activos de datos de la empresa, independientemente de cuáles sean (Oracle Corporation, 2007) En la ilustración 18 se muestra la arquitectura de Oracle Business Intelligence Server.

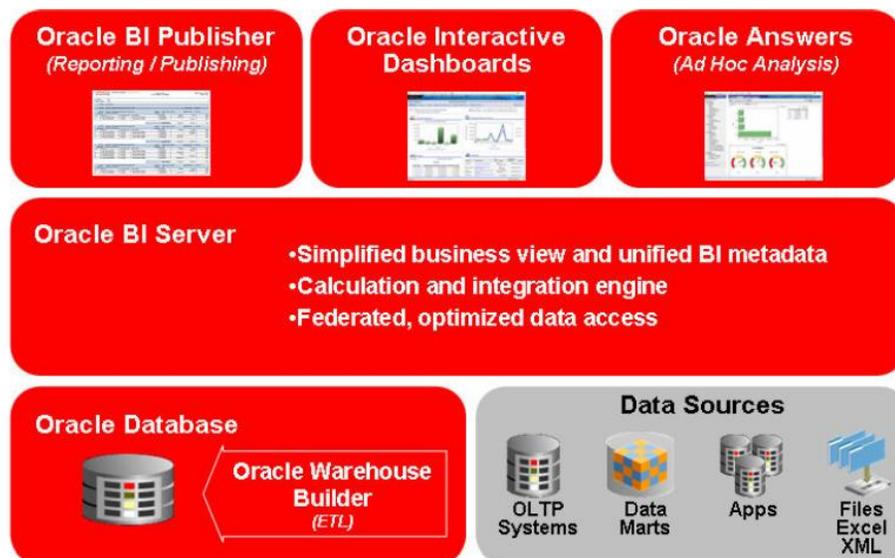


ILUSTRACIÓN 18:  
ARQUITECTURA  
DE ORACLE BI  
SERVER.  
(ORACLE  
CORPORATION,  
S.F.)

- **Oracle Business Intelligence Administrator (Admin Tool)**

De acuerdo con (Oracle Corporation, 2006), Admin Tool es una herramienta que permite crear y administrar los repositorios de datos que se utilizarán dentro de los catálogos en el Oracle Business Intelligence para crear los reportes, gráficos y tableros de control o *DashBoards*. En esta aplicación se definen las tres (3) capas necesarias para completar la creación de los metadatos, las cuales son:

- **Capa Física:** Define las fuentes de datos sobre las cuales Oracle Business Intelligence procesa las consultas. Es manejada para importar los metadatos de las bases de datos usadas como fuente de datos. Al realizar la importación de los datos, todas las propiedades son configuradas automáticamente basándose en la información obtenida durante el proceso de importación. Luego de este proceso, se pueden definir otros atributos sobre los orígenes de datos, como por ejemplo las relaciones que no puedan existir en la fuente de datos, así como la posibilidad de crear objetos manualmente dentro de esta capa.
  - **Capa del Modelo de Negocio y Mapeo:** Define la lógica del negocio. Especifica las asignaciones entre el modelo de negocio y los esquemas de la capa física. En este punto los esquemas físicos se simplifican para formar la base para la visualización de los datos por los usuarios finales. El objetivo principal de esta capa es captar cómo los usuarios piensan en su negocio utilizando su propio vocabulario. En ella, se permite configurar las estructuras lógicas (tablas, jerarquías, dimensiones, cubos, entre otras) a partir de los objetos y elementos importados en la capa física.
  - **Capa de Presentación:** Añade un nivel de abstracción sobre la capa anterior. Proporciona un medio para simplificar aún más o personalizar el modelo de negocio para los usuarios finales. La simplificación de la vista de los datos para los usuarios hace que sea más fácil crear consultas basadas en las necesidades del negocio, ya que se puede exponer sólo los datos que sean significativos para los usuarios, así como organizarlos o cambiarles de nombre según la forma en que
-

ellos piensan acerca de dichos datos. En la Ilustración 19, se puede observar la interfaz gráfica de esta herramienta con las respectivas capas:

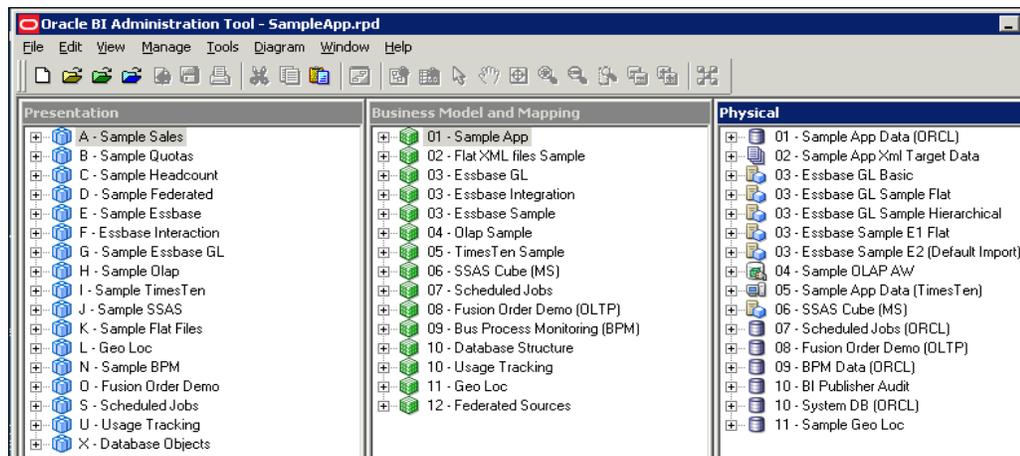
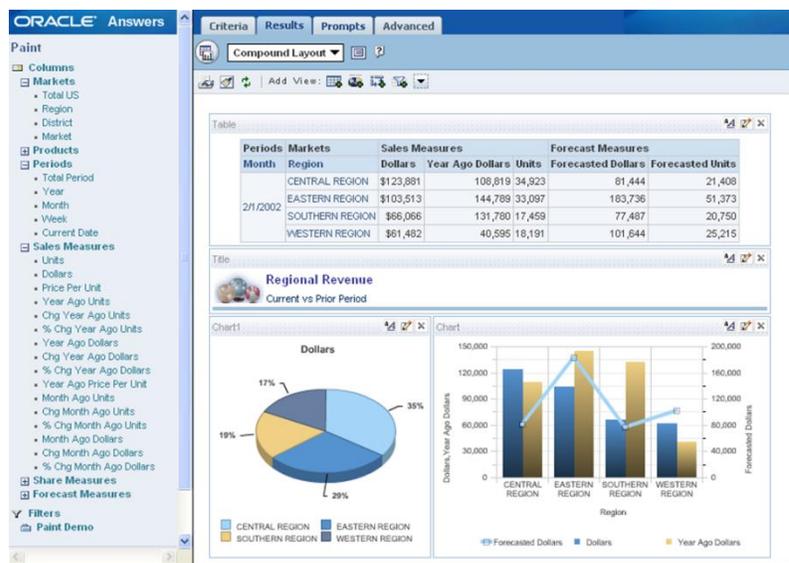


ILUSTRACIÓN 19: INTERFAZ DE ORACLE BI ADMIN TOOL. (Oracle Corporation, s.f.)

▪ **Oracle BI Answers**

Es una herramienta de análisis que procesa los datos desde múltiples orígenes para mostrarlos en un entorno web, y los usuarios pueden consultar y navegar con capacidades de *drill up/down*, compartir los informes, tableros de control, gráficos, entre otros, de una manera ad hoc (al momento) y en todos formatos disponibles. (Oracle Corporation, 2007). Ilustración 20, se muestra un ejemplo de la interfaz gráfica de Oracle Business Intelligence Answer.

ILUSTRACIÓN 20: INTERFAZ DE OBI ANSWERS. (BI-INSIDER, 2011)



- **Oracle BI Interactive Dashboards**

Es una herramienta que brinda una interfaz web personalizable, basada en roles con el fin de proporcionar las tendencias importantes y la inclusión de los análisis. Los tableros de control interactivos permiten a cada usuario controlar el estado del negocio y obtener los conocimientos completos y relevantes que son necesarios para su función, es decir, que les brinda a los usuarios la información de forma dinámica para un entorno de análisis. En la Ilustración 21 se puede observar la interfaz gráfica de esta herramienta.



ILUSTRACIÓN 21: INTERFAZ DE OBI INTERACTIVE DASHBOARDS. (BI-INSIDER, 2011)

- **Oracle BI Publisher**

Es una herramienta que permite crear reportes de las operaciones empresariales y distribución de informes con gran nivel de detalle. Utiliza otras herramientas conocidas de escritorio como Microsoft Office Word, Microsoft Office Excel, PowerPoint, Adobe Acrobat. Se pueden exportar los reportes en distintos formatos tales como, HTML, Microsoft Office Excel, PDF, EFT. En general, es una herramienta de informes operativos que separa los datos de la presentación, por lo tanto, los desarrolladores del departamento de Tecnología de la Información pueden definir las fuentes de datos, y los usuarios pueden especificar el diseño del informe

(Oracle Corporation, 2007). En la Ilustración 22 se muestra la interfaz gráfica de esta herramienta.



ILUSTRACIÓN 22:  
EJEMPLO DE BI  
PUBLISHER. (ORACLE  
CORPORATION, S.F.)

- **Oracle Database Standard Edition One**

Es un Sistema Manejador de Base de Datos (SMBD) objeto-relacional desarrollado por Oracle Corporation. Está basado en la arquitectura cliente/servidor para la gestión de base de datos. Dentro de sus características están el rápido procesamiento de datos, fragmentación de tablas y el uso de vistas materializadas e índices.

Además, es considerado como uno de los SMBD más completos del mercado debido a que provee un alto grado de estabilidad, soporte multiplataforma y buena escalabilidad (Masip, 2002).

- **Oracle Warehouse Builder (OWB)**

Es parte integral de Oracle Database, funciona en todas las versiones (Standard Edition, Standard Edition One, Enterprise Edition). Permite gráficamente la creación del esquema del Almacén de Datos, definición de jerarquías y medidas, mapeo de las fuentes de información, calendarización, ejecución y mantenimiento de las actividades de ETL y herramientas para mejorar la calidad de la información. No solo

está orientada a la definición de Almacenes de Datos o ETL, sino a su administración y mantenimiento (Stackowiak, Rayman, & Greenwald).

Fue concebida para trabajar con SMBD Oracle para la base de datos destino, es decir, donde va a residir el Almacén de Datos, pero puede tener de fuente distintos sistemas manejadores de base de datos accesibles a través de Oracle Database Heterogeneous Services (Oracle Gateway o conexión ODBC) como IBM DB2, Informix, Teradata, SQL Server, entre otros. (Stackiwiak, Rayman, & Greenwald).

Para la construcción del Almacén de Datos, primero se debe crear su estructura lógica, es decir, las dimensiones, jerarquías y cubos definidos en el modelo dimensional previamente diseñado. Luego se deben desplegar (creación física) de las secuencias y las tablas enlazadas a los objetos lógicos.

Los procesos ETL se generan gráficamente mediante el uso de una variedad de componentes que facilitan la extracción de la data, y luego de que se construyan esos procesos, éstos deben ser desplegados para que se genere su código fuente y se almacene en base de datos. En la Ilustración 23, se puede observar la interfaz gráfica de esta herramienta.

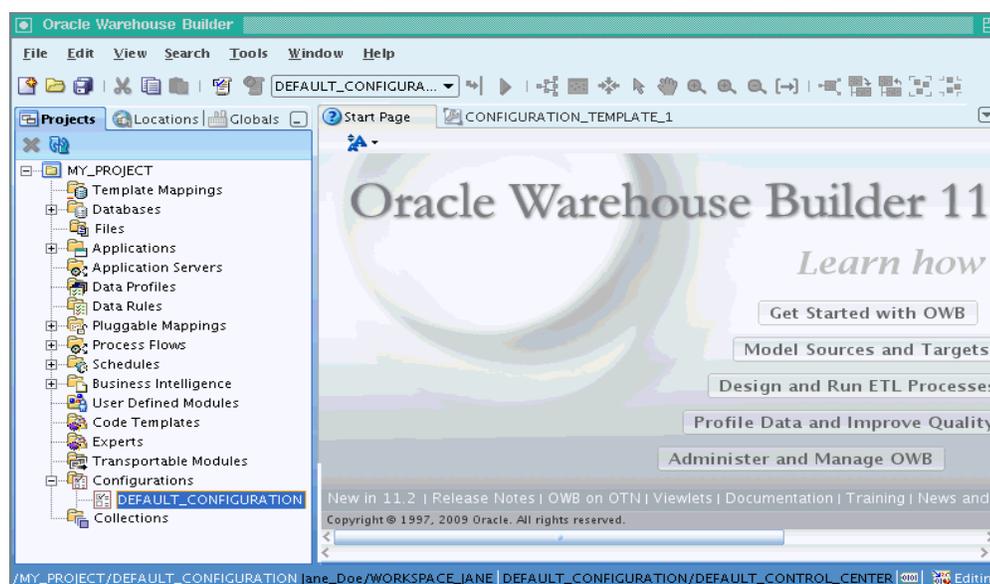


ILUSTRACIÓN 23: AMBIENTE DE TRABAJO DE OWB. (ORACLE CORPORATION, S.F.)

### 2.6.3. Tableau

Es una plataforma de inteligencia de negocios conformada esencialmente por cinco módulos principales que proveen todas las funcionalidades necesarias 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 Su crecimiento en los 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.)

La Solución de inteligencia de negocios de tableau consiste esencialmente en 2 componentes principales, tableau desktop y tableau server. Las demás aplicaciones como tableau online, tableau public, etc. 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 ilustración 24, 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.

---

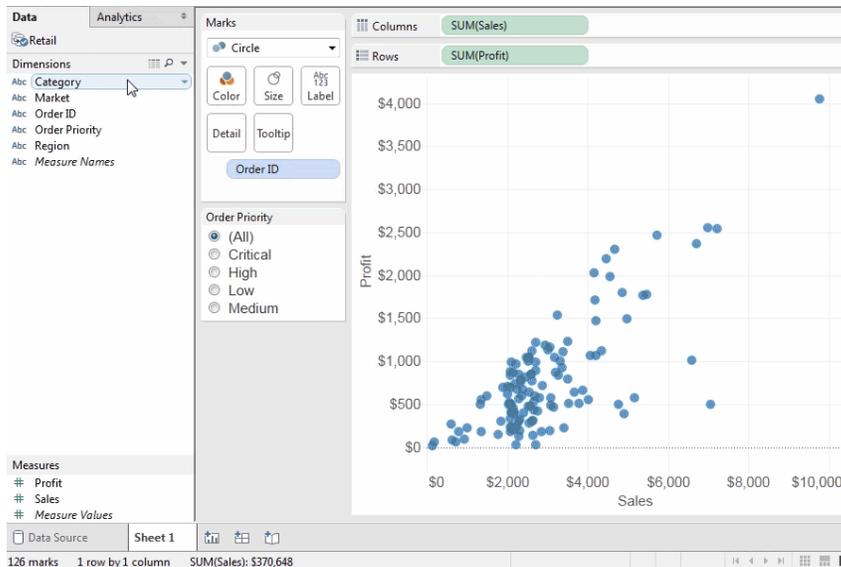
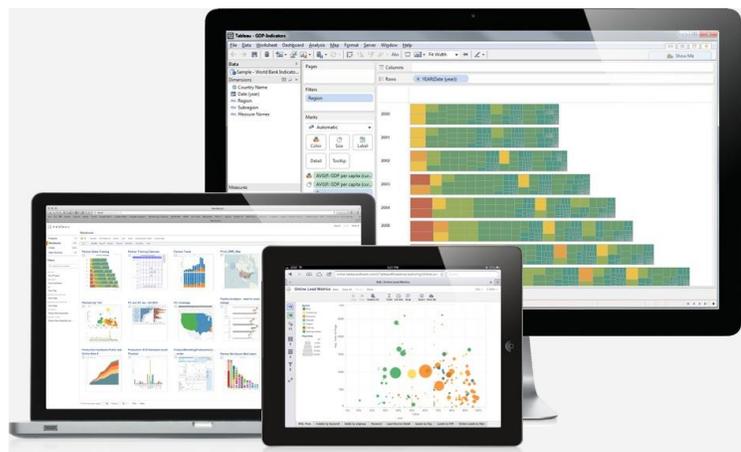


ILUSTRACIÓN 24:  
VISUALIZACIÓN DE  
TABLEAU DESKTOP.  
(TABLEAU, S.F.)

▪ **Tableau Server**

Una vez realizado los tableros de control o cualquier otro elemento dentro de Tableau Desktop, Tableau Server permite compartir la información a través de la 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 25, donde, el gráfico, menús e indicadores se adaptan según el tamaño de la pantalla. (Tableau., s.f.)

ILUSTRACIÓN 25: MODIFICACIÓN DE INTERFAZ SEGÚN RESOLUCIÓN DE PANTALLA EN TABLEAU SERVER. (LOYOLA UNIVERSITY CHICAGO, S.F.)



**2.6.4. Comparación entre las herramientas de inteligencia de negocios.**

Herramienta	Tableros de Control	ETC	Reportes/ Informes	Análisis	Licencia	Orientación
Pentaho	Pentaho Dashboards	Pentaho Data Integration.	Pentaho Reporting	Pentaho Analysis Services.	Comunitaria y Privativa	Pequeñas y Medianas Empresas
Oracle	Oracle BI Interactive Dashboards	Oracle Warehouse Builder	Oracle BI Publisher	Oracle BI Answers.	Privada	Grandes Empresas
Tableau	Tableau Desktop	-----	Tableau Desktop	Tableau Server	Comunitaria y Privativa	Pequeñas y Medianas Empresas

TABLA 2: COMPARACIÓN DE LAS HERRAMIENTAS DE INTELIGENCIA DE NEGOCIOS

**2.7. Proyecto**

El (Project Management Institute, 2013), define un proyecto como un esfuerzo temporal que se lleva a cabo para crear un producto, servicio o resultado único. La naturaleza temporal de los proyectos implica que un proyecto tiene un principio y un final definidos. El final se alcanza cuando se logran los objetivos del proyecto, cuando se termina el proyecto porque sus objetivos no se cumplirán o no pueden ser cumplidos, o cuando ya no existe la necesidad que dio origen al proyecto. Asimismo, se puede poner fin a un proyecto si el cliente (cliente, patrocinador o líder) desea terminar el proyecto.

Que sea temporal no significa necesariamente que la duración del proyecto haya de ser corta. Se refiere a los compromisos del proyecto y a su longevidad. En general, esta cualidad de temporalidad no se aplica al producto, servicio o resultado creado por el proyecto; la mayor parte de los proyectos se emprenden para crear un resultado duradero. Por ejemplo, un proyecto para construir un monumento nacional creará un resultado que se espera perdure durante siglos. Por otra parte, los proyectos pueden tener impactos sociales, económicos y ambientales susceptibles de perdurar mucho más que los propios proyectos.

Un esfuerzo de trabajo permanente es por lo general un proceso repetitivo que sigue los procedimientos existentes de una organización. En cambio, debido a la naturaleza única de los proyectos, pueden existir incertidumbres o diferencias en los productos, servicios o resultados que el proyecto genera. Las actividades del proyecto pueden ser nuevas para los miembros del equipo del proyecto, lo cual puede requerir una planificación con mayor dedicación que si se tratara de un trabajo de rutina. Además, los proyectos se llevan a cabo en todos los niveles de una organización. Un proyecto puede involucrar a una única persona o a varias personas, a una única unidad de la organización, o a múltiples unidades de múltiples organizaciones.

Los proyectos dentro de las empresas son el medio a través del cual se alcanzan los planes estratégicos.

Un proyecto puede generar:

- Un producto, que puede ser un componente de otro elemento, una mejora de un elemento o un elemento final en sí mismo.
- Un servicio o la capacidad de realizar un servicio (p.ej., una función de negocio que brinda apoyo a la producción o distribución).
- Una mejora de las líneas de productos o servicios existentes (p.ej., Un proyecto Seis Sigma cuyo objetivo es reducir defectos).
- Un resultado, tal como una conclusión o un documento (p.ej., un proyecto de investigación que desarrolla conocimientos que se pueden emplear para determinar si existe una tendencia o si un nuevo proceso beneficiará a la sociedad).

Los ejemplos de proyectos, incluyen entre otros:

- El desarrollo de un nuevo producto, servicio o resultado.
- La implementación de un cambio en la estructura, los procesos, el personal o el estilo de una organización.
- El desarrollo o la adquisición de un sistema de información nuevo o modificado (hardware o software).

- La realización de un trabajo de investigación cuyo resultado será adecuadamente registrado.
- La construcción de un edificio, planta industrial o infraestructura.
- La implementación, mejora o potenciación de los procesos y procedimientos de negocios existentes.

### **2.7.1. Ciclo de vida de los proyectos**

El ciclo de vida de un proyecto es la serie de fases por las que atraviesa un proyecto desde su inicio hasta su cierre. Las fases son generalmente secuenciales y sus nombres y números se determinan en función de las necesidades de gestión y control de la organización u organizaciones que participan en el proyecto, la naturaleza propia del proyecto y su área de aplicación. Las fases se pueden dividir por objetivos funcionales o parciales, resultados o entregables intermedios, hitos específicos dentro del alcance global del trabajo o disponibilidad financiera. Las fases son generalmente acotadas en el tiempo, con un inicio y un final o punto de control. Un ciclo de vida se puede documentar dentro de una metodología. Se puede determinar o conformar el ciclo de vida del proyecto sobre la base de los aspectos únicos de la organización, de la industria o de la tecnología empleada. Mientras que cada proyecto tiene un inicio y un final definidos, los entregables específicos y las actividades que se llevan a cabo variarán ampliamente dependiendo del proyecto. El ciclo de vida proporciona el marco de referencia básico para dirigir el proyecto, independientemente del trabajo específico involucrado.

La (Online Business School, 2014) establece que el ciclo de vida de todo proyecto, como se observa en la ilustración 26, se estructura en torno a cinco fases:



ILUSTRACIÓN 26: CICLO DE VIDA DE LOS PROYECTOS

**a) Inicio:** es el momento de definir el alcance y proceder a la selección del equipo.

**b) Planificación:** Ésta es a menudo la fase más difícil para un director de proyecto, ya que tiene que hacer un importante esfuerzo de abstracción para calcular las necesidades de personal, recursos y equipo que habrán de preverse para lograr la consecución a tiempo y dentro de los parámetros previstos. Asimismo, también es necesario planificar comunicaciones, contratos y actividades de adquisición. Se trata, en definitiva, de crear un conjunto completo de planes de proyecto que establezcan una clara hoja de ruta.

**c) Ejecución:** En base a la planificación, habrá que completar las actividades programadas, con sus tareas, y proceder a la entrega de los productos intermedios. Es importante velar por una buena comunicación en esta fase para garantizar un mayor control sobre el progreso y los plazos. Asimismo, es indispensable monitorizar la evolución del consumo de recursos, presupuesto y tiempo, para lo que suele resultar necesario apoyarse en alguna herramienta de gestión de proyectos.

**d) Cierre:** Esta fase comprende todos procesos orientados a completar formalmente el proyecto y las obligaciones contractuales inherentes. Una vez terminado este estadio, se establece formalmente que el proyecto ha concluido.

**e) Control y Seguimiento:** Esta fase comprende los procesos necesarios para realizar el seguimiento, revisión y monitorización del progreso de proyecto. Se concibe como el medio de detectar desviaciones con la máxima premura posible, para poder identificar las áreas en

las que puede ser requerido un cambio en la planificación. La etapa de seguimiento y control se encuentra naturalmente asociada a la de ejecución, de la que no puede concebirse de forma separada.

Las etapas de un proyecto suelen completarse secuencialmente, aunque en algunos momentos puntuales pueden coexistir.

### **2.7.2. Características de un proyecto**

Un proyecto es una idea que surge de una necesidad que puede presentarse en una organización, esta idea tiene como objetivo satisfacer dicha necesidad, y debe ser planificada cuidadosamente para poder obtener los resultados esperados, previendo los posibles obstáculos y dificultades que puedan presentarse, Según él (Project Management Institute, 2013), se definen tres grandes características de un proyecto:

**1) Temporal:** Significa que cada proyecto tiene un inicio definido y un final definido. El final se alcanza cuando se han logrado los objetivos del proyecto o cuando queda claro que los objetivos del proyecto no serán o no podrán ser alcanzados, o cuando la necesidad del proyecto ya no exista y el proyecto sea cancelado. Temporal no necesariamente significa de corta duración; muchos proyectos duran varios años, sin embargo, la duración de un proyecto es limitada. Los proyectos no son esfuerzos continuos.

**2) Productos, servicios o resultados únicos:** Un proyecto crea productos entregables únicos. Productos entregables son productos, servicios o resultados. Los proyectos pueden crear:

- Un producto o artículo producido, que es cuantificable, y que puede ser un elemento terminado o un componente.
- La capacidad de prestar un servicio como, por ejemplo, las funciones del negocio que respaldan la producción o la distribución. Un resultado, tales como salidas o documentos. Por ejemplo, de un proyecto de investigación se obtienen conocimientos que pueden usarse para determinar si existe o no una tendencia o si un nuevo proceso beneficiará a la sociedad.

**3) Elaboración gradual:** Es una característica de los proyectos que acompaña a los conceptos de temporal y único. “Elaboración gradual” significa desarrollar en pasos e ir aumentando mediante incrementos.

### **2.7.3. Gerencia de proyectos**

De acuerdo a (Palacios, 2003), la Gerencia de Proyectos es una aplicación sistemática de una serie de conocimientos, habilidades, herramientas y técnicas para alcanzar o exceder los requerimientos de las partes interesadas de un proyecto.

La gerencia de proyectos requiere de una serie de acciones (procesos) con el objeto de obtener productos o servicios eficazmente acabados a través de emprendimientos temporales para los cuales se cuenta con una serie de recursos que deben ser utilizados de manera óptima si se quiere asegurar la ejecución cabal de cada fase del proyecto, desde su conceptualización hasta su puesta en marcha y cierre administrativo.

Se considera que un proyecto ha sido eficazmente acabado cuando se diferencia de otros por haber alcanzado los objetivos dentro de los parámetros de tiempo, costo, calidad, optimización de los recursos y finalmente logrando la satisfacción del usuario o cliente.

#### **2.7.3.1. Áreas de conocimiento de la gerencia de proyectos**

Un proyecto puede ser visto desde diferentes puntos de vista, como lo son el tiempo, que es un recurso, pero de tal importancia que podemos apartarlo como un punto de vista por sí mismo, la eficacia, que incluye la calidad y la capacidad de resolución de la necesidad que pueda tener el resultado de nuestro proyecto, y la eficiencia, que incluye la manera en que administramos los recursos del proyecto y su buen uso para no desperdiciarlos, como podemos ver en la ilustración 27. Pero estos puntos de vista pueden desglosarse aún más, en lo que conocemos como áreas de conocimiento de la gerencia de proyectos.

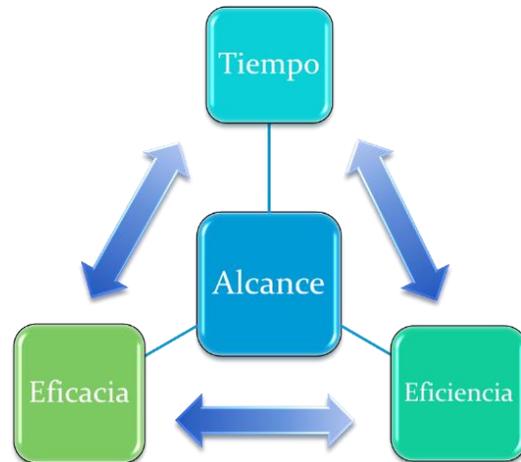


ILUSTRACIÓN 27:  
PUNTOS DE VISTA DE  
LOS PROYECTOS.

Según el (Project Management Institute, 2013), existen nueve áreas de conocimientos a ser cubiertas por la gerencia de proyectos. Las áreas de conocimiento son complementarias entre sí y se integran a las fases de un proyecto través de la gerencia de proyectos. Cada área de conocimiento está referida a una disciplina específica y su aporte al proyecto es significativo, tanto a escala individual, como en su conjunto.

**1) Gerencia de Integración de Proyecto:** Incluye los procesos necesarios para coordinar e integrar todos los elementos de un proyecto. Se contempla dentro de esta área: el desarrollo del plan del proyecto, la ejecución del plan de proyecto y el sistema integrado de control de cambios.

**2) Gerencia de Alcance de Proyecto:** Agrupa los procesos requeridos para garantizar que el proyecto esté bien definido al nivel de alcance. Incluye los procesos de: iniciación, planificación de alcance, definición de alcance hasta el nivel de desagregación necesario para ser gerenciado, verificación de alcance y control de cambios de alcance.

**3) Gerencia de Tiempo:** Considera todos los elementos que aporten a al completado a tiempo de un proyecto. Agrupa los siguientes procesos: definición de actividad, secuencia de actividad, estimación de duración de actividad, desarrollo de cronograma de ejecución y control de cronograma de ejecución.

**4) Gerencia de Costos:** Contempla todos los procesos requeridos para garantizar que el proyecto sea completado dentro del presupuesto aprobado. Contempla los procesos de: planificación de recursos, estimación de costos, presupuestos de costos y control de costos.

**5) Gerencia de Calidad:** Considera los procesos involucrados para alcanzar los niveles de calidad comprometidos en el proyecto. Incluye los siguientes procesos: planificación de calidad, aseguramiento de calidad y control de calidad.

**6) Gerencia de Recursos Humanos:** Incluye los procesos referidos al área de recursos humanos y necesarios para lograr la mayor eficiencia y efectividad de las personas que se involucren e interactúen en el proyecto. Incluye los procesos de: planificación de organización, adquisición de personal y desarrollo de personal individualmente y como equipo de trabajo.

**7) Gerencia de Comunicaciones:** Agrupa los procesos requeridos para garantizar oportunamente la generación, colección, diseminación, almacenamiento y adecuada disposición de la información del proyecto. Contempla los procesos de: planificación de comunicaciones, distribución de información, reporte(s) de desempeño y cierre administrativo.

**8) Gerencia de Riesgo:** Considera el proceso sistemático para identificar, analizar y responder a los riesgos del proyecto. Involucra procesos de: planificación de gerencia de riesgo, identificación de riesgos, análisis cualitativo de riesgos, análisis cuantitativo de riesgos, planificación de respuesta a riesgos y monitoreo y control de riesgos.

**9) Gerencia de Procura:** Se refiere a los procesos requeridos para llevar a cabo la procura de servicios y bienes relacionados a un proyecto. Involucra los siguientes procesos de gerencia de proyectos: planificación de procura, planificación de solicitudes de oferta, proceso de solicitud de oferta, selección de proveedor, administración de contratos y cierre de contrato.

## 2.8. Indicadores

Existen una serie de conceptos de indicador; no obstante, el concepto de indicador varía de unos autores a otros. Para unos cumple una función informativa, para otros es un instrumento en la toma de decisiones, o bien cumple una función de evaluación.

La Organización para la Cooperación y Desarrollo Económico (OCDE, 2000), define el término indicador como un "parámetro o el valor resultante de un conjunto de parámetros, que ofrece información sobre un fenómeno, con un significado más amplio que el directamente asociado a la configuración del parámetro".

(Perez, 2003), dice que un indicador es una medida de la condición de un proceso o evento en un momento determinado. Los indicadores en conjunto pueden proporcionar un panorama de la situación de un proceso, de un negocio, de la salud de un enfermo o de las ventas de una compañía.

Un indicador es un elemento de gestión que proporciona información acerca de aspectos críticos o claves de una organización mediante la relación de dos o más datos. (Pacheco y Castañeda, 2002). Los indicadores deben ser elegidos cuidadosamente para monitorear las áreas más importantes de la empresa, y poder corregir a tiempo lo que no esté saliendo de acuerdo a lo planificado.

Según el Departamento Administrativo Nacional de Estadística (DANE, 2009), un indicador es una expresión cualitativa o cuantitativa observable, que permite describir características, comportamientos o fenómenos de la realidad a través de la evolución de una variable o el establecimiento de una relación entre variables, la que, comparada con períodos anteriores, productos similares o una meta o compromiso, permite evaluar el desempeño y su evolución en el tiempo.

La definición más simple de indicador corresponde a la identificación de una magnitud numérica referida a un evento, que pone en evidencia la intensidad, situación o evolución del mismo. De su tratamiento es posible establecer la explicación, evolución y predicción de un fenómeno estudiado. Constituye un instrumento que pueden permitir evaluar de manera

objetiva aspectos particulares del proceso de ejecución o de los resultados de un programa o proyecto a través de mediciones de carácter cualitativo o cuantitativo.

Un indicador siempre debe estar unido a la definición de los objetivos a alcanzar, ya que los indicadores son medidas cuantitativas de desempeño que sólo cobra significado si se pone en consonancia con el objetivo que previamente se establece.

### 2.8.1. Características de los Indicadores

De acuerdo a (DANE, 2009), los indicadores deben cumplir las siguientes características:

- **Simplificación:** La realidad en la se actúa es multidimensional, un indicador puede considerar alguna de las siguientes dimensiones (economía, social, cultural, política, entre otras), pero no puede abarcarlas todas.
- **Medición:** Permite comparar la situación actual de una dimensión de estudio en el tiempo o respecto a patrones establecidos.
- **Comunicación:** Todo indicador debe transmitir información de un tema en particular para la toma de decisiones.

### 2.8.2. Tipos de Indicadores

Por medio de (Jaramillo, 2006), tenemos la siguiente clasificación para los indicadores:

- **Indicadores de Procesos:** Están relacionados con el conjunto de actividades que forman parte del proceso.
- **Indicadores de Resultados:** Se refiere al comportamiento del proceso como un todo.
- **Indicadores de Gestión:** Son medidas utilizadas para determinar el nivel de éxito de los requerimientos u objetivos de una actividad perteneciente a un proyecto o del proyecto en sí. Los indicadores de gestión están relacionados con la administración.
- **Indicadores de Eficacia:** Eficaz tiene que ver con hacer efectivo un intento o propósito. Estos indicadores miden si los objetivos y metas se cumplieron. Implican que las cosas

que se deben realizar, hacen lo que se tiene que hacer, es decir, miden si las cosas hechas son las cosas que deben ser. Por tal motivo, en el establecimiento de un indicador de eficacia es fundamental conocer y definir operacionalmente los requerimientos del cliente del proceso para comparar lo que entrega el proceso contra lo que él espera. Ejemplo: grado de satisfacción de los clientes con relación a los pedidos.

- **Indicadores de Eficiencia:** Teniendo en cuenta que eficiencia tiene que ver con la actitud y la capacidad para llevar a cabo un proceso o una tarea con el mínimo de recursos. Estos indicadores están relacionados con las razones que indican los recursos invertidos en la consecución de tareas y/o procesos, es decir, indican si se hacen las cosas de forma correcta, lo mejor posible, de la mejor manera, minimizando los recursos en la menor cantidad posible. Ejemplo: Tiempo fabricación de un producto, razón de piezas / hora, rotación de inventarios.
- **Indicador de Efectividad:** este indicador correlaciona los dos anteriores y mide el impacto en el logro de los resultados, se evalúa por el grado de cumplimiento de los objetivos.

### 2.8.3. Objetivos de los Indicadores

Según (DANE, 2009), los indicadores son herramientas útiles para la planeación y la gestión en general de un negocio, y sus objetivos principales son:

- Generar información útil para mejorar el proceso de toma de decisiones, el proceso de diseño, implementación o evaluación de un plan, programa, entre otros.
- Monitorear el cumplimiento de acuerdos y compromisos.
- Cuantificar los cambios en una situación que se considere problemática.
- Efectuar seguimiento a los diferentes planes, programas, y proyectos que permita tomar los correctivos oportunos y mejorar la eficiencia y eficacia del proceso en general.

### 2.8.4. Paradigmas de los indicadores

Los indicadores de gestión representan un paradigma a la hora de la medición:

---

- **La medición precede al castigo:** muchas de las empresas, las personas o administración encargadas de elaborar esas mediciones, toman la medición como un mecanismo de presión o como justificación para sancionar al personal, lo cual crea un rechazo. En cambio, con la medición se busca es generar rasgos de autonomía de decisión y acción razonable para los empleados.
- **No hay tiempo para medir:** algunas veces los empleados creen que usar los indicadores es un trabajo extra a sus tareas diarias, y no se dan cuenta que los mismos los llevan a ese control de la organización.
- **Medir es difícil:** la medición es difícil dependiendo de que se quiera medir o a que quiere llegar la organización con estas mediciones, ya que algunas veces con solo aplicar matemática sencilla, reglas de tres, restas, sumas, divisiones y elementos estadísticos, es suficiente para obtener buena información de la organización.
- **Hay cosas imposibles de medir:** es cierto que en algunos casos la medición de algunos agentes, procesos o variables es sumamente compleja, pero no siempre pasa así.
- **Es más costoso medir que hacer:** esto pasa ya que se tiene que saber que no se puede medir todos los procesos de la organización, hay que tener presente que se debe de medir las variables más representativas o las que mejor identifiquen los aspectos más vitales de la organización.

### 2.8.5. Utilidad de los Indicadores

Las organizaciones actualmente se ven afectadas por diferentes factores en su desempeño, de los departamentos que conformen dicha organización. El comportamiento de estos factores es probabilístico, ya que, estos pueden reducir esa duda que se tiene por medio de la información registrada o captada.

Algunas de las causas de por qué se quiere obtener esta información son estas:

- La exposición de la información.

- La creciente complejidad de la administración.
- El ritmo rápido del cambio.
- La interdependencia de las unidades que conforman la organización.
- El reconocimiento de la información como recursos.
- La evaluación y la disponibilidad de la tecnología y de las telecomunicaciones.
- La necesidad de desarrollo de las organizaciones y de las personas.

### **2.8.6. Ventajas de los Indicadores**

Los indicadores se utilizan ya que trae una ventaja fundamental para la organización, debido a que el uso de los indicadores tiene consigo una reducción drástica de la duda, de la angustia y subjetividad, con el consecuente incremento del bienestar de todos los trabajadores. Estas son algunas ventajas que trae el uso del manejo de los indicadores de gestión:

- Estimular y promover el trabajo en equipo.
- Contribuir con el desarrollo y el crecimiento tanto personal como el equipo dentro de la organización.
- Generar un proceso de innovación y enriquecimientos del trabajo diario.
- Impulsar la eficiencia, la eficacia y la productividad de cada uno de los negocios.

## **2.9. Obra Civil**

### **2.9.1. Definición**

La cosa hecha o producida por el hombre se conoce como obra. Puede tratarse de un producto material o intelectual, protegido por diversas leyes. El concepto también se utiliza para nombrar al proceso de construcción de un edificio o de una infraestructura en general. Civil, por su parte, es un adjetivo que refiere a lo perteneciente a los ciudadanos o la ciudad.

Las obras civiles son todos los trabajos realizados con la finalidad de la construcción, conformación, mantenimiento o demolición de una estructura o elemento en un espacio

determinado. Elaborado por un grupo técnico entre profesionales, personal obrero, materiales y equipo correspondiente. (De Sousa, Las obras civiles., 2014)

La noción de obra civil está vinculada al desarrollo de infraestructuras para la población. En este caso, el uso del término civil procede de la ingeniería civil, que recibe dicha denominación para diferenciarse de la ingeniería militar.

La obra civil, por lo tanto, es la aplicación de nociones de la física, la química, la geología y el cálculo para la creación de construcciones relacionadas con el transporte, la hidráulica, etc.

Las obras civiles tienden a contribuir a la organización del territorio y al aprovechamiento que se hace de éste. Las carreteras que posibilitan la circulación de medios de transporte, las represas que ayudan a gestionar los recursos hídricos, los puentes que permiten atravesar un río y el alcantarillado son algunos ejemplos de obras civiles.

En el amplio grupo de las obras civiles, pueden distinguirse trabajos pertenecientes a la ingeniería geotécnica (que estudia la resistencia entre partículas para garantizar si el suelo puede soportar una determinada estructura), la ingeniería estructural (estima la resistencia de los elementos sometidos a cargas), la ingeniería de transporte e infraestructura vial (especializada en la satisfacción de las necesidades de movilidad) y la ingeniería hidráulica (vinculada a la ejecución de obras relacionadas con el agua). La sociedad se beneficia directamente de una obra civil, y ninguna ciudad moderna puede sostenerse sin esta actuación del ser humano.

### **2.9.2. Características de las obras civiles**

- Optimiza los espacios territoriales para contribuir con la naturaleza humana de esparcimiento, trabajo y vivienda.
- Para la ejecución de los trabajos se cuenta con una serie de equipos, materiales y personal técnico.
- Se pueden desarrollar en cualquier espacio donde pueda interferir o no con la naturaleza.
- Requiere de una densa planificación y control de gestión para el cumplimiento de los objetivos.

### 2.9.3. Tipos de obras civiles

Existen dos tipos de obras civiles:

- **Obra Horizontal:** son construcciones que van a la par de la superficie terrestre y se construye partiendo desde un punto fijo hacia otro punto fijo, tales como carreteras, puentes, rampas, andenes, avenidas, acueductos, entre otros.
- **Obra Vertical:** son construcciones que se ejecutan desde un punto de la superficie terrestre hacia arriba o hacia abajo, se pueden mencionar edificios, hoteles, centros comerciales, escuelas, bibliotecas, universidades, sistemas de transporte público subterráneos, etc.

### 2.9.4. Proyectos de Obra Civil

En el campo de la arquitectura y la ingeniería civil, el proyecto es el conjunto de documentos mediante los cuales se define el diseño de una construcción antes de ser realizada. Es el documento base sobre el que se desarrolla el trabajo de los arquitectos, ingenieros y proyectistas de distintas especialidades.

A lo largo de un proyecto se desarrolla la distribución de usos y espacios, la utilización de materiales y tecnologías, y la justificación técnica del cumplimiento de las especificaciones requeridas por la normativa técnica aplicable. En muchos ámbitos la elaboración de un proyecto completo es obligatoria antes de iniciar el desarrollo de una construcción, y puede tener carácter contractual

El Proyecto de Obra Civil es para determinar el nivel óptimo del terreno, la cantidad mínima a excavar y rellenar, la conformación de las vialidades internas de acuerdo a los factores colindantes y así hacer viable y funcional un terreno. (Forta Ingeniería, 2015)

### 2.9.4.1. Etapas de un Proyecto de Obra Civil

Los proyectos de obra civil se caracterizan por su complejidad; como suponen el diseño y la construcción de grandes estructuras, edificios, avenidas, puentes, ferrocarriles y en general megaobras, las etapas previas suelen ser costosas y requieren de tiempos que a veces son mayores incluso que los de la fase de ejecución.

Para que un proyecto de construcción pueda materializarse deben cumplirse una serie de pasos, los que se detallan a continuación (Constructor Civil, 2016):

**1.- Existencia de una necesidad:** Para que un proyecto se origine debe existir una necesidad insatisfecha, la cual puede ser, por ejemplo, un edificio para solucionar un problema habitacional, un puente que permita la comunicación de una zona aislada, un monumento que recuerde algún hecho o persona importante de la historia, etc.

**2.- Análisis:** En esta segunda etapa se analizan las necesidades, seleccionándose las más relevantes, para lo cual se deben considerar los siguientes aspectos:

- Identificar las causas que originan la necesidad de un proyecto, tales como: modificación del medio, política de desarrollo, modificación de las características de la demanda, obsolescencia de la infraestructura existente y requerimiento de nuevas infraestructuras.
- Establecer los objetivos que debe satisfacer el proyecto, tales como: sociales, económicos, funcionales y de lucro.
- Priorizar las necesidades en función de los objetivos prioritarios establecidos.

**3.- Identificación de soluciones:** En esta etapa se identifican todas las posibles soluciones que permitan resolver el problema planteado en las etapas previas. Se propone en una primera etapa soluciones a nivel de conceptualización privilegiando la imaginación, más que las restricciones.

**4.- Estudios de factibilidad:** Una de las etapas importantes en el ciclo de un proyecto es realizar estudios de factibilidad, los cuales consisten en determinar si el proyecto en estudio es viable desde un punto de vista medioambiental, técnico, económico, administrativo y legal.

**5.- Evaluación:** Se evalúan todas las alternativas posibles que permitan satisfacer las necesidades seleccionadas y se elige por lo general la que presenta una mejor factibilidad técnica y económica, que cumpla con las exigencias.

**6.- Financiamiento:** Una vez decidido el proyecto técnico es importante considerar el aspecto de financiamiento, es decir, como se pagarán los gastos en que se incurrirá en la materialización del proyecto. El financiamiento puede ser propio o a través de un préstamo.

**7.- Diseño:** Una vez determinada la solución que se usará para satisfacer la necesidad se diseña el proyecto, en el diseño normalmente considera los siguientes aspectos:

- Estudio del terreno donde se va a construir la obra, analizando sus condiciones generales y reglamentarias, su topografía, geología, hidrología, ambientales, legales, históricas, etc.
- Diseño arquitectónico, normalmente considera las siguientes etapas: establecimiento de los requerimientos del dueño, preparación de un ante proyecto y. finalmente el diseño del proyecto arquitectónico definitivo, que incluye planos y especificaciones.
- Diseño estructural de la obra para que sea capaz de resistir los esfuerzos a los cuales estará sometida durante su vida útil. Las etapas principales de este diseño son:
  - Determinación de los esfuerzos que solicitarán a la estructura.
  - Estructuración, determinación de los elementos resistentes.
  - Diseño de los elementos estructurales y configuración de planos.
  - Confección de las especificaciones técnicas.
  
- Estudios de impacto ambiental, analizando las consecuencias del proyecto en el medio ambiente.
- Diseño de las instalaciones, que consiste en dar a la estructura la funcionalidad que requerirá para ser ocupada con el fin para el que se la diseñó. Entre las instalaciones típicas

están: las eléctricas, las de gas, las de agua potable y las de alcantarillado y muchas otras (tales como: alarma, climatización, red computacional, red de incendio, etc.).

- Redacción de los documentos de licitación: Finalmente, se deben redactar todos los documentos que permitan llamar a la licitación del proyecto.
- Constructabilidad y Mantenición: Dos aspectos clave que deben considerarse en esta etapa de diseño. El primero debe incluir en forma explícita la forma más eficiente de materializar el proyecto. El segundo, en cambio, debe hacerse cargo anticipadamente de cómo se va a llevar a cabo la conservación del proyecto durante su operación.

**8.- Licitación:** Llamado a licitación y adjudicación. El llamado a licitación puede ser público o privado y la adjudicación puede estar previamente reglamentada o ser de absoluto criterio del mandante. La adjudicación a su vez, puede ser negociada o no, dependiendo de las reglas de licitación.

**9.- Construcción:** Esta etapa es una de las más importantes debido a que en ella se materializa la obra. Las etapas principales incluyen:

- Definición de una estrategia de gestión y calidad.
- Obtención de los permisos para realizar la obra.
- Redacción y aceptación de un contrato, en el cual se fijan plazos, costos y las relaciones entre dueño y contratista.
- Metodología de trabajo, en que se determinan métodos más eficientes y racionales para la construcción, dado los recursos disponibles.
- Planificación y Programación de la obra, en que se fijan plazos parciales y totales, y se planifica el uso de los recursos disponibles a través de la construcción.
- Estudios de presupuestos.
- Contrato de la fuerza laboral necesaria para construir la obra.
- Adquisición de los materiales y arriendo o compra de la maquinaria necesaria para la materialización.
- Materialización física de la obra.

- Control, donde se confronta lo realizado o lo que se debería haber hecho de acuerdo a lo programado y especificado. Este control puede ser interno (o auto-control), externo (normalmente contratado por el mandante) o ambos.

**10.- Puesta en marcha:** En esta etapa se entrega al servicio la obra, realizándose previamente diferentes controles para determinar la calidad de la construcción, entre los que se destacan:

- Verificación de pruebas y ensayos de calidad realizados.
- Revisión detallada de todos los elementos construidos y terminaciones.
- Pruebas de funcionamiento.
- Aprobación final.

**11.- Operación y mantenimiento:** Esta es una etapa que no siempre es considerada adecuadamente. Esta actividad es fundamental para el buen funcionamiento y durabilidad de la estructura y debería ser considerada desde la etapa de diseño.

**12.- Abandono:** Muchos proyectos, una vez que su objetivo y vida útil se cumplen, deben abandonarse. Por lo tanto, esta actividad debe pensarse y diseñarse con anticipación, para minimizar, los impactos ambientales y económicos.

## CAPÍTULO 3

### MARCO METODOLÓGICO

Al momento de desarrollar una solución de inteligencia de negocio es importante tomar en cuenta varios factores, que van desde un buen estudio de los procesos de la organización, hasta la selección de los técnicas, métodos y herramientas que mejor se adapten a las necesidades de la organización y que permitan obtener resultados confiables y satisfactorios. El no darles la importancia requerida a estos factores puede ocasionar que la construcción de la solución de inteligencia de negocio se torne más complicado de lo necesario e incluso pueda fallar en su implementación.

Tomando en consideración lo mencionado anteriormente, estudiar una o varias metodologías que se ajusten adecuadamente, al diseño y construcción de la solución de inteligencia de negocio se vuelve un requerimiento fundamental. Además es esencial saber que la elaboración de una buena solución de inteligencia de negocio, esta intrínsecamente relacionada con la elaboración de un almacén de datos, esto se debe a que dentro de la infraestructura de una solución de inteligencia de negocios, el almacén de datos (y todo lo que tiene que ver con su elaboración) es una pieza fundamental, puesto que su desarrollo abarca desde la captación de datos desde las distintas fuentes, hasta la elaboración de un modelo que permita integrar los datos de una forma idónea para realizar procesos analíticos sobre la data.

Sobre el universo que tiene que ver el desarrollo de una estructura de datos propicia para la toma de decisiones (almacenes de datos), sobresalen dos teorías de desarrollo de dos figuras muy importantes dentro del área, Ralph Kimball y Bill Inmon. Aunque ambos autores tienen puntos en común, los mismos tienen filosofías distintas de trabajo para alcanzar los objetivos y el desarrollo del almacén de datos en general.

Por lo tanto, después de revisar las metodologías posibles a utilizar, se seleccionó como base, la metodología propuesta por Ralph Kimball ya que conduce a una solución completa que se

---

puede implementar en poco tiempo, y su construcción y adaptación a la situación actual de la organización es muy versátil.

### 3.1. Metodología según Ralph Kimball (Ascendente o Bottom-up).

Ralph Kimball define un Almacén de Datos como “una copia de los datos transaccionales específicamente estructurados para consultas y análisis” (Kimball & Caserta, 2008), también da una segunda definición como “la unión de todas las Bodegas de Datos de una entidad”.

La metodología de Kimball es ampliamente utilizada en las corporaciones por su capacidad de ofrecer resultados medibles en un corto periodo de tiempo (en comparación a la metodología de Inmon). Ralph Kimball es un autor considerado como el "gurú" en el tema de almacenes de datos. Se ha dedicado desde hace más de 10 años al desarrollo de su metodología para que este concepto sea bien aplicado en las organizaciones y se asegure la calidad en el desarrollo de proyectos de inteligencia de negocio. Por esta razón, esta metodología se ha convertido en el estándar de facto en el área de apoyo a las decisiones empresariales.

Kimball, plantea un enfoque ascendente (bottom-up), en donde se definen previamente las bodegas de datos para luego integrarlas cuando se diseñe un almacén de datos centralizado (Kimball R. , 1998). Esto permite que la metodología sea flexible y sencilla de implementar, porque se puede construir, como se observa en la Ilustración 28, una bodega de datos como primer elemento del sistema de análisis, y luego ir añadiendo otras bodegas que compartan las dimensiones ya definidas o incluyan otras nuevas, hasta integrarlas y formar un almacén de datos corporativo.

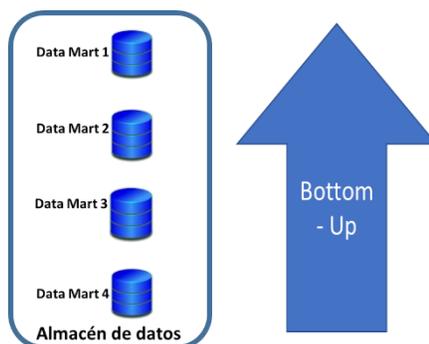


ILUSTRACIÓN 28: METODOLOGÍA  
BOTTOM-UP

La metodología de Kimball esta descrita en su libro publicado en 1998 “The Data Warehouse lyfecicle toolkit”, en este libro Kimbal establece una serie de pasos que deben seguirse para la realización de un almacén de datos. Este conjunto de pasos no es lineal, puesto que en el ciclo de desarrollo ciertas tareas pueden ser desarrolladas de forma paralela.

Con esta metodología Kimball establece los estándares necesarios para el diseño, construcción e implementación de una solución de inteligencia de negocio, la cual también es conocida como el Ciclo de Vida Dimensional del Negocio y está basada en cuatro principios básicos que son:

- **Centrarse en el negocio:** Hay que concentrarse en la identificación de los requerimientos del negocio y su valor asociado.
- **Construir una infraestructura de información adecuada:** Diseñar una base de información única, integrada, fácil de usar, de alto rendimiento donde se refleje la amplia gama de requerimientos de negocio identificados en la empresa.
- **Realizar entregas en incrementos significativos:** Crear el almacén de datos de forma incremental a través de entregables en plazos de 6 a 12 meses aproximadamente.
- **Ofrecer la solución completa:** Proporcionar todos los elementos necesarios para entregar valor a los usuarios de negocios. Esto significa, tener un almacén de datos bien implementado y también, entregar herramientas de consulta, aplicaciones para informes y análisis avanzado, capacitación, soporte y documentación.

De acuerdo con (Kimball & Ross, 2013), el Ciclo de Vida Dimensional del Negocio, representado en la Ilustración 30, describe un flujo de tareas de alto nivel requeridas para el diseño, desarrollo e implementación de un almacén o bodega de datos, cuya ejecución da paso a la construcción de una solución de inteligencia de negocio. Por lo tanto, este ciclo asegura que las piezas del proyecto a iniciar bajo esta metodología, se reúnan en el orden correcto y en el momento adecuado.

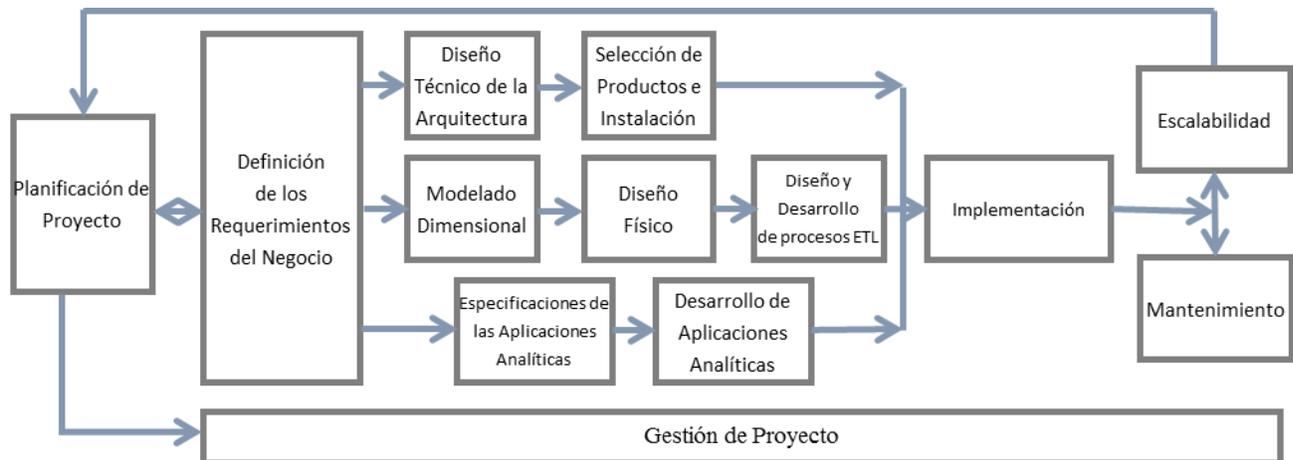


ILUSTRACIÓN 29: CICLO DE VIDA DIMENSIONAL DEL NEGOCIO. (KIMBALL & ROSS, 2013)

Sin embargo, hay que destacar que no todos los detalles de las tareas del ciclo de vida deben ser ejecutados en todos los proyectos, ya que esto dependerá del alcance y de la planificación que se defina sobre los mismos, por esta razón la primera de las tareas sirve como base principal para el desarrollo del sistema.

No obstante, antes de entrar en detalle con cada una de las tareas involucradas, es necesario conocer lo siguiente:

- La planificación de proyectos es el soporte inicial de las tareas subsiguientes. Durante este módulo se evalúa la disposición de la organización para poder llevar a cabo una solución de inteligencia de negocios.
- La gestión de proyectos sirve como base para mantener el resto de las tareas que comprenden el ciclo de vida.
- La planificación del proyecto y la definición de los requisitos de negocio, están estrechamente relacionadas porque de ellas parte la comprensión y definición de lo que se quiere y se va a desarrollar.
- El flujo superior de tareas, comprendido por el diseño de arquitectura y la selección de productos e instalación, se refiere a la tecnología. En esta parte se establece el marco

general para apoyar la integración de las múltiples tecnologías que se utilizaran para el desarrollo del proyecto.

- El flujo del medio de tareas, conformado por el modelado dimensional, diseño físico y, el diseño y desarrollo de procesos ETL, se centra en los datos. Con estas tareas se diseña e implementa el modelo dimensional, y se desarrollan los procesos de extracción, transformación y carga (ETL) para poblar con datos las bodegas y/o almacenes de datos.
- El flujo final de tareas, generado por la definición de los requerimientos de negocio, comprende el diseño y desarrollo de aplicaciones de análisis, las cuales permiten satisfacer las necesidades analíticas de los usuarios finales.
- La unión de los tres flujos descritos anteriormente (tecnología, datos y aplicaciones analíticas), junto con una buena base teórica, permiten el desarrollo de un buen proyecto.
- El mantenimiento continuo permite asegurar que el proyecto se mantiene en buen estado.

El conjunto de pasos que plantea Kimball son descritos a continuación:

### **3.1.1. Planificación del proyecto.**

La planificación de un proyecto de construcción de un almacén de datos marca el inicio de su desarrollo. Da pie a una serie de actividades como lo son la definición del proyecto, el alcance, evaluación y la justificación del proyecto. En esta fase se evalúa cual será el alcance del almacén de datos o mercado de datos, cuales son los recursos, nivel o habilidades del equipo que desarrollara el proyecto, así como los distintos roles que asumirán dentro de la fase de desarrollo. También se establecen los recursos o los elementos que se poseen para desarrollar el proyecto (datos, acceso a elementos de la organización, dinero, etc.).

La planificación de proyecto y el consecuente proyecto que busca generarse tiene que justificarse, es decir, en esta fase se tiene que establecer cuáles serán los beneficios de la organización. La justificación comprende la estimación de los beneficios y costos asociados

al desarrollo e implementación de este tipo de solución. Es necesario considerar los costos aproximados para el hardware y software que serán usados.

### **3.1.2. Definición de requerimientos del negocio.**

Esta fase es el centro de desarrollo del almacén de datos puesto que sirve como punto de partida para definir el alcance, los modelos de datos operacionales que se utilizaran, las características de dichos modelos operacionales y cuáles son las necesidades analíticas del grupo de usuarios que serán afectados por el almacén de datos.

Según (Kimball R. , 1998) “las aplicaciones de usuarios finales están definidas por los requerimientos. Finalmente, tu implementación, mantenimiento y planes de crecimientos deben estar impulsada por los usuarios. Uno debe comenzar por formular respuesta a todas estas preguntas durante el ciclo de vida basado en el entendimiento de los requerimientos de negocio de los usuarios”.

Uno de los aspectos fundamentales en esta área es poder levantar los requerimientos o necesidades de los usuarios, entendiendo que los usuarios muchas veces no tienen claramente dichas necesidades establecidas, que se les dificulta expresar sus requerimientos y darlos a entender y que muchas veces no son especialistas en el área de informática o manejo de datos, por lo que no comprenderán las implicaciones de sus requerimientos (requerimientos muy complejos en una primera fase por los modelos de datos que se están manejando).

En esta fase es vital la utilización de entrevistas, encuestas, charlas, llegar al usuario final y lograr captar todos los requerimientos, entendiendo que estos requerimientos pueden cambiar y que la puesta en escena del almacén de datos hará que la utilización del mismo por parte de los usuarios levante nuevos requerimientos.

### **3.1.3. Modelado dimensional.**

El levantamiento de los requerimientos define el modelo dimensional requerido para resolver las necesidades de los usuarios. Las fuentes de datos y los requerimientos van a definir la granularidad del modelo, los indicadores, dimensiones, jerarquías y atributos.

Es importante comprender que cada área de negocio posee sus propios requerimientos, por lo que se deben realizar modelos dimensionales departamentales (mercados de datos) que se adapten a las necesidades de los distintos grupos de usuarios existentes dentro de la organización. Una buena forma de establecer cuales dimensiones son necesitadas o utilizadas en los distintos mercados de datos es aplicar una matriz de dimensiones. En esta matriz por cada proceso de negocio son establecidas las dimensiones a utilizar.

Esta fase tiene por resultado los diseños multidimensionales y sus conexiones lógicas con las distintas fuentes de datos.

### **3.1.4. Diseño físico.**

Consiste en la construcción de las estructuras físicas dentro de la base de datos que soportaran el modelo dimensional. Dentro de esta etapa se definen estándares para el nombre de los atributos o elementos, se configura la base de datos, se crean índices o demás estructuras auxiliares requeridas.

### **3.1.5. Diseño y construcción de procesos ETL.**

Una vez creada la base de datos, se procede a crear los procesos de ETL que extraerán los datos requeridos desde las distintas fuentes de datos. Aquí se evalúan las fuentes, tipos de datos, formato de los datos entre otros. Se recomienda que la extracción se realice hacia un área intermedia en donde se realizaran todas las transformaciones dentro de la data para generar datos íntegros.

Una vez transformada la data se procede a cargarla al almacén de dato, por lo general en horas o momento de poca utilización del almacén de datos para no afectar su rendimiento, sin embargo, la frecuencia en como los procesos ETL son ejecutados dependerá de los requerimientos, es decir, a partir del área de negocio, frecuencia de los cambios dentro de los datos y la necesidad del usuario a tener datos actualizados, estos procesos se pueden ejecutar diariamente, semanalmente o cada hora.

### **3.1.6. Diseño Técnico de la arquitectura.**

La infraestructura será la encargada de soportar o permitir la realización del almacén de datos. La infraestructura incluye el hardware, las redes, elementos de seguridad, funciones de bajo nivel requeridas para la existencia del almacén de datos. Dentro de la infraestructura también se puede señalar la metada, porque es una estructura lógica da soporte a los datos y la cual puede ser utilizada por distintas herramientas para extraer información de los modelos de datos definidos dentro del sistema (metada de las fuentes de datos y metada del modelo dimensional).

Dentro de la infraestructura es importante establecer las variables que van a afectar la elección de la infraestructura y documentar la arquitectura que se posee.

La elección de la arquitectura dependerá del nivel o aspecto que sea visualizado. En este sentido a nivel del servidor, que es lugar donde se tendrá el almacén de datos, se tiene que evaluar la concurrencia de usuarios, el tamaño de los datos, la capacidad de realizar procesamiento paralelo, el sistema operativo, la volatilidad del sistema entre otros. Dentro del servidor tiene que establecerse el tipo de base de datos a utilizar, relacional o multidimensional.

A nivel de red debe evaluarse el ancho de banda que se posee desde el servidor hasta las distintas terminales de los usuarios, así como los distintos conectores que se posean (conectores a bases de datos o alguna aplicación específica).

Por último, se tiene que evaluar el terminal donde el usuario va a acceder al almacén de datos. Los distintos sistemas operativos o aplicaciones que se pueden tener, evaluando que el desempeño de las aplicaciones de usuario no se vea afectado, es decir, existe una independencia de la arquitectura. También es importante evaluar las capacidades de procesamiento y memoria del terminal del usuario final con el fin de mantener tiempos óptimos de respuesta.

### 3.1.7. Selección e instalación de productos.

Una vez pasada la parte teórica de diseño de la arquitectura (punto anterior), se pasa a la selección de los productos requeridos para soportar la infraestructura. Muchas veces estos dos procesos se ejecutan de forma paralela en el sentido de que en la evaluación de los componentes dentro del diseño de la arquitectura (como el sistema operativo), ya se selecciona la mejor opción para el almacén de datos a desarrollar.

Para cada componente se debe realizar una evaluación de los distintos productos existentes en el mercado, esto atado a escoger herramientas que puedan soportar los requerimientos de los usuarios. Dentro de la selección de productos Kimball (1998) establece que existen 4 áreas importantes donde se tiene que adquirir componentes y que son los que harán posibles la implementación del almacén de datos, estas son:

- **Plataforma de hardware:** los equipos redes que soportaran el almacén de datos.
- **Sistema manejador de base datos:** encargado de gestionar la base de datos.
- **Herramientas de ETL y área intermedia:** los elementos que tomaran las fuentes de datos y pasaran estos al almacén de datos, haciendo las transformaciones pertinentes para que estos datos puedan ser utilizados.
- **Herramienta de acceso a los datos:** todo el conjunto de herramientas que servirán a los usuarios finales para realizar análisis.

Para cada uno de estos elementos se debe evaluar los distintos productos, quienes son los fabricantes, que referencia se tiene sobre ellos, la facilidad de implementación, costes, la capacidad de que dicho componente pueda ser útil para satisfacer los requerimientos.

La selección de productos puede estar definida por las habilidades que posea el grupo de trabajo o por políticas propias de la organización, por ejemplo, ciertas organizaciones solo aceptan el uso de aplicaciones de software libre.

Dentro de la parte de selección se pueden utilizar distintas herramientas que pueden ayudar a la elección de los productos, entre estas se destaca la matriz de evaluación del producto, la cual es una matriz donde se ubican los productos que pueden complacer determinado punto (por ejemplo una matriz de evaluación de sistema manejadores de bases de datos), donde se establecen los criterios de evaluación (costo, rendimiento, capacidad de concurrencia, etc) y por cada producto se realiza su ponderación respectiva. La fase de selección de productos puede incluir además la realización de prototipos que permitan evaluar el desempeño de los distintos componentes, lo cual es especialmente útil si no se posee algún conocimiento previo de la herramienta y se desea evaluar el desempeño de dicho componente para satisfacer los requerimientos específicos en función de los demás componentes existentes dentro de la arquitectura, es decir, puede querer evaluarse el desempeño de una herramienta de acceso a los datos a partir de un terminal de un usuario, o el desempeño de un sistema manejador de base de datos en la plataforma de hardware que se posee para soportar el almacén de datos.

### **3.1.8. Especificaciones de las Aplicaciones Analíticas**

Los requerimientos de los usuarios vienen dados por la necesidad de utilizar herramientas que permitan fácilmente obtener información, esto implica crear aplicaciones que sean de fácil uso, eficientes, eficaces, flexibles entre otros aspectos.

Dentro de la especificación de las aplicaciones existen cuatro componentes básicos descritos a continuación:

**Determinar el conjunto inicial de plantillas (reportes):** Consiste en determinar un conjunto inicial de plantillas de reportes u otros elementos (tableros de control, informes) que serán desarrollados. Este grupo inicial (de unas 10-20 plantillas) tienen que ser flexibles y abarcar la mayor cantidad de reportes requeridos por los usuarios. Para esto se deben establecer un conjunto candidato de reportes, hacer prototipos, realizar presentación con los usuarios sobre los tipos de reportes y generar todas aquellas plantillas que sean fundamentales para los usuarios. El conjunto de plantillas será el punto inicial de la utilización del almacén de datos, esto implica que no necesariamente se abarquen todos los requerimientos con las mismas o que no puedan necesitarse plantillas adicionales según las necesidades de los clientes aumenten. Esto ocurre porque el almacén no se encuentra construido o todavía no se ha utilizado datos de la organización para generar información, por lo que ciertos elementos en el uso, manejo y requerimiento de las herramientas por parte de los usuarios no pueden conocerse.

**Diseño del enfoque de navegación de las plantillas:** Luego de tener un conjunto inicial de plantillas para crear reportes, es muy probable que este conjunto cambie. La puesta en ejecución del almacén va a generar cambios en el modelo en los requerimientos por lo que los usuarios van a requerir nuevas plantillas. A medida que el número de plantillas aumente es importante mantener un orden, una estructura para navegar sobre los distintos componentes. En este sentido se pueden crear un esquema de plantillas y agrupar a las mismas por sector, es decir, para el departamento de ventas, se puede tener un sector para promociones y dentro de este todos los reportes de promociones, otro sector para ventas de productos propios, otro para venta de productos internacionales y así sucesivamente. La búsqueda en este punto es lograr una navegabilidad que permita al usuario visualizar y acceder de forma eficiente a cualquier plantilla existente.

**Determinar estándares dentro de las plantillas:** Aunque las plantillas son un esquema del reporte a generar y van a existir una gran variedad de las mismas. Es importante mantener un estándar entre todas las plantillas generadas. En este sentido es aconsejable mantener el nombre de objetos o localizaciones, todas las plantillas deben de tener un formato parecido (colores, letras, alineación del texto, entre otros).

**Especificar detalladamente las plantillas:** Luego de realizar todos los puntos anteriores, se procede a realizar bosquejos de las plantillas, así como una especificación de las mismas a los usuarios. En este punto se consolidan las plantillas iniciales a utilizar.

### **3.1.9. Desarrollo de Aplicaciones Analíticas**

Una vez especificadas las plantillas a generar se procede a la fase de desarrollo. En este punto es importante establecer cuál es el método de acceso, es decir, si las aplicaciones serán una aplicación web, una aplicación de escritorio (en cual sistema operativo), entre otros. También debe seleccionarse la herramienta de acceso y la forma como esta herramienta se conectará con el almacén de datos, en este caso, se requieren pruebas de validación para comprobar que los informes están dando datos correctos.

Esta fase implica la documentación de las aplicaciones generadas, esto puede hacerse a través de manuales, diagramas UML entre otros.

### **3.1.10. Implementación.**

La parte de implementación conlleva a su vez dos grupos de tareas. La primera es la instalación y configuración de elementos. Por lo general el desarrollo de las aplicaciones, el almacén de datos y los procesos ETL no se realizan directamente en la organización, sino que se ejecutan a partir de un subconjunto de datos dados por la organización en un ambiente de ejecución distinto. Por lo que una vez terminado el desarrollo, todos los elementos deben de ser puestos en el ambiente de producción, lo cual requiere la configuración del servidor, las aplicaciones, entre otros.

Una vez puesto en producción el almacén de datos se pasa a la segunda tarea, la cual es la educación a los usuarios sobre la herramienta. Dentro de los usuarios que usaran el almacén de datos existen distintos roles (administradores, soportes, analistas), cada uno de estos roles utilizara el almacén de forma distinta, lo que requiere una instrucción que permita a estos usuarios comprender el almacén de datos, así como el conjunto de herramientas que giran en

torno al almacén de datos (aplicaciones de visualización, herramientas de administración, etc.) y que utilizaran según sea el caso.

La implementación puede darse a partir de un lanzamiento beta, es decir de un lanzamiento de prueba donde solo accedan un grupo de usuarios definidos, esto permite que la puesta en producción no abarque toda la organización, entendiendo que una primera versión de la misma puede contener múltiples errores o bugs. Los usuarios beta sirven entonces para probar la herramienta y asegurar que el producto que será puesto en producción posee ciertos criterios de calidad previamente definidos (eficiencia, efectividad, usabilidad, entre otros).

### **3.1.11. Mantenimiento**

En esta fase ya se ha desarrollado e implementado el almacén de datos, sin embargo, esto no significa que el ciclo de desarrollo haya terminado.

El mantenimiento implica la evaluación del rendimiento del sistema, de la eficiencia en generar los reportes, en evaluar si los procesos ETL se están ejecutando correctamente, es decir, los datos están siendo cargados de forma correcta y los datos presentados en los reportes corresponden. También conlleva el seguir brindando soporte a la organización, ya sea con personal que gestione el sistema, así como profundizando la educación sobre el uso del almacén de datos.

### **3.1.12. Escalabilidad**

Posterior a su implementación los procesos de mantenimiento y crecimiento permiten al almacén de datos evolucionar.

El crecimiento parte del principio de evaluar en primera instancia cual ha sido el éxito del almacén de datos, es decir, como el almacén de datos ha mejorado la toma de decisiones y como esto ha repercutido en los beneficios netos de la organización. Por ejemplo, si el almacén de datos orientado al área de inventario ha disminuido la cantidad de pérdidas por concepto de alimentos perecederos que se vencieron en el almacén, ese beneficio sustenta la

existencia del almacén de datos y su posterior crecimiento. El crecimiento puede darse por la necesidad de ampliar la capacidad del almacén de datos involucrando nuevos procesos dentro de la organización o porque se han creado nuevos requerimientos. Cualquiera sea el caso, el crecimiento del almacén de datos, la generación de una nueva versión del mismo implica repetir el ciclo de desarrollo, lo que implica que el desarrollo de un almacén de datos es un proceso cíclico definido por la necesidad de adaptarse a los requerimientos cambiantes dentro de la organización.

### **3.1.13. Gestión del proyecto**

Es una tarea que se ejecuta durante todo el desarrollo del proyecto. Esta tarea implica la verificación de que las actividades del proyecto se están ejecutando según lo planificado. Por tanto, la gestión del proyecto se basa en monitorizar todas las fases del proyecto en busca de que la ejecución del proyecto no se salga de los parámetros establecidos (tiempo, costo, producto final) y en dado caso de que exista algún elemento que pueda afectar el desarrollo del proyecto, dicho factor pueda ser detectado y corregido a tiempo. En caso de que un elemento no pueda ser corregido y afecte la planificación del proyecto, se tienen que reajustar las demás actividades y sus tiempos de ejecución. Todo esto para tener como última opción la modificación de la planificación establecida, la cual debe modificarse si y solo si no existe forma de realizar el proyecto con los parámetros preestablecidos.

## CAPÍTULO 4

### MARCO APLICATIVO

La implementación de este Trabajo Especial de Grado, está basada en la metodología definida por Ralph Kimball, la cual conduce a una solución de inteligencia de negocio adaptable, que se puede realizar en poco tiempo y obtener resultados rápidos.

A continuación, se procede a describir las actividades realizadas a lo largo del proceso de desarrollo de la solución propuesta, distribuidas en cada una de las fases que comprende la metodología aplicada, que abarca 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, finalizando con la realización de un portal web el cual contiene el tablero de control (Dashboard), el conjunto de reportes y el cubo de análisis de información Mondrian (ver punto 2.6) que cumplirán con los requerimientos analíticos del área de control y seguimiento a los proyectos de obras civil, planificados por Entidades Político-Territorial.

#### **4.1. Fases del Proyecto**

##### **4.1.1. Planificación del Proyecto**

Dentro de esta fase se establecen unos lineamientos y actividades a seguir, que permiten implementar la solución de inteligencia de negocio planteada, además de cumplir con los objetivos establecidos para esta investigación en el punto 1.2.2; siguiendo un orden de ejecución basado en el ciclo de vida dimensional del negocio definido por Ralph Kimball. Dentro de estos, destacan: el levantamiento y análisis de los requerimientos, la definición de la arquitectura de la solución, la instalación y configuración de los productos tecnológicos seleccionados, el diseño e implementación del almacén de datos, la elaboración y ejecución de los procesos de extracción, transformación y carga de los datos ETLs (junto con la programación de las tareas automáticas que se encargan de ejecutarlos periódicamente), y

---

por último la construcción y publicación del tablero de control, de los reportes y del cubo de análisis de información Mondrian con los indicadores propuestos.

Las actividades como se observa en la tabla 3, son las siguientes:

Objetivo	Actividad	Duración
Definir los indicadores y reportes relacionados con la planificación, ejecución y gestión de los proyectos de obra civil en base a la identificación de los requisitos del negocio.	Reuniones, Entrevistas y Consultas	3 Semanas
	Definición de los Requerimientos del Negocio	3 Semanas
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 Semana
	Selección de Productos e Instalación	1 Semana
	Modelado Dimensional	3 Semanas
	Diseño Físico	1 Semana
	Diseño y Desarrollo de procesos de extracción, transformación y carga de datos.	2 Semanas
Desarrollar 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
	Desarrollo de Aplicaciones Analíticas	2 Semanas
Realizar pruebas de Calidad de los Datos	Implementación	3 Semanas

TABLA 3: ACTIVIDADES A REALIZAR PARA IMPLEMENTAR LA SOLUCIÓN DE INTELIGENCIA DE NEGOCIOS

La fase de Gestión del Proyecto es una fase transversal en todo el desarrollo del proyecto, por lo cual se mantiene desde el inicio hasta el final mediante las correcciones y la verificación de cumplimiento de las actividades

#### **4.1.2. Definición de los Requerimientos del Negocio**

La solución de inteligencia de negocio a implementar, pretende ser útil para las Entidades Político-Territoriales en particular a la Dirección de Seguimientos y Políticas públicas del Fondo de Compensación Interterritorial y para aquellas instituciones que realicen control y seguimiento sobre los proyectos de obra civil que éstas planifican, con el fin de tener una herramienta que les facilite la toma de decisiones. Por ello, se han realizado diversas investigaciones con el objetivo de conocer cuál es el proceso de planificación, ejecución, administración y gestión de los proyectos de obra civil a través de algunas entidades e instituciones expertas en el área y conocer cuál es el déficit que tienen a la hora de poder visualizar los datos para el análisis de sus operaciones; obteniendo como resultado la necesidad de conocer:

- Un valor numérico que refleje la planificación, ejecución y gestión de un proyecto de obra civil planificado por una Entidad Político-Territorial.
- Conocer en porcentaje del avance físico de la ejecución de un proyecto de obra civil.
- Conocer en porcentaje el avance financiero (monto gastado) de un proyecto de obra civil.
- Promediar los valores numéricos y porcentajes mencionados anteriormente y agruparlos por Entidad Político-Territorial, por Estados, y por Región.
- Conocer la cantidad de proyectos de obra civil según su estatus de avance físico y agruparlos por Entidad Político-Territorial, por Estados, y por Región.
- Conocer la duración en días de la ejecución del proyecto de obra civil según su planificación, así como la duración real de la ejecución del mismo, y los días de retraso en el inicio de la ejecución del proyecto en caso dado.
- Mostrar los datos básicos del proyecto de obra civil como el nombre, la descripción, la categoría, el monto del proyecto, estatus de avance físico, estatus de rendición, entre otros.

Es importante señalar que se tomaran en cuenta solo aquellos proyectos cuyos estatus de carga del proyecto sea “aprobado” o “culminado”. Además, si un proyecto tiene como estatus

---

de avance físico “sin iniciar” y la fecha en curso está dentro de la banda de protección (más o menos el 20% de la duración teórica del proyecto sumado a la fecha de inicio del proyecto) entonces, ese proyecto aún no debe ser medido. Por lo tanto, los indicadores solo serán aplicados a los proyectos que no cumplan con la condición antes mencionada.

Partiendo de lo anterior, se plantea que esta solución permita a las entidades e instituciones antes mencionadas, mediante la definición de indicadores y variables cuantitativas, obtener de forma rápida y precisa información respecto a la planificación, ejecución y gestión de los proyectos de obra civil con la finalidad de dar apoyo a la toma de decisiones.

Como se puede observar en la tabla 4, algunas de las variables provenientes del sistema transaccional utilizadas para obtener los indicadores son:

Variable	Siglas
Fecha de Inicio del proyecto estipulada en la planificación.	TI
Fecha de Finalización del proyecto estipulada en la planificación.	TF
Monto del proyecto estipulada en la planificación.	MP
Monto total anual asignado a la Entidad Político-Territorial	ME
Fecha de Inicio real del proyecto según el acta de inicio, y disponible al momento de realizar una actualización en la rendición de cuenta.	TIR
Fecha de Finalización real del proyecto según el acta de culminación, y disponible al momento de rendir el proyecto como culminado.	TFR
Avance físico de la ejecución del proyecto hasta el momento de realizar una actualización rendición.	AFisc
Monto del proyecto utilizado hasta el momento de realizar una actualización en la rendición.	Afinan
Fecha de la última actualización de la rendición de cuentas	FR
Estatus de avance físico	EAF
Estatus de la rendición	ER

TABLA 4: VARIABLES PROVENIENTES DEL SISTEMA TRANSACCIONAL

Ahora, conociendo la fecha de inicio y la fecha de finalización del proyecto según fue planificado, así el porcentaje de avance físico de la ejecución del proyecto, la fecha de inicio real y la fecha de finalización real del proyecto según su rendición de cuentas, se puede observar, el comportamiento teórico y real del proyecto.

Usando lo mencionado anteriormente como premisa, algunos de los indicadores implementados (Pérez, 2015) y procedimientos de cálculos son:

- **Promedio de Avance Financiero (PAFinan):** Es el porcentaje del monto que ha sido utilizado en la ejecución del proyecto hasta un momento determinado y tendrá valores entre 0% y 100%.

Formula: 
$$PAFinan = \frac{(AFinan*100)}{MP}$$

- **Promedio de Avance de la Ejecución Física (PAFisi):** Es el porcentaje de la ejecución física que lleva el proyecto en un momento determinado y tendrá valores entre 0% y 100%.

Este dato es suministrado en la rendición de cuentas.

- **Ponderación por Monto (PM):** Es un valor numérico que se asigna al proyecto tomando en cuenta el monto del proyecto respecto al monto total que posea la Entidad Político-Territorial para la ejecución de proyectos de Obra Civil, esta ponderación tendrá valores entre 0 y 1 siendo 1 el valor óptimo.

Formula: 
$$(PM) = \frac{(MP - 0,2 * ME)^2}{(0,2 * ME)^2}$$

- **Ponderación por Categoría (PC):** Es un valor numérico que se asigna al proyecto tomando en cuenta la percepción de la población sobre la prioridad que tiene esa categoría del proyecto, y tendrá valores entre 0 y 1 siendo 1 el valor óptimo.

Este dato es suministrado por el Sistema Transaccional.

---

- **Duración del Proyecto (DP):** Es el número de días que durara la ejecución del proyecto según la fecha de inicio y fecha de finalización planificada.

Formula: 
$$DP = TF - TR$$

- **Duración del Proyecto Real (DPR):** Es el número de días que duró la ejecución del proyecto según la fecha de inicio estipulada en el acta de inicio y la fecha de finalización estipulada en acta de finalización.

Formula: 
$$DPR = TFR - TIR$$

- **Desviación en inicio ( $\Delta T$ ):** Es el número de días que se retrasó el inicio de la ejecución del proyecto.

Formula:

- Si ya realizo una rendición de cuentas y cargo la fecha de inicio según acta de inicio:

$$\Delta T = TIR - TI$$

- Si no ha realizado ninguna rendición de cuenta o ya realizo una rendición de cuentas, pero no ha cargado la fecha de inicio según el acta de inicio, y la fecha en curso es mayor a la banda de protección (más o menos el 20% de la duración teórica del proyecto sumado a la fecha de inicio del proyecto) \*:

$$\Delta T = Fecha\ en\ Curso - TI$$

- Si no ha realizado ninguna rendición de cuenta, y se encuentra dentro de la banda de seguridad:

$$\Delta T = 0$$

\* Banda de Protección: 
$$BPN = TI \pm (TF - TI * 0,2)$$

- **Angulo de la Recta Teórica ( $\theta T$ ):** Es el ángulo que forma la recta teórica de la ejecución del proyecto.

Formula: 
$$\theta T = atan\left(\frac{1}{TF-TI}\right)$$

- **Angulo de la Recta Real ( $\theta R$ ):** Es el ángulo que forma la recta real de la ejecución del proyecto tomando en cuenta del avance físico registrado en la última fecha de rendición.

Formula:

- Si se ha rendido cuentas, y la fecha de la última actualización de esta rendición es menor a la fecha de inicio del proyecto y el avance físico es mayor a 0%:

$$\theta R = atan\left(\frac{TI - TR}{AFisc}\right) + \pi/2$$

- Si no se cumple la condición anterior:

$$\theta R = atan\left(\frac{AFisc}{TR-TI}\right)$$

- **Índice de Desviación (Id):** Es la relación angular existente entre el ángulo de la recta teórica y el ángulo de la recta real, el cual permite cuantificar el valor del retraso del inicio del proyecto. Este índice tendrá un valor entre 0 y 1, siendo 1 el valor ideal.

Formula:

- Si el ángulo de la recta real es menor o igual al ángulo de la recta teórica:

$$Id = \theta R / \theta T$$

- Si el ángulo de la recta teórica es menor o igual al ángulo de la recta real:

$$Id = \theta T / \theta R$$

- **Índice de Avance Real (Iavr):** Se toma en cuenta la Ponderación por Monto, la Ponderación por Categoría, el avance físico y el índice de desviación con el propósito de

ver el comportamiento de la ejecución física del proyecto en el tiempo. Este índice tendrá un valor entre 0 y 1, siendo 1 el valor ideal.

Formula: 
$$I_{avr} = \frac{1.5 * AFisc * PM + 0.5 * AFisc * PC + Id}{3}$$

- **Índice Físico Teórico (Iavt):** Representa el valor ideal de avance en el tiempo del proyecto tomando en cuenta la Ponderación por Monto, la Ponderación por Categoría, el avance físico el cual debe ser igual al planificado para el tiempo de la última rendición de cuenta y con un índice de desviación igual a 1.

Formula: 
$$I_{avt} = \frac{1.5 * AFT * PM + 0.5 * AFT * PC + 1}{3}$$

- **Índice de Ejecución (Iej):** Refleja cómo ha sido la evolución y rendimiento en el tiempo del desarrollo y ejecución del proyecto, sin tener dependencia de su categoría o monto, ya que compara el desarrollo del mismo, en un tiempo determinado, con su valor ideal en ese instante. Este Índice tendrá un valor entre 0 y 100, siendo 100 el valor ideal.

Formula: 
$$I_{ej} = \left( \frac{i_{avr}}{i_{avt}} \right) * 100$$

- **Índice de Planificación (Ip):** Muestra cuan ajustado esta un proyecto a las necesidades de la población y su nivel estratégico tomando en cuenta la ponderación por categoría (PC) y la ponderación por Monto (PM). Este Índice tendrá un valor entre 0 y 100, siendo 100 el valor ideal.

Formula:

- Si el proyecto es declarado como culminado pero su avance físico fue menor al 50%:

$$I_p = \frac{(1.5 * PM + 0.5 * PC)}{2} * 100 * I_{ej}$$

- En caso de que desviación en tiempo (en adelantamiento) sea mayor a 40% del intervalo de duración del proyecto:

$$I_p = \frac{(1.5 * PM + 0.5 * PC)}{2} * 100 * Id$$

- En caso de que no se cumplan ninguna de las condiciones antes mencionadas:

$$I_p = \frac{(1.5 * PM + 0.5 * PC)}{2} * 100$$

- **Índice de Gestión:** Tiene como propósito cuantificar los resultados obtenidos en el desempeño del proyecto tomando en cuenta que se haya realizado una buena planificación y una buena ejecución.

Formula: 
$$I_g = \frac{0.8 * I_p + 1.2 * I_ej}{2}$$

Por lo tanto, una vez establecidos los indicadores a implementar para cumplir con los requerimientos establecidos, se procede a realizar el desarrollo de la aplicación.

#### 4.1.3. Diseño Técnico de la Arquitectura

El diseño de la arquitectura de la solución de inteligencia de negocio que se desarrolla se puede observar en la ilustración 30:

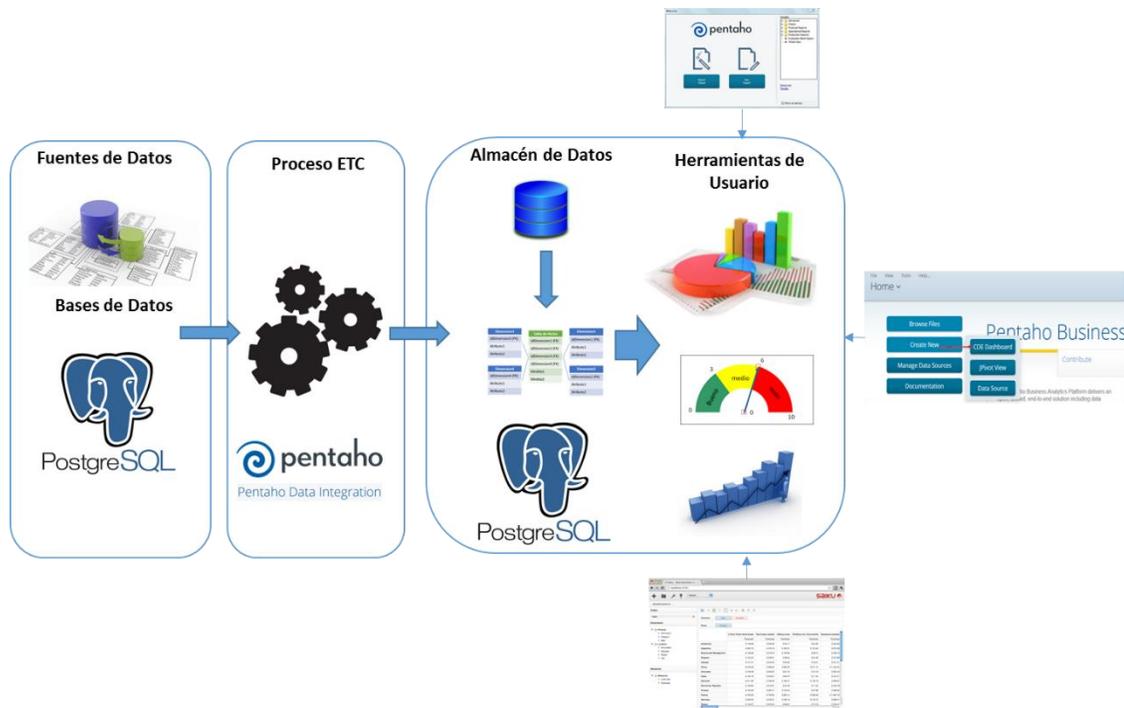


ILUSTRACIÓN 30: ARQUITECTURA DE LA SOLUCIÓN DE INTELIGENCIA DE NEGOCIOS

Como se puede observar en la imagen, la arquitectura está conformada por tres (3) ambientes y diversos procesos y elementos que permiten llevar a cabo el flujo de trabajo de la misma.

El primer ambiente corresponde a las fuentes de datos que se consideran para la solución de inteligencia de negocio. Estas fuentes provienen del sistema transaccional relacionado con los procesos de negocio previamente estudiados, cuyos datos se almacenan en una base de datos Postgresql.

El segundo ambiente comprende los procesos de extracción, transformación y carga (ETL) que se ejecutarán periódicamente. Mediante estos procesos, se comienza a poblar el almacén con los datos obtenidos de las fuentes de datos previamente procesados y transformados. En este ambiente se utiliza la herramienta de Pentaho Data Integration (PDI) 6.1 para crear los procesos ETL.

Por último, está el tercer ambiente, el cual se encuentra conformado por dos componentes:

- El primer componente es el almacén de datos, donde se localizan los datos históricos, que son la integración de las diversas fuentes de datos; para así acelerar el proceso de análisis que se puede realizar a través de las herramientas analíticas, ya que el almacén está estructurado en base a un modelo dimensional. Dicho almacén será periódicamente poblado por los procesos ETL del ambiente anterior. Para este componente se hace uso de la Base de Datos Postgresql 9.4 en conjunto con la herramienta pgAdmin III para crear con sentencias SQL, el ambiente y las estructuras necesarias para construir el almacén.
- El segundo componente se refiere a la capa de presentación, es decir, a las herramientas de acceso a los datos dispuestos en el almacén para permitirle a los usuarios finales visualizar y analizar los distintos indicadores y reportes contemplados para esta solución de inteligencia de negocio. En este componente se recurre al conjunto de herramientas de la suite Pentaho Business Intelligence en su versión Pentaho Community, utilizando la herramienta Pentaho BI Server para el desarrollo del tablero de control, la herramienta Pentaho Report Designer para generar los reportes, y por último la herramienta Pentaho Schema Workbench para crear el esquema de datos y el cubo de información Mondrian que utilizaran las herramientas mencionadas anteriormente.

#### **4.1.4. Selección de Productos e Instalación**

El proceso de selección de una herramienta tecnológica para el desarrollo e implementación de una solución de inteligencia de negocio, puede resultar complejo debido a la diversidad de herramientas que existen en el mercado y a la variada gama de funcionalidades que cada una de ellas ofrece hoy en día, siendo Pentaho Business Intelligence, Tableau Software, Oracle Business Intelligence (OBI), IBM Cognos, SpagoBI y Microsoft Power Business Intelligence algunas de las herramientas más utilizadas en la actualidad.

Sin embargo, independientemente de la herramienta a seleccionar, ésta debe proveer un repositorio centralizado, visualización de un tablero de control y la posibilidad de construir

consultas a la medida de acuerdo a los requerimientos de información que poseen las instituciones.

Luego de estudiar las distintas herramientas presentes en el mercado, se optó utilizar para el desarrollo de la solución de inteligencia de negocio planteada, las herramientas de la suite Pentaho Business Intelligence Community y la base de datos Postgresql como se observa en la tabla 5:

Herramienta	Uso	Requisitos de Instalación
Pentaho Data Integration 6.1	Creación y Administración de Procesos ETL	<ul style="list-style-type: none"> <li>• <b>Sistema Operativo:</b> Windows o Linux</li> <li>• <b>Memoria RAM:</b> min 1 GB, recomendable &gt; 2 GB</li> <li>• <b>Memoria en disco duro:</b> 200 MB</li> </ul>
Pentaho BI Server 6.1	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.	<ul style="list-style-type: none"> <li>• <b>Sistema Operativo:</b> Windows o Linux</li> <li>• <b>Memoria RAM:</b> min 1 GB, recomendable &gt; 2 GB</li> <li>• <b>Memoria en disco duro:</b> 500 MB</li> </ul>
Pentaho Report Designer 6.1	Creación y Administración de Reportes, se enlaza con Pentaho BI Server para hacer uso de los esquemas de datos allí alojados.	<ul style="list-style-type: none"> <li>• <b>Sistema Operativo:</b> Windows o Linux</li> <li>• <b>Memoria RAM:</b> min 2 GB, recomendable &gt; 4 GB</li> <li>• <b>Memoria en disco duro:</b> 500 MB</li> </ul>
Pentaho Schema Workbench 6.1	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.	<ul style="list-style-type: none"> <li>• <b>Sistema Operativo:</b> Windows o Linux</li> <li>• <b>Memoria RAM:</b> min 1 GB, recomendable &gt; 2 GB</li> <li>• <b>Memoria en disco duro:</b> 200 MB</li> </ul>
Postgresql 9.4	Creación y administración de bases de datos.	<ul style="list-style-type: none"> <li>• <b>Sistema Operativo:</b> Windows o Linux</li> <li>• <b>Memoria RAM:</b> min 1 GB, recomendable &gt; 2 GB</li> <li>• <b>Memoria en disco duro:</b> &gt; 40 MB</li> </ul>

TABLA 5: HERRAMIENTAS PARA DESARROLLAR LA SOLUCIÓN DE INTELIGENCIA DE NEGOCIO PROPUESTA

Destacando que, estas herramientas son seleccionadas con la finalidad de obtener las fortalezas de cada una de ellas y a su vez, realizar un experimento en cuanto al desarrollo de una solución de inteligencia de negocio usando únicamente software.

#### **4.1.5. Modelado Dimensional**

En base a los requerimientos del negocio planteados 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 a los proyectos de Obra Civil planificados por Entidades Político-Territoriales.

Sin embargo, para poder desarrollar en esta fase el diseño del modelo dimensional hay que seguir un conjunto de actividades propuestas por Ralph Kimball, siendo éstas las siguientes:

##### **4.1.5.1. Definir el Proceso de Negocio**

Se modela el proceso de Control y Seguimiento a los proyectos de Obra Civil planificados por Entidades Político-Territoriales a través de indicadores.

##### **4.1.5.2. Identificar el Nivel de Granularidad**

Para determinar el nivel de granularidad, es necesario especificar cómo se representaría una fila en la tabla de hechos, porque con ello, se define el nivel de detalle de la información que se desea almacenar del proceso a modelar, lo cual permite identificar fácilmente, lo que se espera medir con cada indicador propuesto. Partiendo de este hecho, se procede a establecer el nivel de granularidad:

- Monto del proyecto, avance financiero, porcentaje de avance físico, porcentaje de avance financiero, duración del proyecto según lo planificado, duración del proyecto real,

desviación en inicio, índice de avance real y teórico, índice de desviación, índice de planificación, índice de ejecución, índice de gestión de un proyecto que posee una categoría, un estatus de avance físico, un estatus de rendición, una fecha de inicio y fecha de finalización planificada, y una fecha de inicio y fecha de finalización real perteneciente a una Entidad Político-Territorial.

- Monto del proyecto, avance financiero, porcentaje de avance físico, porcentaje de avance financiero, duración del proyecto según lo planificado, duración real del proyecto, desviación en inicio, índice de avance real y teórico, índice de desviación, índice de planificación, índice de ejecución, índice de gestión de los proyectos que poseen una categoría, un estatus de avance físico, un estatus de rendición, una fecha de inicio y fecha de finalización planificada, y una fecha de inicio y fecha de finalización real pertenecientes a un Municipio.
- Monto del proyecto, avance financiero, porcentaje de avance físico, porcentaje de avance financiero, duración del proyecto según lo planificado, duración real del proyecto, desviación en inicio, índice de avance real y teórico, índice de desviación, índice de planificación, índice de ejecución, índice de gestión de los proyectos que poseen una categoría, un estatus de avance físico, un estatus de rendición, una fecha de inicio y fecha de finalización planificada, y una fecha de inicio y fecha de finalización real pertenecientes a un Estado.
- Monto del proyecto, avance financiero, porcentaje de avance físico, porcentaje de avance financiero, duración del proyecto según lo planificado, duración real del proyecto, desviación en inicio, índice de avance real y teórico, índice de desviación, índice de planificación, índice de ejecución, índice de gestión de los proyectos que poseen una categoría, un estatus de avance físico, un estatus de rendición, una fecha de inicio y fecha de finalización planificada, y una fecha de inicio y fecha de finalización real pertenecientes a una región REDI<sup>3</sup>.

---

<sup>3</sup> **Región Estratégica de Desarrollo Integral (REDI):** son instancias con autoridades que facilitan la implementación de las políticas públicas nacionales a lo largo de todo el territorio nacional. Establecen mecanismos de coordinación con el Gobierno Central, los gobiernos regionales y locales, permitiendo una mayor eficiencia, seguimiento y control en la ejecución de estas

---

### 4.1.5.3. Definir las Dimensiones

De acuerdo con los requerimientos de información recolectados y el análisis del sistema transaccional, en la fase de diseño, se identificaron las siguientes dimensiones:

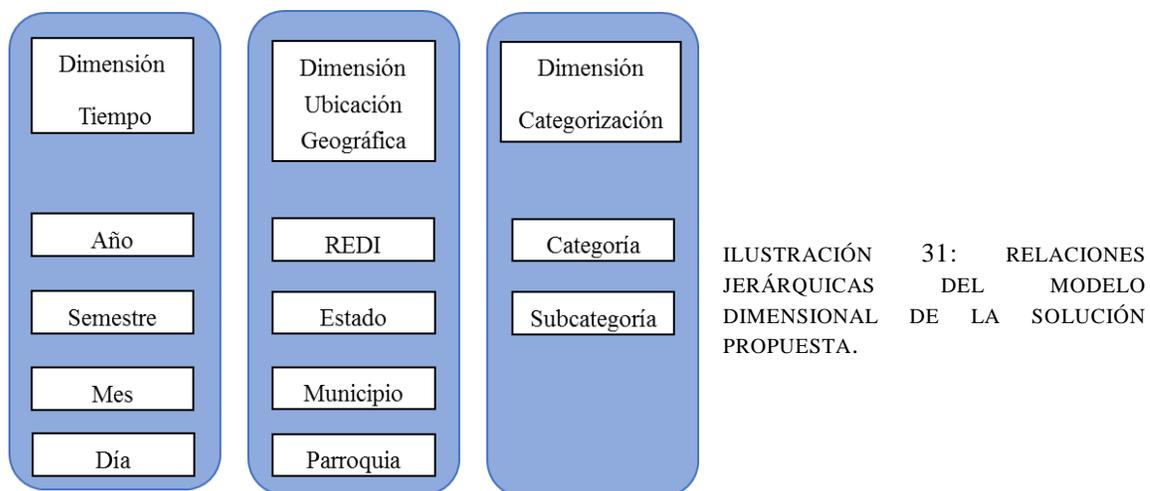
Dimensión	Nombre en el almacén de datos	Descripción
<b>Tiempo</b>	<b>dim_tiempo</b>	En esta tabla se guardan todos los datos referentes al tiempo (año, semestre, mes, día).
<b>Ubicación Geográfica</b>	<b>dim_ubicacion_geo</b>	Esta tabla contiene información sobre los municipios que pertenecen a un estado dentro de la región de un país.
<b>Categorización</b>	<b>dim_categorizacion</b>	Almacena información sobre las categorías y subcategorías que posee un proyecto.
<b>Dificultad</b>	<b>dim_dificultad</b>	Contiene las principales dificultades que se presentan durante la ejecución de un proyecto
<b>Entidad</b>	<b>dim_entidad</b>	Posee la información de las Entidades Político-Territoriales, su Registro Único de Información Fiscal (RIF) <sup>4</sup> y el Responsable Legal de esa entidad.
<b>Proyectos</b>	<b>dim_proyectos</b>	Esta tabla contiene los datos básicos de los proyectos de obra civil, como su código, nombre y descripción.
<b>Estatus de avance Físico</b>	<b>dim_estatus_af</b>	En esta tabla se encuentra los diferentes estatus de avance físico de los proyectos.
<b>Estatus de Rendición</b>	<b>dim_estatus_rnd</b>	En esta tabla se encuentra los diferentes estatus de avance físico de los proyectos.
<b>Estatus de carga del proyecto</b>	<b>dim_estatus_cp</b>	En esta tabla se encuentra los diferentes estatus que puede tener un proyecto cuando es registrado en el sistema transaccional.

TABLA 6: DIMENSIONES DEL MODELO DIMENSIONAL DE LA SOLUCIÓN PROPUESTA.

Además, se estudiaron e identificaron las relaciones jerárquicas que surgen entre los atributos de las diversas dimensiones según la lógica de negocio, obteniendo como resultado, como se

<sup>4</sup> **Registro Único de Información Fiscal (RIF):** es un registro destinado al control tributario de los impuestos, tasas y contribuciones administrados por el Servicio Nacional Integrado de Administración Aduanera y Tributaria (SENIAT)

puede apreciar en la ilustración 31, para la dimensión tiempo, la dimensión ubicación geográfica, y la dimensión categorización, las siguientes jerarquías:



Igualmente se identificó que la dimensión tiempo juegan diferentes roles (role-playing) dentro del modelo dimensional ya que esta dimensión está relacionada con todas las fechas mencionadas anteriormente como se muestra en la ilustración 32:

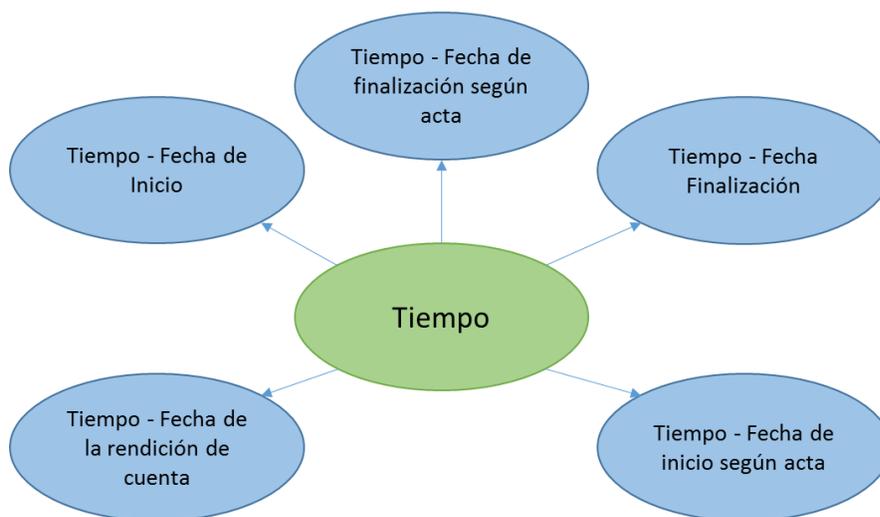


ILUSTRACIÓN 32: DIMENSIÓN ROLE-PLAYING TIEMPO

#### 4.1.5.4. Identificación de los Hechos y las Tablas de Hechos

Una vez definidas todas las dimensiones, se generó la tabla de hechos:

- **th\_cys:** Contiene toda la información necesaria para generar los reportes y el tablero de control que permitirán la visualización de los indicadores de control y seguimiento a los proyectos de Obra Civil.

Los hechos que se toman en cuenta para monitorear este proceso como se muestran en la tabla 6, son los siguientes:

Hecho	Nombre en el almacén de datos
Monto del proyecto	monto_proyecto
Avance financiero	avance_financ
Porcentaje de avance físico	porc_avance_financ
Porcentaje de avance financiero	porc_avance_fisico
Duración del proyecto según lo planificado	duración_dias
Duración real del proyecto	duración_dias_real
Desviación en inicio	desviacion_inicio
Índice de avance real	Índice_avt
Índice de avance teórico	índice_avr
Índice de desviación	índice_desviacion
Índice de planificación	iej
Índice de ejecución	ip
Índice de gestión	Ig

TABLA 7: HECHOS UTILIZADOS EN LA SOLUCIÓN PROPUESTA.

En la ilustración 33, se puede observar gráficamente la estructura que conforma el modelo dimensional.

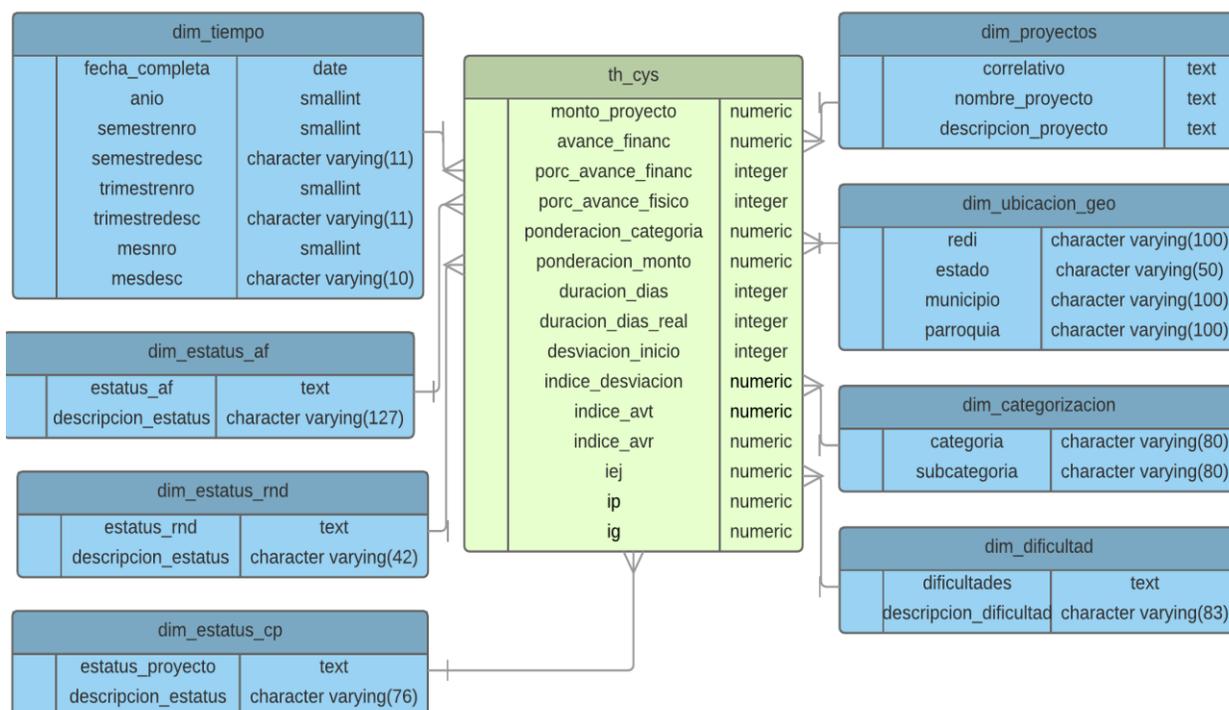


ILUSTRACIÓN 33: MODELO DIMENSIONAL DE LA SOLUCIÓN PROPUESTA.

#### 4.1.6. Diseño Físico

Una vez definido el modelo dimensional, utilizando Postgresql y la herramienta pgAdmin III provista por dicha base de datos, se procede a la construcción de las estructuras físicas (Dimensiones y tabla de hechos) del almacén de datos dentro de la base de datos por medio de sentencias SQL. Para ello se definen las siguientes propiedades de la base de datos:

**Servidor:** localhost

**Esquema:** public

**Puerto:** 4532

**Usuario:** Postgresql

**Nombre de la base de datos:** dw\_tesis

En la ilustración 34, se puede observar el modelo físico del almacén de datos ya implementado dentro de la base de datos.

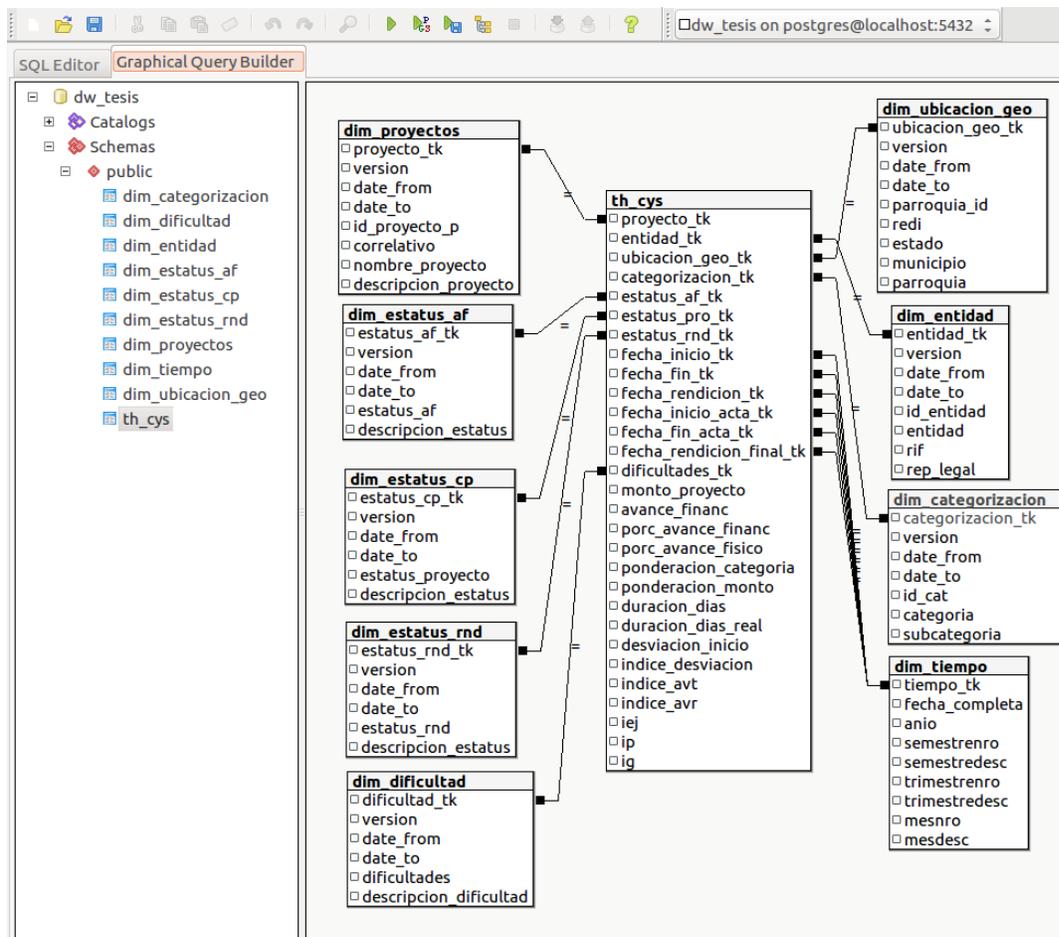


ILUSTRACIÓN 34: MODELO FÍSICO DEL ALMACÉN DE DATOS IMPLEMENTADO EN LA BASE DE DATOS.

#### 4.1.7. Diseño y Desarrollo de Procesos ETL

Después de crear en físico el almacén de datos en la base de datos Postgresql, se comienza a diseñar e implementar los procesos ETC (ETL) para extraer, transformar y cargar los datos procedentes de las fuentes de datos (suministradas el sistema transaccional que utilizan las Entidades Político-Territoriales) al almacén, para obtener fácilmente por medio de dicho almacén, información de manera precisa y oportuna para la toma de decisiones estratégicas en cuanto al control y seguimiento de proyectos de obra civil. Sin embargo, cabe señalar que los datos utilizados como fuentes de datos fueron suministrados a través de un respaldo del

sistema que se encuentra en producción, es decir, que los datos no están actualizados y toda la información que arrojará la aplicación es simulada.

Para el desarrollo de los procesos ETL, se utiliza la herramienta Pentaho Data Integration (PDI), como se mencionó en la fase de selección de productos. A través de esta herramienta se crean las transformaciones (*transformations*) que se conectan a las distintas fuentes de datos a las cuales se le realizan consultas para obtener los datos necesarios y procesarlos, con el fin de tomar el resultado de ese proceso e insertarlo o actualizarlo en las correspondientes fuentes de destino (tablas que constituyen al almacén de datos). Asociado a esto, también se crean trabajos (*Jobs*), los cuales permiten ejecutar un conjunto de transformaciones de forma automática sin necesidad de ejecutar una a una y se pueden crear condiciones sobre la ejecución de las mismas e incluso el envío de correo electrónico.

Las transformaciones y los trabajos se crean dentro de un repositorio de archivos, tal como se muestra en la ilustración 35. Este repositorio es un directorio específico que permite ubicar en un mismo lugar a través de carpetas, todas las transformaciones y los trabajos (*Jobs*) desarrollados.

Name	Type	User	Changed date	Description
/				
CyS	JOB	-	2016/10/11 15:	
dimension_categorizacion	TRANSFORMATION	-	2016/10/10 13:	
dimension_dificultades	TRANSFORMATION	-	2016/10/10 13:	
dimension_entidad	TRANSFORMATION	-	2016/10/11 15:	
dimension_proyectos	TRANSFORMATION	-	2016/10/10 13:	
dimension_status_avance_fisico	TRANSFORMATION	-	2016/10/10 13:	
dimension_status_proyecto	TRANSFORMATION	-	2016/10/10 13:	
dimension_status_rnd	TRANSFORMATION	-	2016/10/10 13:	
dimension_tiempo	TRANSFORMATION	-	2016/10/10 13:	
dimension_ubicacion_geografica	TRANSFORMATION	-	2016/10/10 13:	
envio_correo_neg	TRANSFORMATION	-	2016/10/04 09:	
envio_correo_pos	TRANSFORMATION	-	2016/10/04 09:	
tabla_de_hechos	TRANSFORMATION	-	2016/10/11 15:	

ILUSTRACIÓN 35: REPOSITORIO DE ARCHIVOS

Para cargar con datos las tablas que conforman al almacén de datos, se tienen transformaciones con una estructura similar a la reflejada en la ilustración 36:

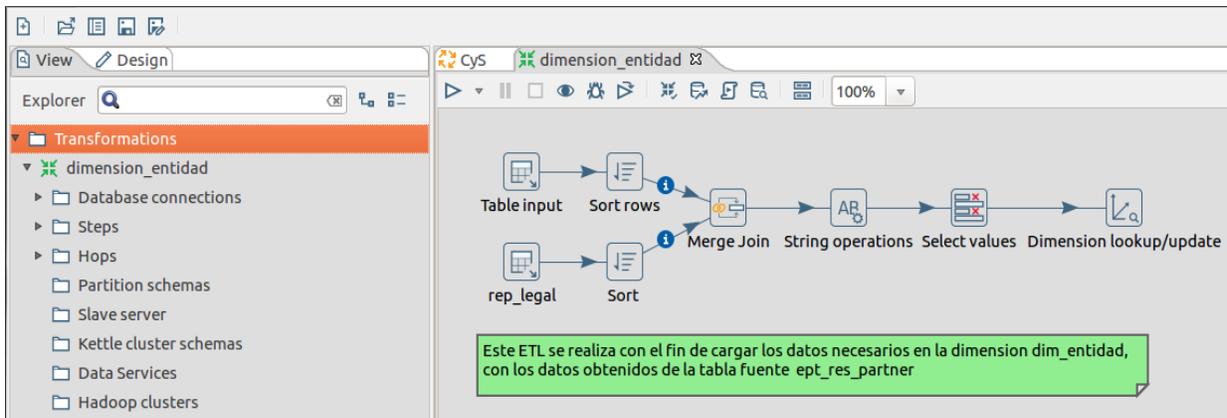


ILUSTRACIÓN 36: ETL PARA CARGAR LA DIMENSIÓN ENTIDAD

En la Ilustración #, se puede observar que, para cargar una dimensión en el almacén de datos, se realizan consultas a una o varias tablas fuente, que en este caso es la tabla fuente llamada *ept\_res\_partner* (tabla proveniente del sistema transaccional que posee la organización). Con los datos obtenidos como salida en dicha consulta, se procede a realizar una limpieza y transformación de ellos, con el fin de estandarizar los datos y obtener sólo los campos necesarios para formar un registro e insertarlo o actualizarlo a través de un paso llamado *dimension lookup* en una tabla destino, que, para este proceso, es la dimensión llamada *dim\_entidad*. Asimismo, todas las transformaciones que permiten cargar con datos el resto de las tablas que integran al almacén de datos poseen una estructura similar a la de esta transformación.

Aunado a lo mencionado anteriormente, como lo muestra la ilustración 37, se crea una transformación para poblar la tabla de hechos, esta transformación, aunque mantiene una estructura similar a la anterior, es un poco más compleja en el sentido en que se realizan consultas tanto al sistema transaccional como a las dimensiones ya previamente cargadas, y así poder obtener todos los registros en base a sus claves primarias. En este caso, es importante mencionar que en esta transformación se realizaron una serie de cálculos y operaciones matemáticas que generan los indicadores a implementar, estos indicadores van asociados al nivel más fino de granularidad como lo son los proyectos.

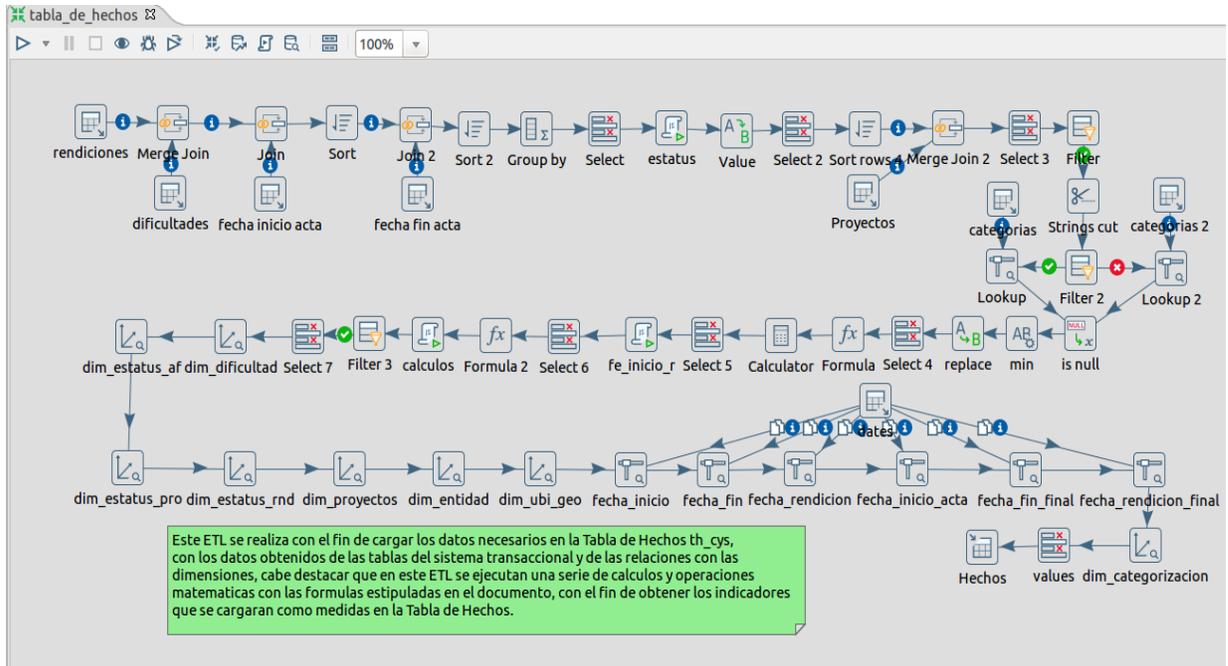


ILUSTRACIÓN 37: ETL PARA CARGAR LA TABLA DE HECHOS

Asociado a lo expuesto previamente, para no ejecutar de forma individual cada transformación, se crean uno o más trabajos (*Jobs*) para poblar el almacén. Como se mencionó en el ítem 2.6.1, 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 ilustración 38:

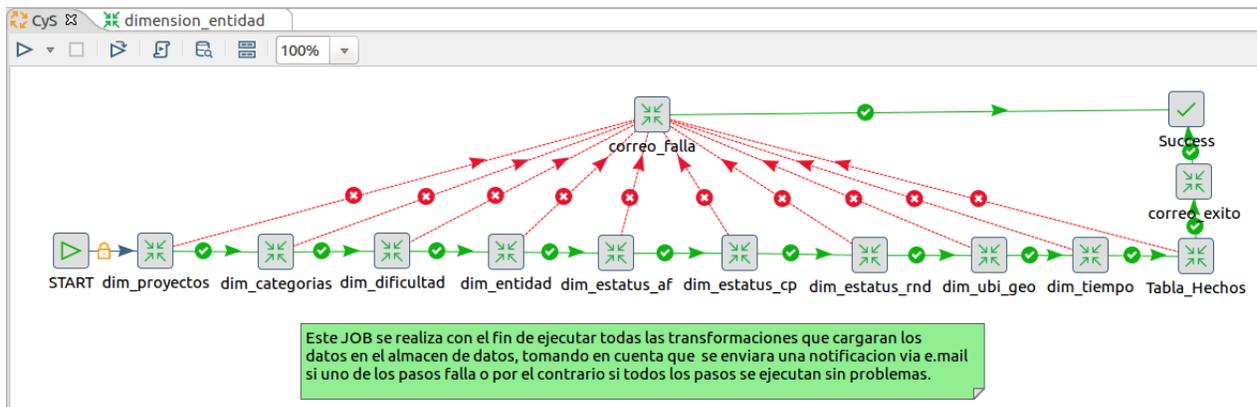


ILUSTRACIÓN 38: TRABAJO (JOB) QUE EJECUTA TODAS LAS TRANSFORMACIONES

En la anterior ilustración, se puede observar que el *job* comienza desde un punto de partida (*Start*) y va luego, se ejecutan todas las transformaciones que representan cada uno de los procesos ETL que cargan con datos las tablas del modelo dimensional previamente mencionado. Posteriormente, se eliminan todos los registros de las dimensiones y tablas de hechos cuya clave primaria es igual a 0, ya que al utilizar los *dimension lookup*, éstos generan una primera fila con todos los campos en nulo lo cual es innecesario; para continuar finalmente con el control de ejecución del ETL.

No obstante, hay que resaltar que, las flechas verdes siguen un flujo con ejecución exitoso, pero si ocurre algún error en alguna transformación, el flujo de ejecución se irá por el camino en rojo y no continuaran ejecutándose las demás transformaciones. Sin embargo, después del acontecimiento de un error, se le envía a la persona encargada de monitorear la ejecución del *job*, un correo indicando que fallo la ejecución del *job*, con el fin de que pueda solucionar el inconveniente ocurrido. En caso contrario, dicha persona recibirá también un correo, pero en éste será indicado que la ejecución del *job* fue exitosa.

Por último, es importante destacar que para poder realizar la ejecución periódica del *job*, se crea un archivo *.sh* para realizar el procesamiento por lotes de los datos mediante comandos linux, y luego una tarea programada que ejecuta ese archivo *.sh* todos los días a las 4:00 a.m.

#### **4.1.8. Especificaciones de las Aplicaciones Analíticas**

Los requerimientos de los usuarios vienen dados por la necesidad de utilizar herramientas que permitan fácilmente obtener información, esto implica crear aplicaciones que sean de fácil uso, eficientes, eficaces, flexibles entre otros aspectos. Por lo tanto, se planteó el conjunto inicial de reportes como se muestra en la ilustración 39, y la estructura del tablero de control a utilizar.

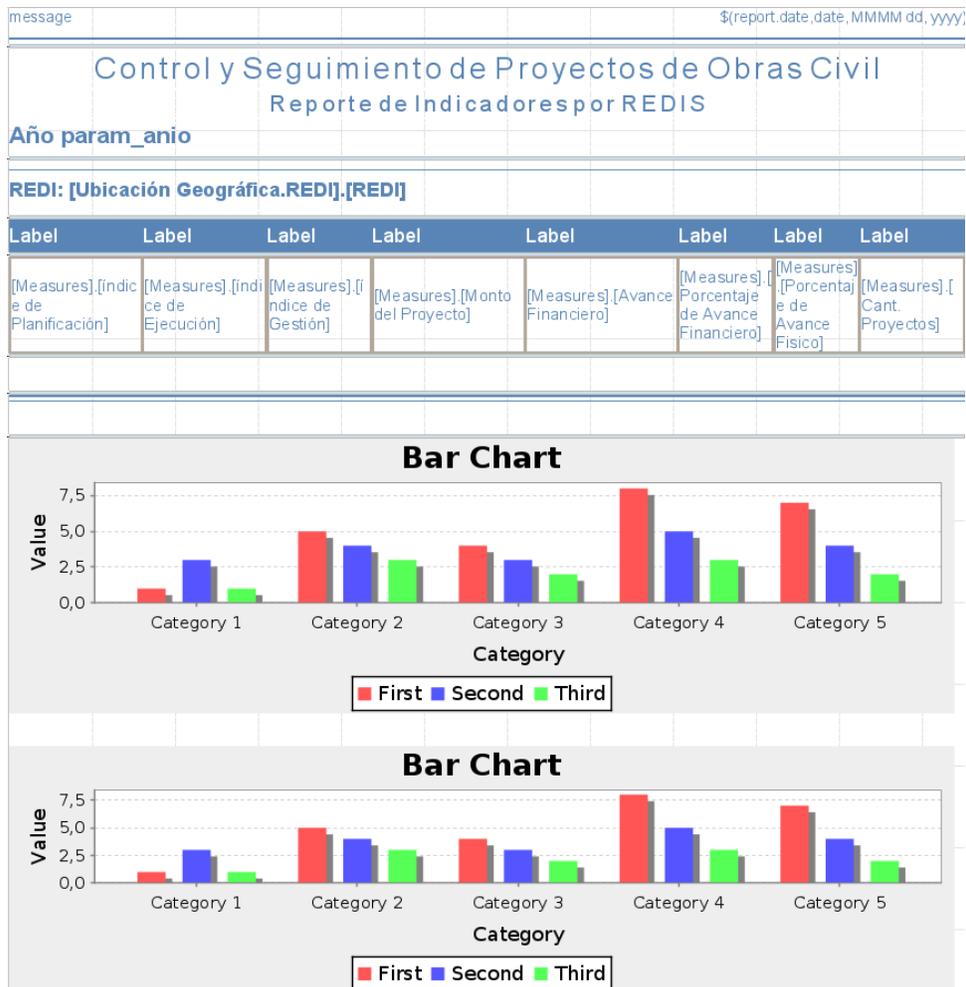


ILUSTRACIÓN 39: PLANTILLA ESTÁNDAR PARA LA CREACIÓN DE REPORTES

Además, tomando en consideración lo expuesto anteriormente, se decidió que un portal web es el medio por el cual el usuario de la aplicación puede interactuar con el tablero de control, con los reportes y con el cubo de información Mondrian que contiene los indicadores desarrollados en esta solución de inteligencia de negocio. Por tanto, para que la aplicación sea segura permitiéndole acceder sólo a los usuarios autorizados y la interfaz gráfica del portal, sea agradable, intuitiva y de fácil entendimiento para todos esos usuarios, se define un conjunto de pautas y estándares, siendo éstos, los descritos a continuación:

- **Autenticación:** Todos los usuarios que hacen uso de la aplicación deben autenticarse al momento de acceder al tablero de control, a los reportes o al cubo de información mediante el nombre de usuario y contraseña asignados.

- **Diseño Simple:** El portal web, el tablero de control y los reportes se muestran sin sobrecarga de información agrupando los indicadores por ubicación geográfica.
- **Vistas de los Reportes e Indicadores:** La representación de los reportes e indicadores está dada por vistas de tablas combinadas con elementos gráficos, tales como: gráficos de barras, gráficos de líneas, entre otros.
- **Refinamiento de Consultas:** Cada reporte y el tablero de control posee un conjunto de filtros, los cuales le facilita al usuario obtener la información que requiere analizar a través de los reportes e indicadores creados.
- **Formato Estándar:** La aplicación posee un formato que permite lograr homogeneidad en la representación de los indicadores y reportes mostrados. Dichos reportes e indicadores contienen un encabezado compuesto por un título, un subtítulo y la fecha de consulta.
- **Portabilidad de Información:** La aplicación le permitirá al usuario exportar en diferentes formatos (PDF, hoja de cálculo, entre otros) la información que desee de los reportes y/o indicadores que consulta.

#### 4.1.9. Desarrollo de las Aplicaciones Analíticas

Por medio de las herramientas que ofrece Pentaho Community, específicamente Pentaho Schema Workbench, se da inicio al desarrollo de la aplicación analítica de la solución de inteligencia de negocio planteada, en donde como se observa en la ilustración 40, se genera el esquema de datos con toda la estructura y los datos del almacén de datos implementado anteriormente, que sirve como base para crear las consultas utilizadas en el desarrollo de los reportes y el tablero de control.

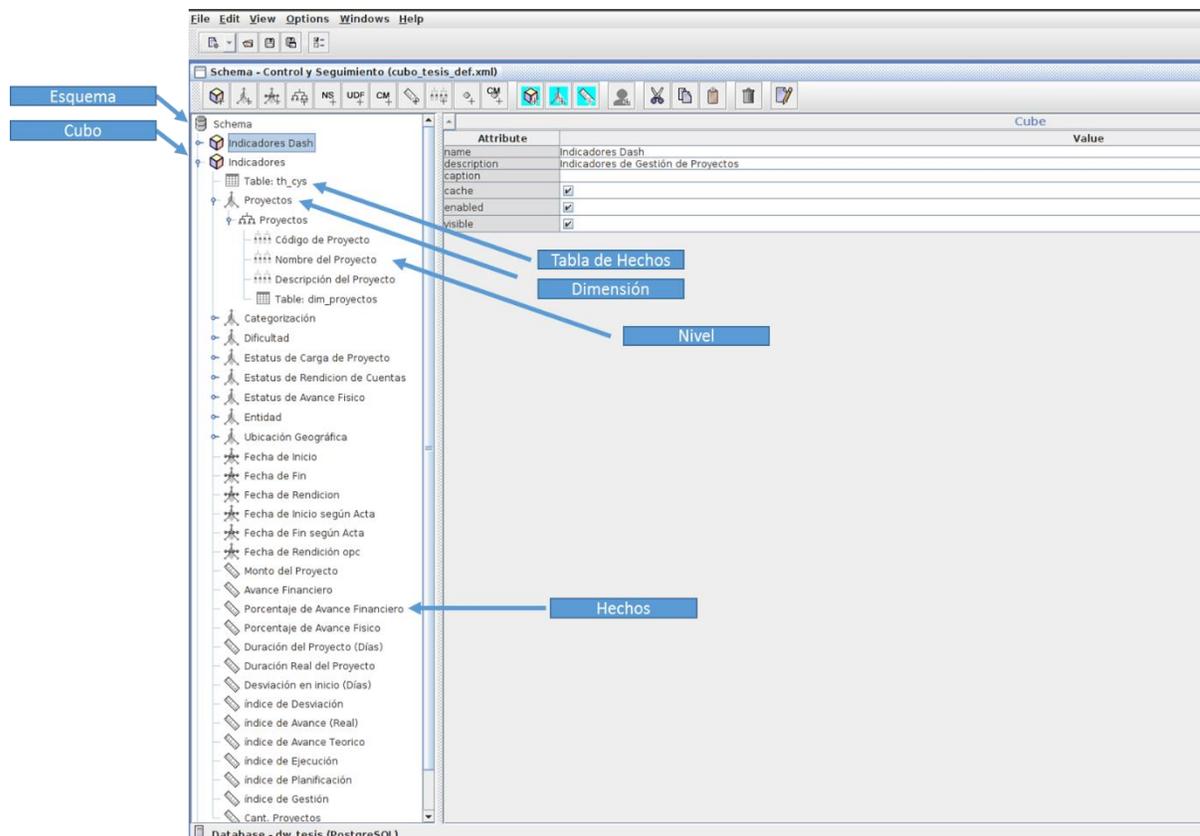


ILUSTRACIÓN 40: VISTA DEL ESQUEMA DE DATOS DESARROLLADO CON PENTHAO SCHEMA WORKBENCH

Como se muestra en la ilustración anterior, un esquema de datos está conformado por uno o más cubos de información, dentro de los cuales se selecciona la tabla de hechos con la que se desea trabajar y se definen las dimensiones con sus respectivas jerarquías, además se crean las funciones de agregación sobre los hechos contenidos en la tabla de hechos. Por último, como se muestra en la ilustración 41, ya realizado el esquema de datos, éste debe ser publicado en el servidor de Pentaho BI Server, con lo cual estará disponible para ser utilizado por las otras herramientas de construcción de reportes y del tablero de control. Cabe destacar que el archivo que se genera al crear el esquema de datos está estructurado en lenguaje MDX<sup>5</sup>.

<sup>5</sup> Las expresiones multidimensionales o MDX (acrónimo de MultiDimensional eXpressions) son un lenguaje de consulta para bases de datos multidimensionales OLAP.

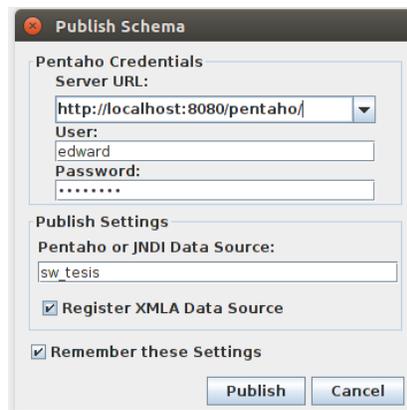


ILUSTRACIÓN 41: PUBLICACIÓN DE UN ESQUEMA DE DATOS EN EL SERVIDOR PENTAHO BI SERVER.

A continuación, una vez establecida el esquema de fuentes de datos se puede construir los reportes y el tablero de control; en el caso de los reportes, se hace uso de la herramienta Pentaho Report Designer en la cual se debe elegir el esquema de datos con el que se va a trabajar, y luego diseñar la estructura del reporte. Es importante señalar que, aunque se puede realizar consultas con sentencias SQL directamente sobre la base de datos del sistema transaccional o incluso sobre el almacén de datos para generar los reportes, lo recomendable es utilizar consultas con sentencias MDX sobre el esquema de datos ya que son consultas más sencillas y además optimiza la velocidad de consulta y se hace menos uso de recursos. En las ilustraciones 42 y 43, se puede observar el proceso de construcción de uno de los reportes a implementarse en la solución de inteligencia de negocios

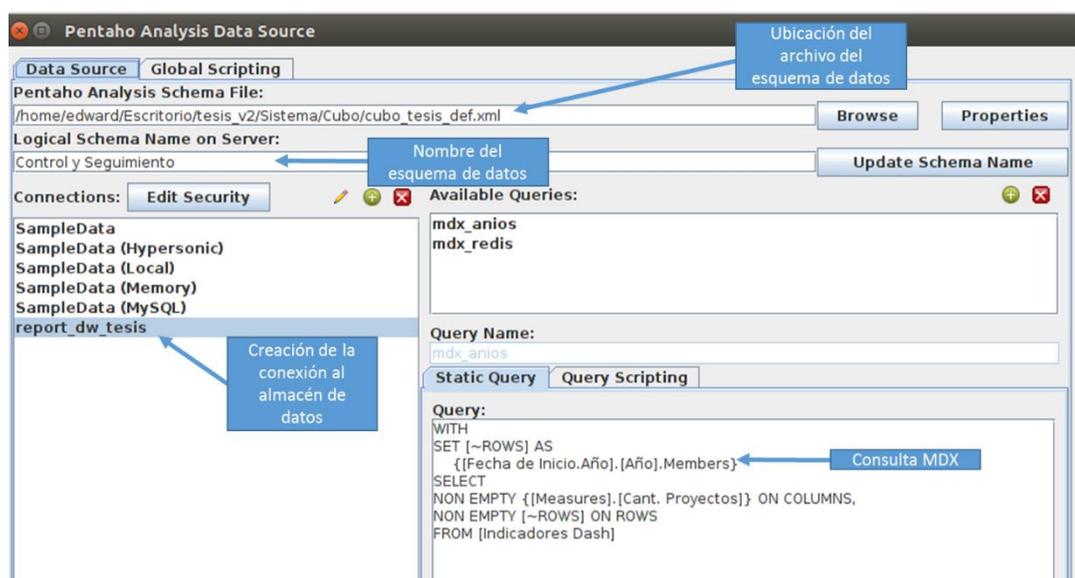


ILUSTRACIÓN 42: SELECCIÓN DEL ESQUEMA DE DATOS Y CREACIÓN DE LAS CONSULTAS MDX.

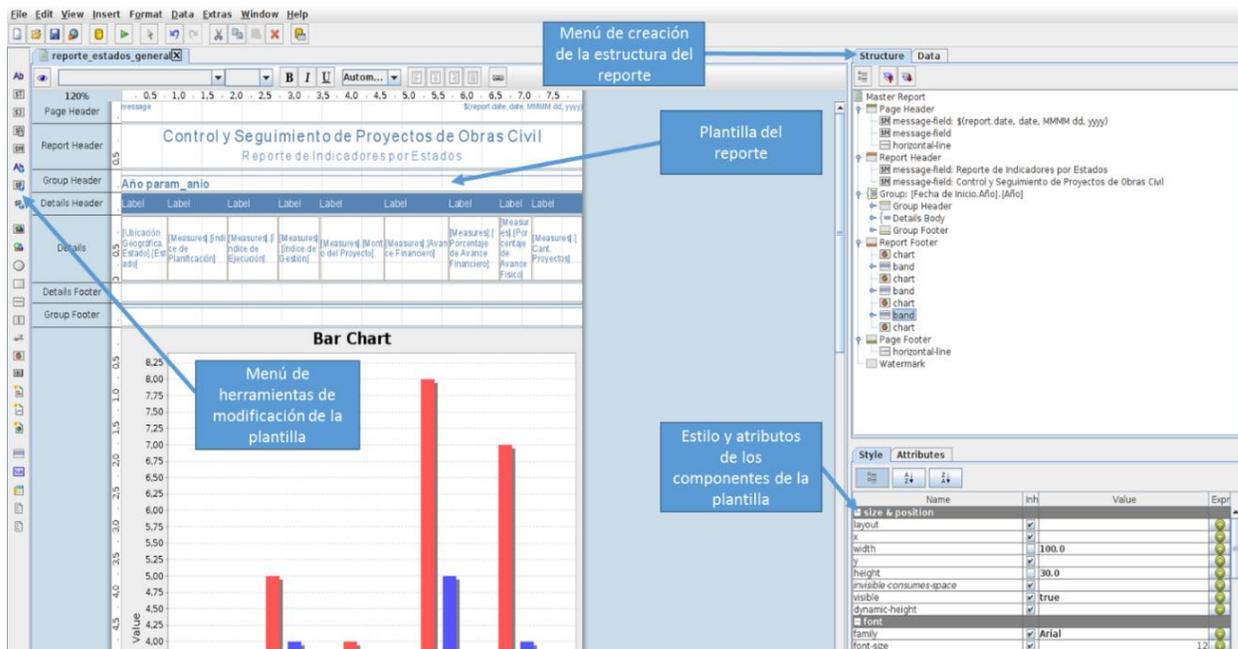


ILUSTRACIÓN 43: DISEÑO DE LA ESTRUCTURA DEL REPORTE

Posteriormente, se crea el tablero de control, a través de la herramienta Community Dashboard Editor (CDE) contenido dentro de Pentaho BI Server. Esta herramienta está compuesta por tres capas, la primera es Community Dashboard Framework (CDF) que permite la manipulación grafica del diseño web, la segunda es el Community Data Access (CDA) que se utiliza para crear las consultas para el acceso a los datos y la tercera es la Community Chart Components (CCC) que es el gestor de los componentes, gráficos, tablas y librerías implementas para crear el tablero de control.

Tomando en cuenta lo anterior, las ilustraciones 44, 45, 46, muestran la utilización de las tres capas para la creación del tablero de control propuesto en la solución de inteligencia de negocios.

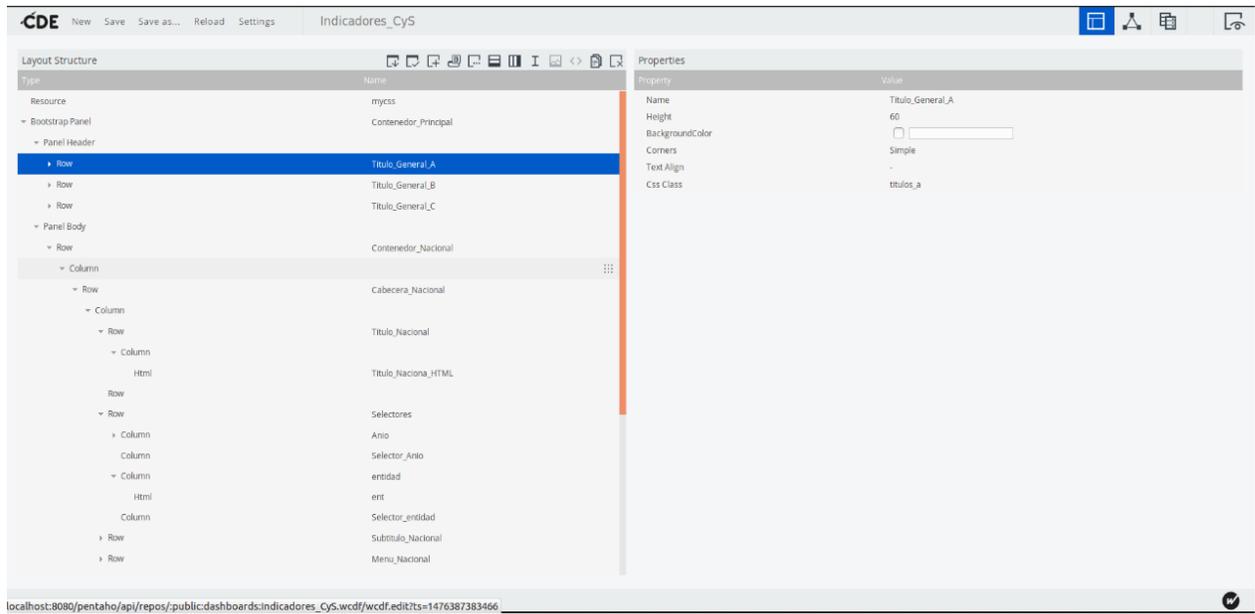


ILUSTRACIÓN 44: CAPA COMMUNITY DASHBOARD FRAMEWORK (CDF)

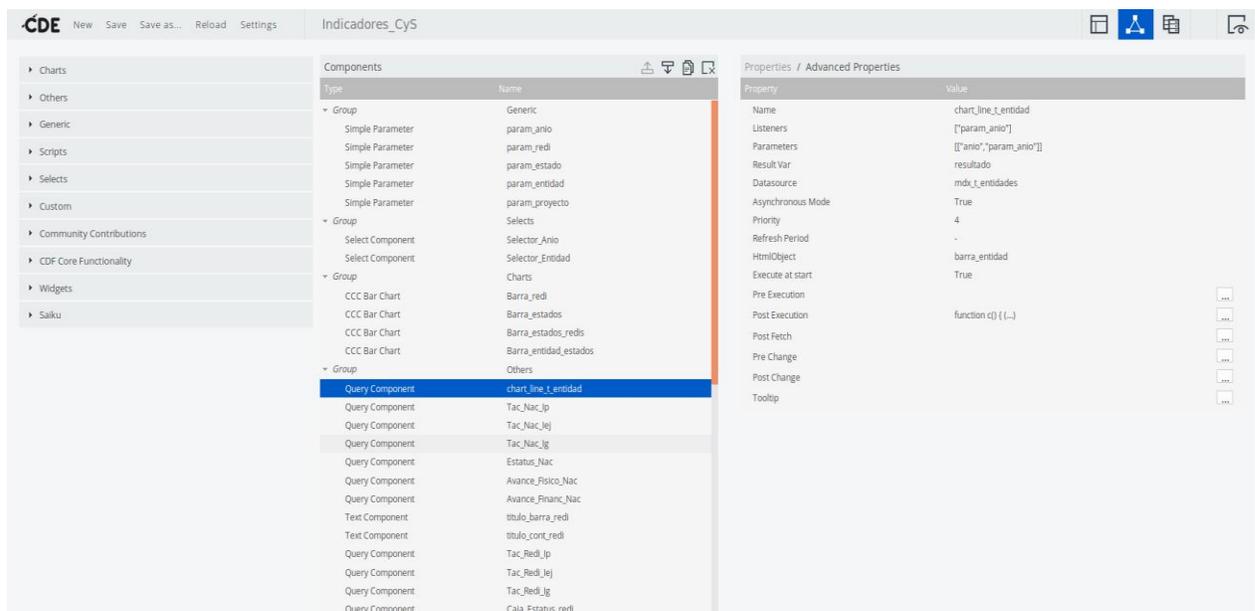


ILUSTRACIÓN 45: CAPA COMMUNITY CHART COMPONENTS (CCC)

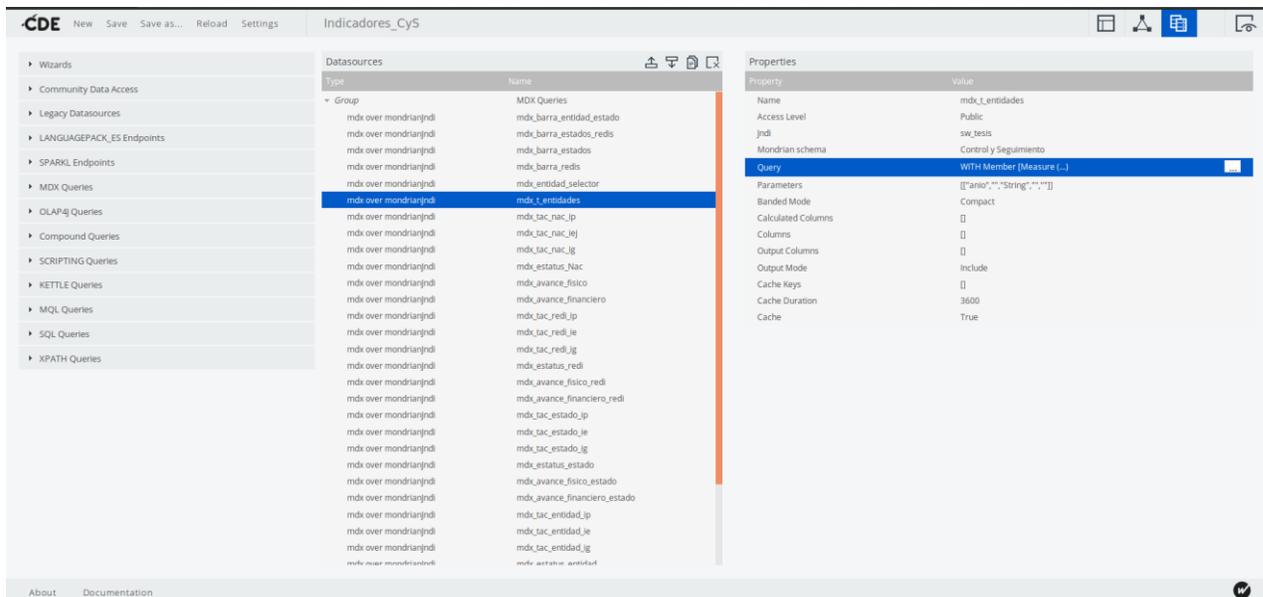


ILUSTRACIÓN 46: CAPA COMMUNITY DATA ACCESS (CDA)

Por último, el tablero de control permite la visualización de la información a través de gráficos de barras, gráficos de línea, tablas y tacómetros, y permite como se observa en las ilustraciones 47, 48, 49, 50 y 51 visualizar la información a través de diferentes niveles de detalle en función a la ubicación geográfica y del tiempo.



ILUSTRACIÓN 47: INDICADORES A NIVEL NACIONAL



ILUSTRACIÓN 48: INDICADORES A NIVEL DE REDI



ILUSTRACIÓN 49: INDICADORES A NIVEL DE ESTADO



ILUSTRACIÓN 50: INDICADORES A NIVEL DE ENTIDAD POLÍTICO-TERRITORIAL

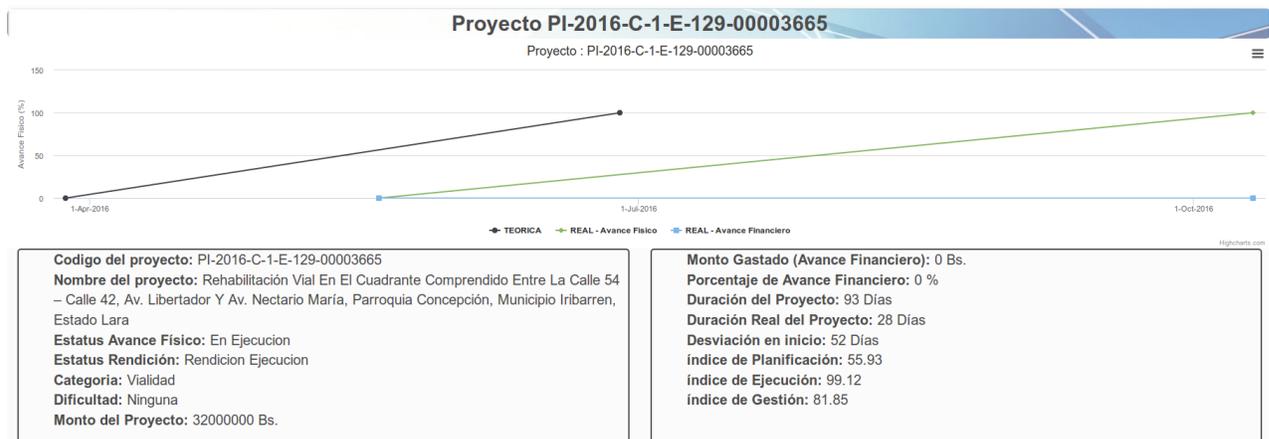


ILUSTRACIÓN 51: INDICADORES A NIVEL DE PROYECTO

Finalmente, el acceso al conjunto de reportes, al tablero de control y al cubo de información Mondrian, se puede realizar como se muestra en la ilustración 52 a través de un portal web.

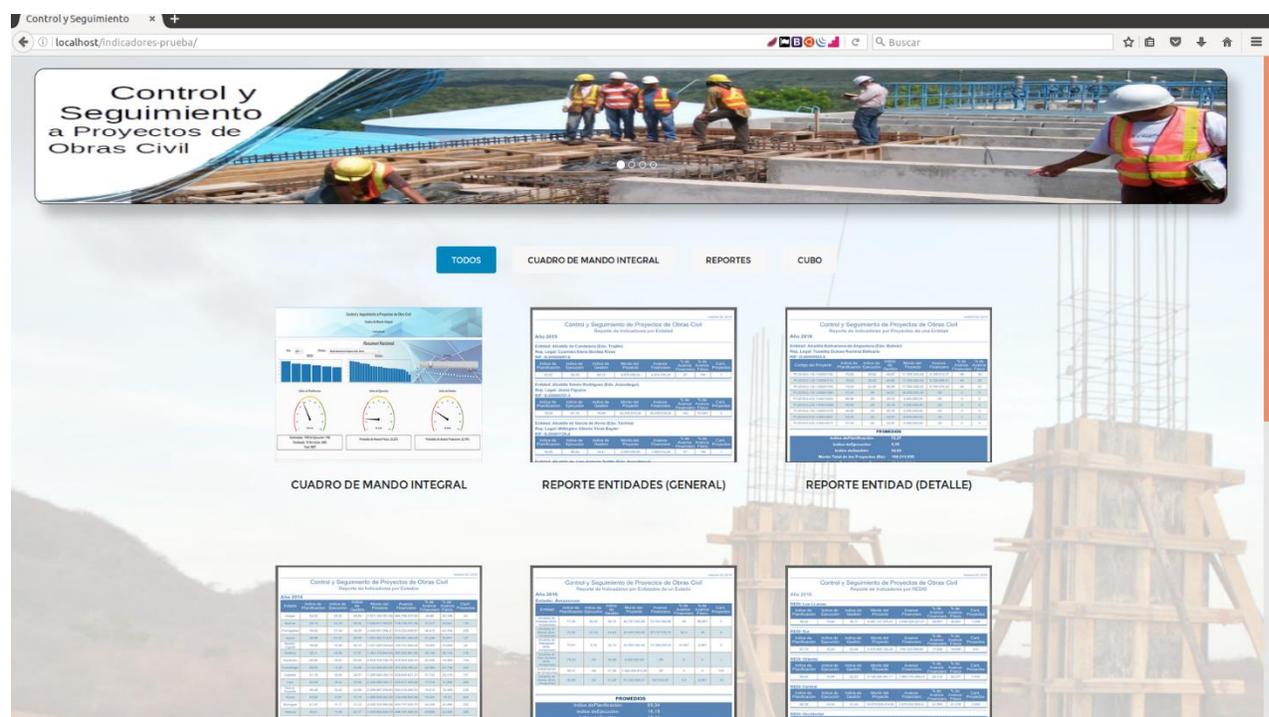


ILUSTRACIÓN 52: PORTAL WEB PARA ACCEDER AL TABLERO DE CONTROL, A LOS REPORTES Y A CUBO DE INFORMACIÓN MONDRIAN

#### 4.1.10. Implementación

La parte de implementación conlleva a su vez dos grupos de tareas. La primera es la instalación y configuración de elementos en el ambiente de producción. Por lo general el

desarrollo de las aplicaciones, el almacén de datos y los procesos ETC no se realizan directamente en la organización, sino que se ejecutan a partir de un subconjunto de datos dados por la organización en un ambiente de ejecución distinto.

En el caso de esta investigación, en el ambiente de producción se instalaron todas las herramientas mencionadas en el punto 4.1.4, y posteriormente se ejecutaron todos los procesos ETL para la carga de los datos al almacén de datos, posteriormente se importaron los archivos correspondientes al esquema de datos (cubo de información) a los reportes y al tablero de control para de esta manera tener la solución de inteligencia de negocios ya implementada en el ambiente de producción, finalmente se realizaron pruebas para verificar que la implementación se ejecute de manera correcta y se mantenga consistente.

#### 4.1.10.1. Pruebas de Calidad

Una vez que puesto en marcha en almacén de datos, se deben realizar un conjunto pruebas que permitan evaluar el rendimiento del sistema, es decir, que los datos están siendo cargados de forma correcta y los datos contenidos en el almacén de datos corresponden con los que se encuentran en el modelo dimensional. A continuación, se establece el plan de calidad de datos con los resultados obtenido:

- Aplicar una consulta para verificar la coincidencia en la cantidad de proyectos entre la fuente de datos y la dimensión **dim\_proyectos** del almacén de datos. (Ilustración 53)

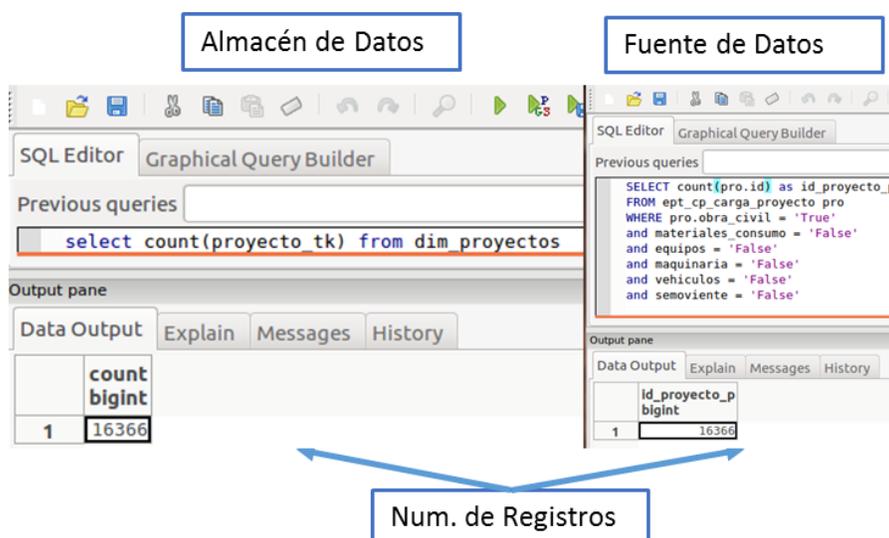


ILUSTRACIÓN 53: COMPARACIÓN ENTRE LA FUENTE DE DATOS Y LA DIMENSIÓN PROYECTOS.

- Aplicar una consulta para verificar la coincidencia en la cantidad de entidades entre la fuente de datos y la dimensión **dim\_entidad** del almacén de datos. (Ilustración 54)

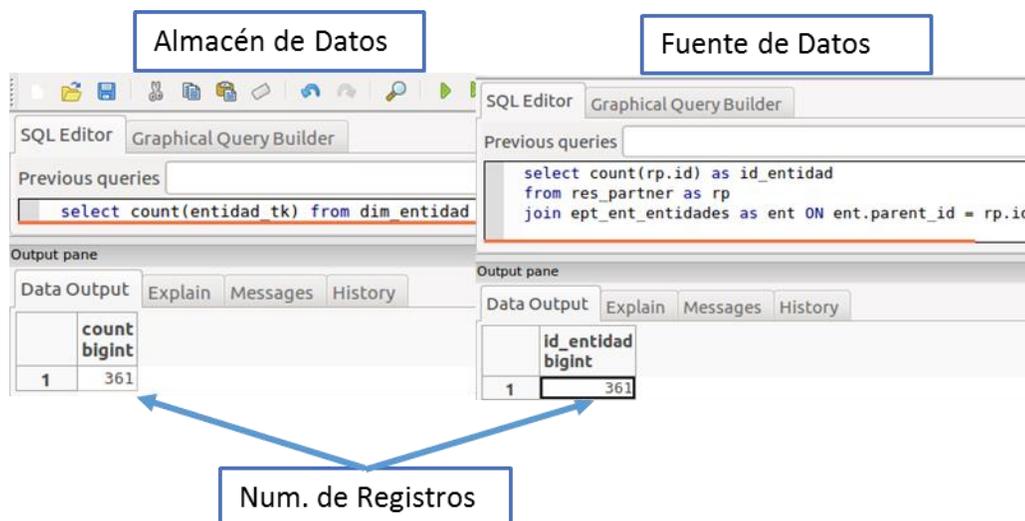


ILUSTRACIÓN 54: COMPARACIÓN ENTRE LA FUENTE DE DATOS Y LA DIMENSIÓN ENTIDAD.

- Aplicar una consulta para verificar la coincidencia en la cantidad de parroquias entre la fuente de datos y la dimensión **dim\_ubicación\_geo** del almacén de datos. (Ilustración 55)

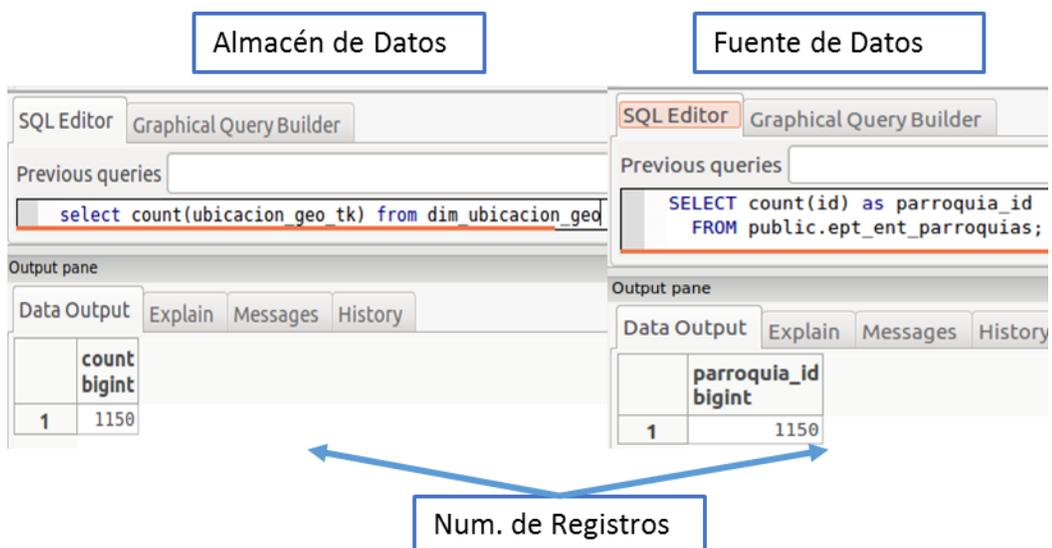


ILUSTRACIÓN 55: COMPARACIÓN ENTRE LA FUENTE DE DATOS Y LA DIMENSIÓN UBICACIÓN GEOGRÁFICA

En esta consulta se toma en cuenta los códigos de las parroquias por ser el nivel más bajo en la jerarquía de la ubicación geográfica

- Aplicar una consulta para verificar la coincidencia en la cantidad de parroquias entre la fuente de datos y la dimensión **dim\_estatus\_rnd** (estatus de rendición) del almacén de datos. (Ilustración 56)

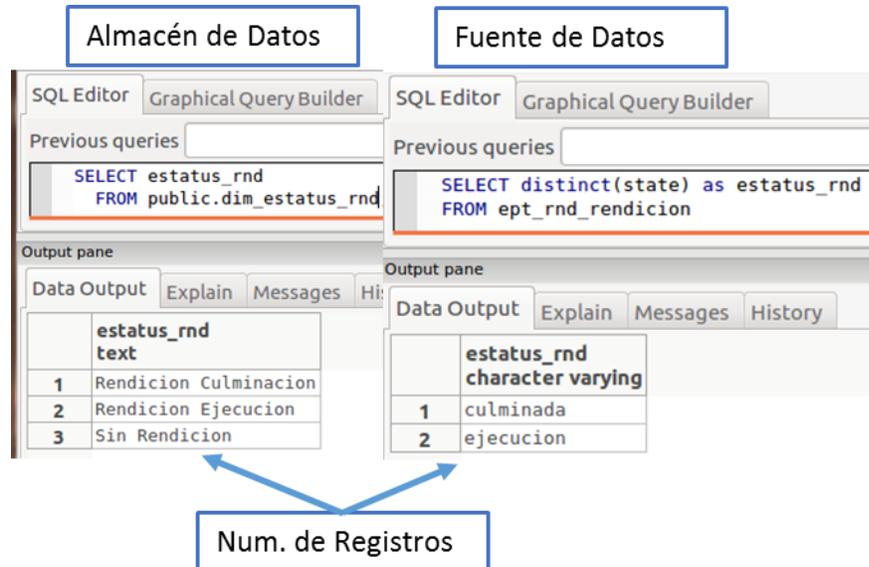


ILUSTRACIÓN 56: COMPARACIÓN ENTRE LA FUENTE DE DATOS Y LA DIMENSIÓN ESTATUS RENDICIÓN

En esta comparación la diferencia en el número de registro viene dada a que en el proceso ETL de esta dimensión se realizan transformaciones para mejorar la presentación de los datos y se agrega el registro “Sin Rendición” para cumplir con los requerimientos.

- Aplicar una consulta para verificar la coincidencia en la cantidad de parroquias entre la fuente de datos y la dimensión **dim\_estatus\_cp** (estatus de carga de proyecto) del almacén de datos. (Ilustración 57)

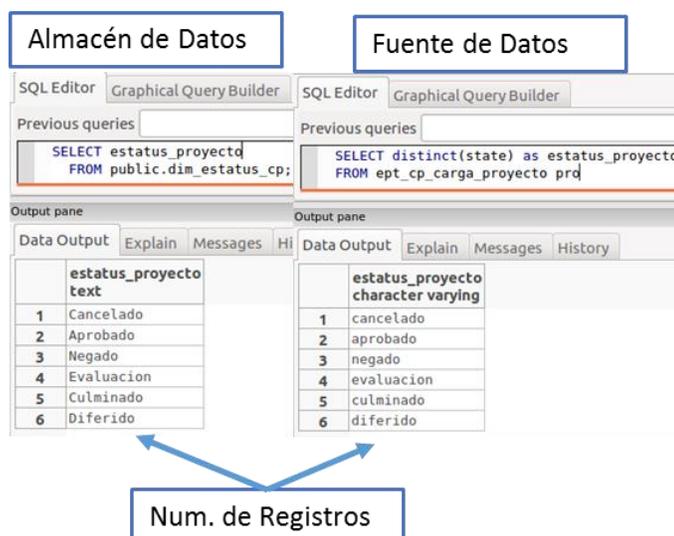


ILUSTRACIÓN 57: COMPARACIÓN ENTRE LA FUENTE DE DATOS Y LA DIMENSIÓN ESTATUS CARGA DE PROYECTO.

- Aplicar una consulta para verificar la coincidencia en la cantidad de parroquias entre la fuente de datos y la dimensión **dim\_estatus\_af** (estatus de avance físico) del almacén de datos. (Ilustración 58)

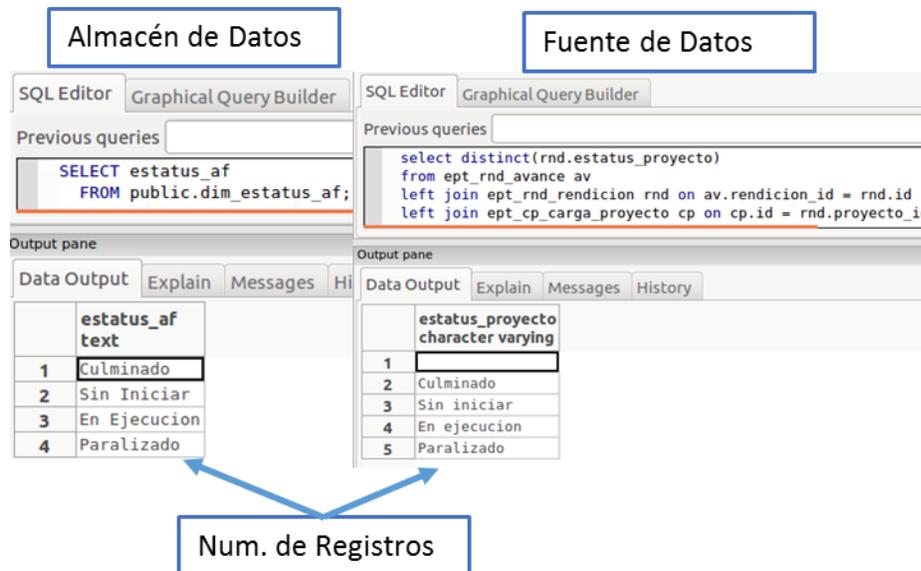


ILUSTRACIÓN 58: COMPARACIÓN ENTRE LA FUENTE DE DATOS Y LA DIMENSIÓN ESTATUS DE AVANCE FÍSICO

La diferencia en esta comparación viene dada a que en la fuente de datos existe un registro nulo, el cual no es relevante y por ello eliminado.

- Aplicar una consulta para verificar la coincidencia en la cantidad de parroquias entre la fuente de datos y la dimensión **dim\_categorización** del almacén de datos. (Ilustración 59)

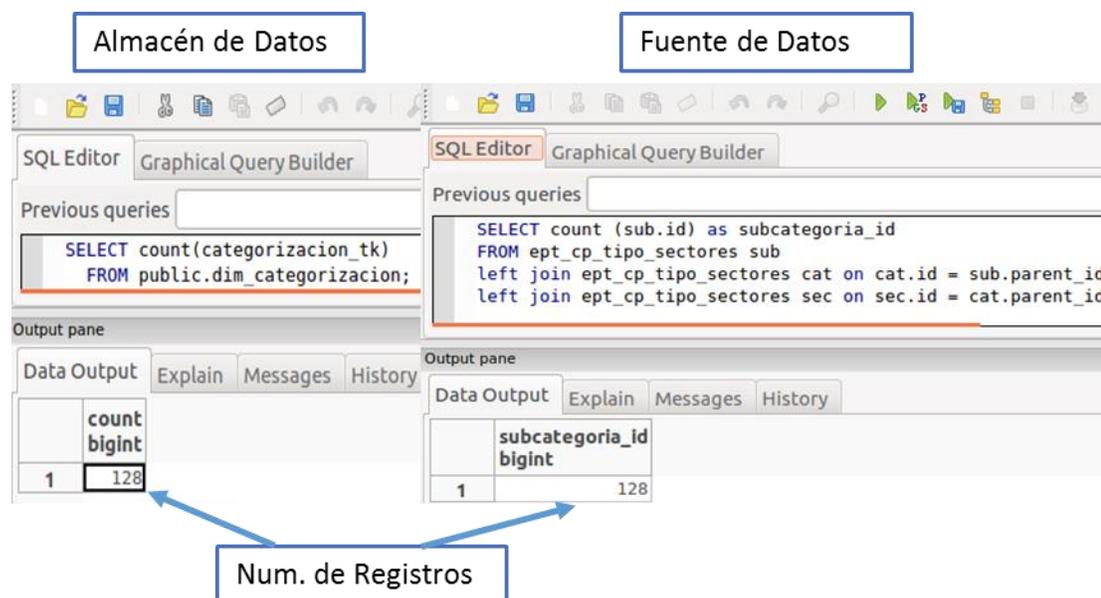


ILUSTRACIÓN 59: COMPARACIÓN ENTRE LA FUENTE DE DATOS Y LA DIMENSIÓN CATEGORIZACIÓN

En esta consulta se toma en cuenta los códigos de las subcategorías por ser el nivel más bajo en la jerarquía de la categorización.

- Aplicar una consulta para verificar la coincidencia en la cantidad de parroquias entre la fuente de datos y la dimensión **dim\_dificultad** del almacén de datos. (Ilustración 60)

Almacén de Datos

Fuente de Datos

Num. de Registros

ILUSTRACIÓN 60:  
COMPARACIÓN ENTRE  
LA FUENTE DE DATOS Y  
LA DIMENSIÓN  
CATEGORIZACIÓN

La diferencia en esta comparación viene dada a que en la fuente de datos existe un registro nulo, el cual es transformado en “Ninguna” en el proceso ETL.

- Aplicar una consulta para verificar la coincidencia en los hechos entre la fuente de datos y el almacén de datos. (Ilustración 61)

Sistema Transaccional

Cubo de información Mondrian

ILUSTRACIÓN 61: COMPARACIÓN ENTRE LA FUENTE DE DATOS Y LOS HECHOS DEL ALMACÉN DE DATOS

Para llevar a cabo esta comparación se utilizaron el sistema transaccional y el cubo de información Mondrian, en los que se pueden observar que los hechos contenidos en el sistema transaccional (fuente de datos) son iguales a los que muestra el cubo de información Mondrian (el cual se alimenta del almacén de datos).

Una vez puesto en producción el almacén de datos se pasa a la segunda tarea, la cual es la educación a los usuarios sobre la herramienta, lo que requiere una instrucción que permita a estos usuarios comprender el almacén de datos, así como el conjunto de herramientas que giran en torno al almacén de datos (aplicaciones de visualización, herramientas de administración, etc) y que utilicen según sea el caso. Es por ello, que, con el apoyo de varios usuarios, se realizaron inducciones y pruebas sobre la solución de inteligencia de negocios implementada, con lo cual se pudo apreciar que la solución es intuitiva y permite un rápido aprendizaje.

#### **4.1.10.2. Manual de usuarios**

Para hacer uso de las herramientas de visualización de la información (tablero de control, reportes y el cubo de información Mondrian) el usuario deberá seguir los siguientes pasos:

1. Ingresar en la barra de búsqueda del explorador de internet el siguiente enlace:  
**localhost/indicadores.:**

Seleccionar la herramienta a utilizar (hacer click sobre la imagen)  
(Ilustración 62):

- a. Tablero de Control.
- b. Reportes.
- c. Cubo de Información.

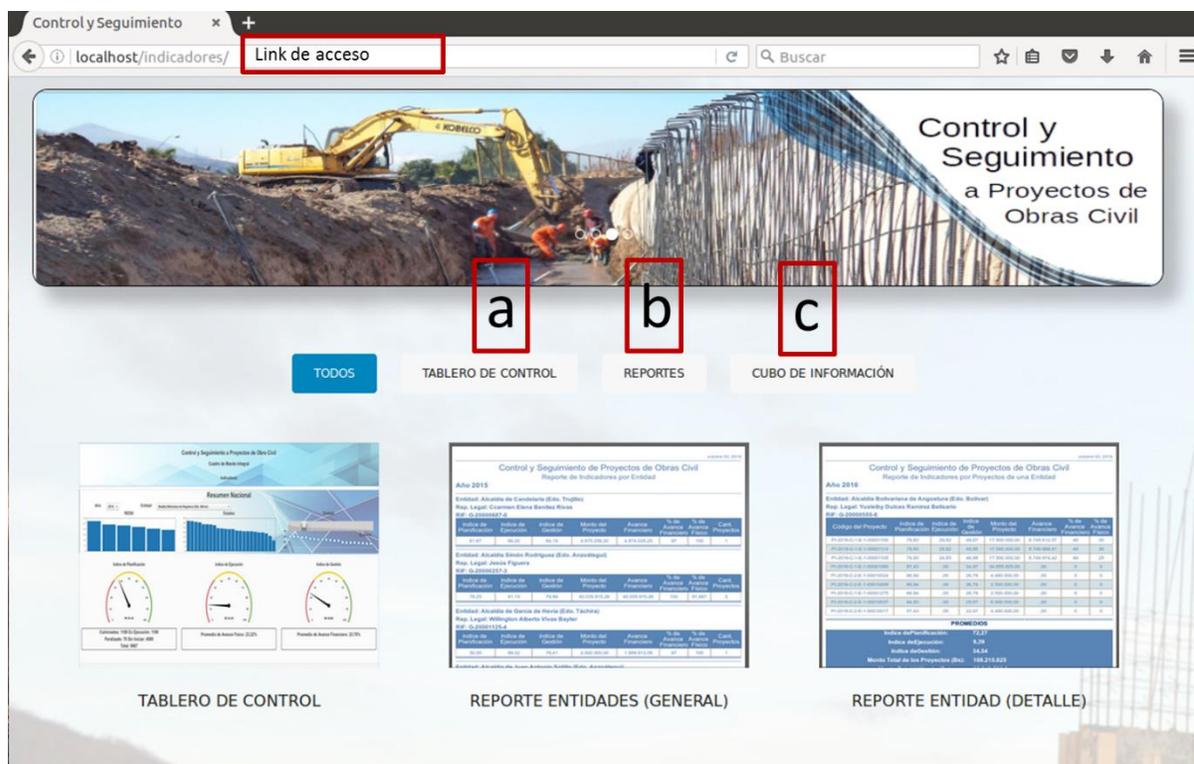


ILUSTRACIÓN 62: PÁGINA DE INICIO PARA INGRESAR A LAS HERRAMIENTAS DE VISUALIZACIÓN DE LA INFORMACIÓN

2. Al seleccionar el Tablero de Control el explorador de internet abrirá una ventana nueva con el Tablero de Control (Ilustración 63):

El tablero de control contiene los indicadores propuestos en función a la jerarquía de la ubicación geográfica, y al año de inicio de los proyecto. Está compuesta por cinco (5) niveles de despliegue que comienzan desde la información nacional, hasta la información de un proyecto específico.

En el primer menú se encuentran los siguientes componentes:

- a. Selector donde podrá elegir el año.
- b. Selector para elegir directamente una Entidad Política-Territorial, una vez elegida se desplegará un segundo nivel con la información de la Entidad Político-Territorial, además un gráfico de línea con los proyectos pertenecientes a esa entidad.
- c. Gráfico de barra donde se muestra las Regiones (REDI) en orden descendente por Índice de Gestión. Se puede seleccionar una región

- (hacer click sobre la columna del gráfico) y se desplegara un segundo nivel con la información de la región, además de un gráfico de barra con los Estados pertenecientes a esa región.
- d. Gráfico de barra donde se muestran los Estados en orden descendente por Índice de Gestión. Se puede seleccionar un Estado (hacer click sobre la columna del gráfico) y se desplegara un segundo nivel con la información del Estado, además de un gráfico de barra con las Entidades Político-Territorial pertenecientes a ese Estado.
  - e. Gráfico de línea donde se muestran las Entidades Político-Territoriales en orden descendente por Índice de Gestión. Se puede seleccionar una Entidad Político-Territorial (hacer click sobre la columna del gráfico) y se desplegara un segundo nivel con la información de la Entidad Político-Territorial, además un gráfico de línea con los proyectos pertenecientes a esa entidad.
  - f. Indicador de planificación: Un tacómetro que muestra el índice de planificación. Este Índice tendrá un valor entre 0 y 100, siendo 100 el valor ideal, si se encuentra en la franja roja significa que esta deficiente, si está en la franja amarilla significa que esta regular, si está en la franja verde significa que está bien.
  - g. Indicador de Ejecución: Un tacómetro que muestra el índice de ejecución. Este Índice tendrá un valor entre 0 y 100, siendo 100 el valor ideal, si se encuentra en la franja roja significa que esta deficiente, si está en la franja amarilla significa que esta regular, si está en la franja verde significa que está bien
  - h. Indicador de Gestión: Un tacómetro que muestra el índice de gestión. Este Índice tendrá un valor entre 0 y 100, siendo 100 el valor ideal, si se encuentra en la franja roja significa que esta deficiente, si está en la franja amarilla significa que esta regular, si está en la franja verde significa que está bien
  - i. Cantidad de Proyectos Según Estatus de Avance Físico a nivel nacional.
  - j. Promedio de Avance Físico: Es el promedio del porcentaje de la ejecución física a nivel nacional en un momento determinado y tendrá valores entre 0% y 100%
  - k. Promedio de Avance Financiero: Es el promedio del porcentaje del monto utilizado en la ejecución de los proyectos a nivel nacional en un momento determinado y tendrá valores entre 0% y 100%

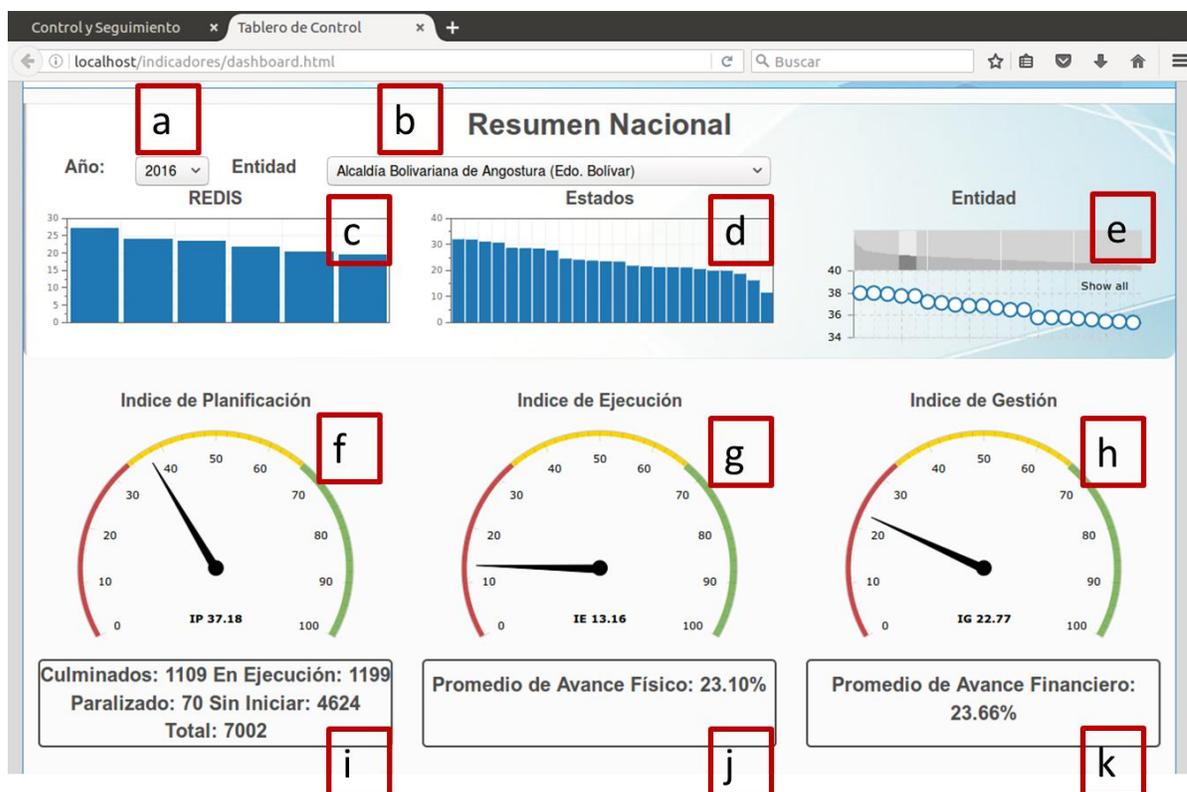


ILUSTRACIÓN 63: COMPONENTES DEL TABLERO DE CONTROL

El segundo, tercer y cuarto nivel de despliegue se presenta la información de manera similar a la ilustración 64:

- Gráfico de barra donde se muestra el nivel de ubicación geográfica en orden descendente por Índice de Gestión. Se puede hacer click sobre la columna del gráfico y se desplegara otro nivel con la información correspondiente.
- Indicador de planificación: Un tacómetro que muestra el índice de planificación. Este Índice tendrá un valor entre 0 y 100, siendo 100 el valor ideal, si se encuentra en la franja roja significa que esta deficiente, si está en la franja amarilla significa que esta regular, si está en la franja verde significa que está bien.
- Indicador de Ejecución: Un tacómetro que muestra el índice de ejecución. Este Índice tendrá un valor entre 0 y 100, siendo 100 el valor ideal, si se encuentra en la franja roja significa que esta deficiente, si está en la franja amarilla significa que esta regular, si está en la franja verde significa que está bien.

- d. Indicador de Gestión: Un tacómetro que muestra el índice de gestión. Este Índice tendrá un valor entre 0 y 100, siendo 100 el valor ideal, si se encuentra en la franja roja significa que esta deficiente, si está en la franja amarilla significa que esta regular, si está en la franja verde significa que está bien
- e. Cantidad de Proyectos Según Estatus de Avance Físico a nivel nacional.
- f. Promedio de Avance Físico: Es el promedio del porcentaje de la ejecución física a nivel nacional en un momento determinado y tendrá valores entre 0% y 100%
- g. Promedio de Avance Financiero: Es el promedio del porcentaje del monto utilizado en la ejecución de los proyectos a nivel nacional en un momento determinado y tendrá valores entre 0% y 100%

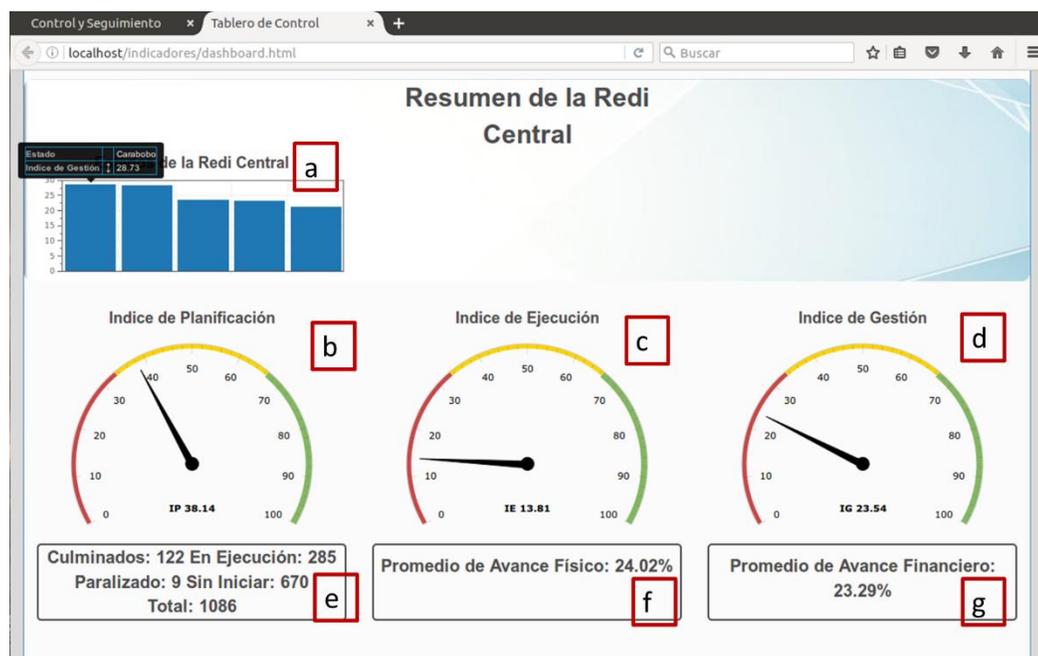


ILUSTRACIÓN 64: COMPONENTES DEL SEGUNDO, TERCER Y CUARTO NIVEL DEL TABLERO DE CONTROL

En el último nivel de despliegue se presenta la información referente a un proyecto como se muestra en la ilustración 65:

- a. Gráfico de líneas, donde se observa la línea teórica que refleja el comportamiento teórico del proyecto, la línea real de avance físico que refleja el comportamiento real del proyecto según la fecha de inicio y

- fin real del proyecto, y la lineal real de avance financiero que refleja el comportamiento del monto utilizado en la ejecución de la obra.
- b. Tabla que contiene la información básica del proyecto.
- c. Tabla que contiene los hechos e indicadores del proyecto.

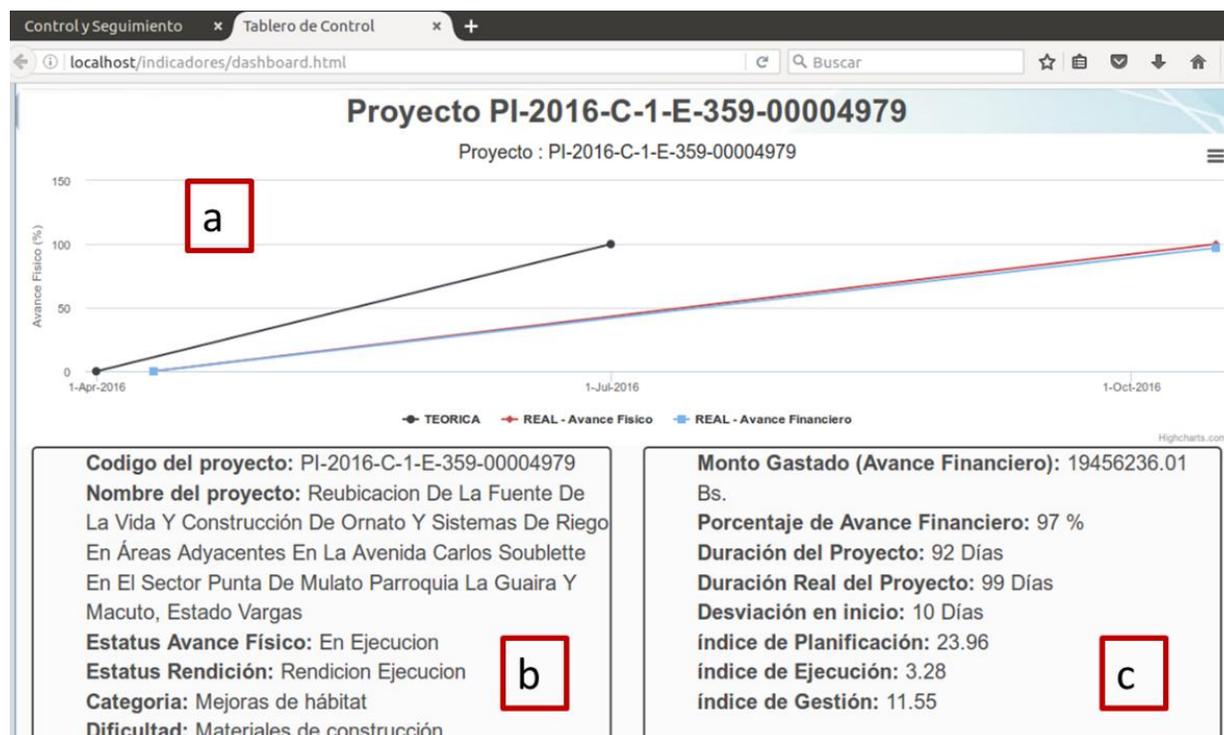


ILUSTRACIÓN 65: COMPONENTES DEL ÚLTIMO NIVEL DE DESPLIEGUE.

3. Al seleccionar alguno de los reportes el explorador de internet abrirá una ventana nueva con el reporte seleccionado:

Los reportes contienen los indicadores propuestos en función a la jerarquía de la ubicación geográfica, y al año de inicio de los proyectos, siguiendo una estructura similar a la ilustración 66:

- a. Selector del año.
- b. Selector de la entidad.
- c. Selector del tipo de archivo de salida del reporte ( HTML, PDF, CSV, Hoja de cálculo, RTF, Texto)
- d. Opción para ocultar los parámetros del informe (a, b, c) y así visualizar en pantalla completa el reporte.
- e. Opción para actualizar el reporte y generar la salida seleccionada.

Control y Seguimiento x Reporte de Entidad (D... x Tablero de Control x Cuadro de Análisis x +

localhost/indicadores/reporte\_entidades\_detalle.html

1 / 1

Año: 2016

Entidad: Alcaldía Metropolitana de Caracas

Tipo de Salida: HTML (Una Página)

Ver Informe Auto-Enviar

October 16, 2016

### Control y Seguimiento de Proyectos de Obras Civil

Reporte de Indicadores por Proyectos de una Entidad

**Año 2016**

Entidad: Alcaldía Metropolitana de Caracas  
Rep. Legal: Helen Torrealba  
RIF: G-20000153-4

Código del Proyecto	Índice de Planificación	Índice de Ejecución	Índice de Gestión	Monto del Proyecto	Avance Financiero	% de Avance Financiero	% de Avance Físico
PI-2016-C-1-E-335-00009271	75.00	.00	30.00	8,028,219.00	.00	0	0
PI-2016-C-1-E-335-00002643	75.00	.00	30.00	6,245,530.00	.00	0	0
PI-2016-C-2-E-335-00010221	68.86	.00	27.54	4,342,000.00	.00	0	0

**PROMEDIOS**

Índice de Planificación: 72.95  
Índice de Ejecución: 0

ILUSTRACIÓN 66: COMPONENTES DE LOS REPORTES

4. Al seleccionar el cubo de información el explorador de internet abrirá una ventana nueva con el cubo de información Mondrian para realizar análisis de información:

Permite a los usuarios explorar el almacén de datos, utilizando una interfaz familiar de arrastrar. Se seleccionan los datos de interés, y se pueden observar desde diferentes perspectivas. Una vez que se tiene la consulta se pueden guardar los resultados, compartíros, o exportarlos a Excel o PDF.

- a. Campo donde se encuentran todas las medidas (hechos) del almacén de datos.
- b. Campo donde se encuentran todas las dimensiones del almacén de datos.

- c. Campo donde se encuentran todas las medidas seleccionadas para realizar la consulta.
- d. Campo donde se colocan las dimensiones seleccionadas para realizar la consulta a nivel de columna.
- e. Campo donde se encuentran las dimensiones seleccionadas para realizar la consulta a nivel de fila.
- f. Campo donde se encuentran las dimensiones seleccionadas para realizar un filtrado en la consulta.
- g. Matriz con la información resultante de la ejecución de la consulta.
- h. Barra de opciones donde se puede realizar diferentes funciones como guardar la consulta, exportarla o compartirla.

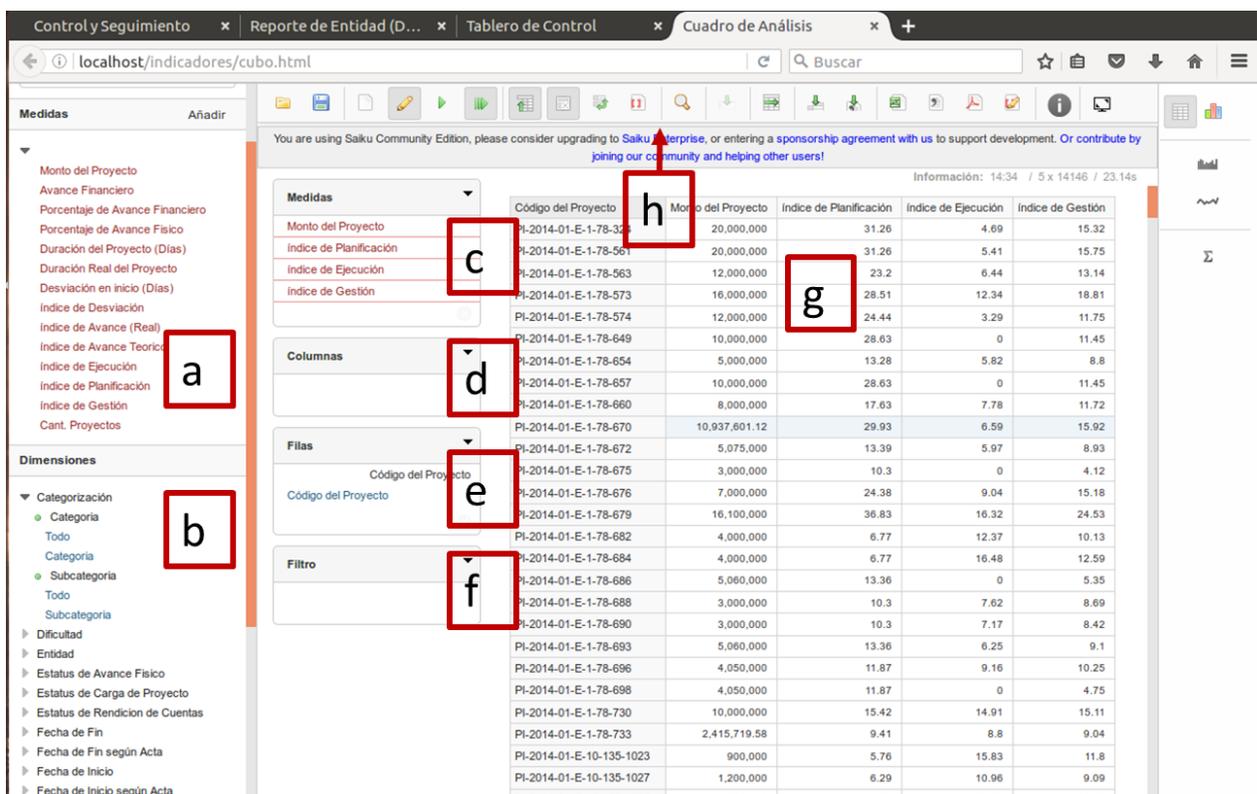


ILUSTRACIÓN 67: COMPONENTES DEL CUBO DE INFORMACIÓN

#### 4.1.11. Gestión del proyecto

La gestión del proyecto es una tarea que se ejecuta durante todo el desarrollo del proyecto e implica la verificación de que las actividades del proyecto se están ejecutando según lo planificado. Por tanto, en el caso de la solución de inteligencia de negocios propuesta, las

actividades se desarrollaron según lo establecido en el punto 4.1.1 dándose cumplimiento a la duración de veinte (21) semanas para la implementación de la misma.

## CONCLUSIONES Y RECOMENDACIONES

Se ha logrado satisfactoriamente el objetivo del Trabajo Especial de Grado, el cual consistió en el desarrollo de una Solución de Inteligencia de Negocio para la obtención de indicadores de Control y Seguimiento a Proyectos de Obra Civil, planificados por Entidades Político-Territoriales; con el fin de apoyar y respaldar las decisiones estratégicas de estas entidades, proporcionando información de calidad basada en los hechos reales del negocio

En este sentido, es importante mencionar que se realizaron estudios sobre el proceso de control y seguimiento a los proyectos de obra civil planificados por Entidades Político-Territoriales y fueron analizadas sus necesidades con el objetivo de cubrir satisfactoriamente los procesos de levantamiento y análisis de requerimientos, para así tener como resultado la construcción e implantación de los indicadores que facilitan el monitoreo de estos proyectos.

Además, se debe resaltar que, el desarrollo de esta solución se planteó bajo la metodología de Ralph Kimball cumpliendo con cada una de las etapas plasmadas en la misma, esta metodología es idónea cuando se está empezando a desarrollar una infraestructura de inteligencia de negocios porque se puede instaurar una solución de negocios completa sobre un área determinada y permite que el proyecto pueda ser puesto rápidamente en producción.

También es importante señalar, que las herramientas utilizadas para la implementación de la solución de inteligencia de negocios fueron las que componen 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 fue PostgreSQL. Se debe destacar que todas estas herramientas son de software libre (open source) por lo cual la solución de inteligencia de negocios está totalmente desarrollada en este tipo de software. Sin embargo, es importante resaltar, que la versión Pentaho Community posee limitaciones, como la imposibilidad de realizar reportes vía web, lo que conlleva a que solo se pueden hacer reportes mediante la herramienta Pentaho Report Designer. Por ello es

recomendable en lo posible utilizar las herramientas de la Suite Pentaho Business Analytics Enterprise Edition la cual si posee todas las funcionalidades mencionadas en el punto 2.6.1. Para finalizar, se obtuvo como resultado una solución de inteligencia de negocio mediante la cual se integran datos de diversas fuentes en un almacén de datos, permitiendo con ello, aprovechar un almacenamiento eficiente de los datos que ayuda a tener acceso a grandes volúmenes de información de forma rápida y sencilla, optimizando así el proceso de análisis, de esta forma se puede obtener a través de reportes, un tablero de control y un cubo de información Mondrian una visión general y específica sobre el desempeño de los proyectos que son planificados por Entidades Político-Territoriales al contar con indicadores de planificación, 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 Entidades Político-Territoriales.

## REFERENCIAS BIBLIOGRÁFICAS Y DÍGITALES

- Ackoff, R. (1994). *The Democratic Corporation: A radical prescription for recreating corporate America and Rediscovering success*. Nueva York: Oxford University Press.
- Alavi, M., & Leidner, D. (2003). *Sistemas de gestion de conocimiento: cuestiones, retos y beneficios en sistema de gestion del conocimiento. Teoria y Practica*. España: Thompson editores.
- Antonorsi, M. (1999). *Guia practica de la empresa competitiva*. Caracas: Centro de artes Integradas.
- Beltrán, J. (2006). *Indicadores de Gestión: Herramientas para lograr la competitividad*. Bogotá: 3R Editores.
- BI-INSIDER. (2011). *BI-INSIDER*. Obtenido de <http://bi-insider.com/portfolio/components-of-oracle-bi-enterprise-edition-obiee/>
- BI-INSIDER. (2011). *BI-INSIDER*. Obtenido de <http://bi-insider.com/business-intelligence/components-of-oracle-bi-enterprise-edition-obiee-oracle-answers/>
- Blanchard, B. (1993). *Aministracion de ingenieria de sistemas*. Mexico: Limusa.
- Burch, J., & Starter, F. (1981). *Sistemas de informacion: Teoría y práctica*. Mexico: Limusa.
- Cano, J. L. (2007). *Business Intelligence: Competir con Información*. España: ESADE.
- Codd, E. (1971). *Further Normalization of the Data Base Relational Model*. Nueva Jersey: Prentice Hall.
- Constructor Civil. (2016). *Constructor Civil/Etapas en un proyecto de construccion*. Obtenido de <http://www.elconstructorcivil.com/2013/04/etapas-en-un-proyecto-de-construccion.html>
- DANE. (2009). *Guia de Construccion e Interpretacion de Indicadores*. Bogota: DANE.
- Davenport, T., & Prusak, L. (1998). *Working Knowledge*. Boston: Harvard Business Scholl Press.
- De Sousa, P. (2014). *Las obras civiles son todos los trabajos realizados con la finalidad de la construcción, conformación, mantenimiento o demolición de una estructura o*

*elemento en un espacio determinado. Elaborado por un grupo técnico entre profesionales, personal obrero.*, Caracas.

- Espiñeira, Sheldon y Asociados. (2008). *PWC*. Obtenido de [www.pwc.com/](http://www.pwc.com/)
- Forta Ingenieria. (2015). *fortaingeneria/Proyecto de Obra Civil*. Obtenido de <http://fortaingeneria.com/proyecto-obra-civil/>
- Fuentes, E. J. (2003). *Sistemas de Información Gerencial*. Mexico.
- Gibert Ginestà, M., & Pérez Mora, O. (s.f.). *Universidad Abierta de Cataluña*. Obtenido de [http://ocw.uoc.edu/computer-science-technology-and-multimedia/bases-de-datos/bases-de-datos/P06\\_M2109\\_02152.pdf](http://ocw.uoc.edu/computer-science-technology-and-multimedia/bases-de-datos/bases-de-datos/P06_M2109_02152.pdf)
- Gil, P. (1997). *Sistemas y Tecnologías de la información para gestión*. Madrid: McGraw-Hill.
- Gravitar. (12 de Octubre de 2014). *Gravitar - Información sin límites*. Recuperado el 22 de julio de 2015, de Pentaho: <http://gravitar.biz/pentaho/>
- Harjinder, G. (1996). *Data Warehousing: La integración de información para la mejor toma de decisiones*. 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.
- Jaramillo, B. (2006). *Indicadores de Gestión - Herramientas para lograr la competitividad*. Bogota: 3R Editores.
- 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.
- Kimball, R., & Ross, M. (2013). *The Data Warehouse Toolkit: The complete guide to dimensional modeling*. EEUU: Wiley Computer Publishing.
- Laudon , K. C., & Laudon, J. P. (2012). *Sistemas de Información Gerencial*. New York: Always Learning Pearson.

- Laudon, K., & Laudon, J. (2008). *Sistemas de informacion gerencial: Administracion de la empresa digital*. Mexico: Pearson Education.
- Laudon, K., & Laudon, J. (2012). *Sistemas de informacion Gerencial*. Mexico: Pearson.
- Loyola University Chicago. (s.f.). *BUSINESS INTELLIGENCE*. Obtenido de <http://www.luc.edu/businessintelligence/ourtools/tableau/>
- Martínez, R. (2010). *Portal en español sobre Postgresql*. Obtenido de <http://www.postgresql.org.es/>
- Martins, S. (12 de septiembre de 2014). *Revista Latinoamericana de Ingenieria de Software*. Recuperado el 21 de julio de 2015, de Derivación del Proceso de Explotación de Información desde el Modelado del Negocio: <http://sistemas.unla.edu.ar/sistemas/redisla/ReLAIS/relais-v2-n1-53-76.pdf>
- Masip, D. (19 de julio de 2002). *desarrollobeb.com*. Recuperado el junio de 2015, de Qué es Oracle: <http://www.desarrolloweb.com/articulos/840.php>
- Mauricio, M. (2012). *Data Warehousing y Business Intelligence con herramientas Open Source*. Obtenido de <http://ingmmurillo-dwh-bi.blogspot.com/2012/06/arquitectura-de-flujos-de-datos-con.html>
- Molina, H. (2008). *Database System*. EEUU: Pearson.
- Murdick, R., & Ross, J. (1990). *Sistemas de Informacion Basados en Computadoras*. Mexico: Diana.
- O'Brien, J. (2001). *Sistemas de Informacion Gerencial*. Bogota: McGraw-Hill Interamericana, S.A.
- OCDE. (2000). *Indicadores Ambientales*. Madrid: Direccion General de Calidad y Evaluacion Ambiental.
- Online Business School. (2014). *OBS Business School*. Obtenido de Etapas de un Proyecto: [www.obs-edu.com/blog-project-managemen/etapas-de-un-proyecto/conoces-cuales-son-las-etapas-de-un-proyecto/](http://www.obs-edu.com/blog-project-managemen/etapas-de-un-proyecto/conoces-cuales-son-las-etapas-de-un-proyecto/)
- Oracle. (2007). *ARSON Group SAC*. Obtenido de <http://www.arsongroup.com/PDFs/OracleBISE1espaniol.pdf.pdf>
- Oracle. (2011). *Oracle*. Obtenido de [https://docs.oracle.com/cd/E24693\\_01/server.11203/e22487/chapter1.htm](https://docs.oracle.com/cd/E24693_01/server.11203/e22487/chapter1.htm)

- 
- Oracle. (Septiembre de 2014). *Oracle*. Obtenido de [http://docs.oracle.com/html/E18558\\_01/fusion\\_requirements.htm#BGBBDBBH](http://docs.oracle.com/html/E18558_01/fusion_requirements.htm#BGBBDBBH)
- Oracle. (2014). *Oracle*. Obtenido de [http://docs.oracle.com/pdf/E25460\\_01.pdf](http://docs.oracle.com/pdf/E25460_01.pdf)
- Oracle Corporation. (2006). *Oracle Business Intelligence Server Administration Guide*. Obtenido de Oracle Business Intelligence: [docs.oracle.com/cd/E10415/doc/bi.1013/b31770.pdf](https://docs.oracle.com/cd/E10415/doc/bi.1013/b31770.pdf)
- Oracle Corporation. (2007). *Oracle Business Intelligence Standard Edition One Tutorial*. Recuperado el Junio de 2015, de Oracle Business Intelligence: [http://docs.oracle.com/cd/E10352\\_01/doc/bi.1013/e10312.pdf](http://docs.oracle.com/cd/E10352_01/doc/bi.1013/e10312.pdf)
- Oracle Corporation. (2010). *Oracle® Fusion Middleware User's Guide for Oracle Business Intelligence Enterprise Edition*. Obtenido de Oracle Business Intelligence: [https://docs.oracle.com/cd/E21043\\_01/bi.1111/e10544/getstart.htm](https://docs.oracle.com/cd/E21043_01/bi.1111/e10544/getstart.htm)
- Oracle Corporation. (s.f.). *Oracle Business Intelligence Standard Edition One Tutorial*. Obtenido de Oracle Business Intelligence: <http://www.oracle.com/technetwork/middleware/bi-foundation/bi-standard-edition-one-overview-131081.pdf>
- Oracle Corporation. (s.f.). *Oracle Fusion Middleware Online Documentation Library 11g Release 1*. Obtenido de Oracle Documentation: [https://docs.oracle.com/cd/E28280\\_01/upgrade.1111/e16452/bi\\_plan.htm#FUGBI436](https://docs.oracle.com/cd/E28280_01/upgrade.1111/e16452/bi_plan.htm#FUGBI436)
- Oracle Corporation. (s.f.). *Setting Up Usage Tracking in Oracle BI 11g*. Obtenido de <http://www.oracle.com/webfolder/technetwork/tutorials/obe/fmw/bi/bi1116/ut/ut.html>
- Oracle Corporation. (s.f.). *Warehouse Builder Installation and Administration Guide*. Obtenido de Oracle Documentation: [https://docs.oracle.com/cd/E11882\\_01/owb.112/e17130/multi\\_design\\_mgmt.htm](https://docs.oracle.com/cd/E11882_01/owb.112/e17130/multi_design_mgmt.htm)
- Oracle. (s.f.). *Oracle*. Obtenido de <http://www.oracle.com/technetwork/es/documentation/317527-esa.pdf>
- Oracle. (s.f.). *Oracle*. Obtenido de <http://www.oracle.com/us/solutions/business-analytics/business-intelligence/enterprise-edition/features/index.html>
- Palacios, L. (2003). *Principios Esenciales para realizar proyectos; un enfoque latino*. Caracas: Universidad Catolica Andres Bello.
-

- Pentaho. (2012). *PENTAHO*. Recuperado el 12 de julio de 2015, de <http://www.pentaho.com/>
- Pentaho Corporation. (2013). *Infocenter 4.8*. Obtenido de [http://infocenter.pentaho.com/help48/index.jsp?topic=/install\\_ziptar/reference\\_support48.html](http://infocenter.pentaho.com/help48/index.jsp?topic=/install_ziptar/reference_support48.html)
- Pentaho Corporation. (2013). *Pentaho Aggregation Designer Overview*. Obtenido de Pentaho Infocenter 5.0: [http://infocenter.pentaho.com/help/index.jsp?topic=%2Faggregation\\_designer\\_guide%2Fconcept\\_pad\\_overview.html](http://infocenter.pentaho.com/help/index.jsp?topic=%2Faggregation_designer_guide%2Fconcept_pad_overview.html)
- Pentaho. (s.f.). *Gravitar*. Obtenido de [gravitar.biz/pentaho/](http://gravitar.biz/pentaho/)
- Pentaho. (s.f.). *Pentaho*. Obtenido de [www.pentaho.com/](http://www.pentaho.com/)
- Pentaho, A. H. (2005). <http://www.pentaho.com>. Obtenido de [http://www.pentaho.com/service/custom-dashboards:](http://www.pentaho.com/service/custom-dashboards)  
[http://ww3.pentaho.com/pentaho/api/repos/:public:Steel%20Wheels:CTools\\_dashboard.wcdf/generatedContent?userid=pentaho&password=demo](http://ww3.pentaho.com/pentaho/api/repos/:public:Steel%20Wheels:CTools_dashboard.wcdf/generatedContent?userid=pentaho&password=demo)
- Perez, C. (2003). *Los indicadores de Gestion*. Caracas.
- Pérez, J. (2015). *Indicadores de Avance, Rendimiento, Planificacion y Gesdtion de Proyectos delas Entidades Politico-Territoriales. PI-2015*. Caracas: Fondo de Compensación Interterritorial.
- Pinto, J. C. (2013). *Sistemas de Informacion*. Obtenido de Modulo 3: Sistemas de Apoyo a la Toma de Decisiones - Inteligencia de Negocio: [http://clasesdesi.blogspot.com/2013\\_01\\_01\\_archive.html](http://clasesdesi.blogspot.com/2013_01_01_archive.html)
- Project Management Institute. (2013). *Guia de los Fundamentos para la direccion de proyectos*. Pensilvania: Project Management Institute Inc.
- Radatz, J. (1997). *The IEEE Standar Dictionary of Electrical and Electronics Terms*. Nueva York: IEEE Standars Office.
- Rivera R., J. (26 de Julio de 2012). *Correo.ryrhermanos.com*. Recuperado el Junio de 2015, de [http://correo.ryrhermanos.com/home/jrivera@ryrhermanos.com/Publico/PDI\\_Parte\\_1: PDI\\_Parte\\_1.pdf](http://correo.ryrhermanos.com/home/jrivera@ryrhermanos.com/Publico/PDI_Parte_1: PDI_Parte_1.pdf)
- Sandra, A., & Lemus, J. (2009). *Pentaho - BI*. Bogotá.

- Stackiwiak, R., Rayman, J., & Greenwald, R. (s.f.). Oracle Data Warehousing and Business Intelligence Solutions.
- Stair, R., & Reynolds, G. (2000). *Principios de Sistemas de informacion*. Mexico: Ciencias THOMSON.
- Tableau. (s.f.). *Tableau Desktop*. Obtenido de <http://www.tableau.com/es-es/products/desktop>
- Tableau. (s.f.). *Tableau Products*. Obtenido de <http://www.tableau.com/es-es/products/>
- TechTarget. (2014). *SearchDataManagement*. Recuperado el 15 de julio de 2015, de Pentaho BI Suite Enterprise Edition: <http://searchdatamanagement.techtarget.com/review/Pentaho-BI-Suite-Enterprise-Edition>
- Tecnología de Gestión y Comunicación S.A. (2015). *TGC*. Recuperado el 20 de junio de 2015, de Inteligencia de Negocio (BI): <http://www.tgc.mx/web/bi-inteligencia-de-negocios.html>
- William, I. (2007). *The father of data warehousing*. San Diego: Inmon Consulting Services.