



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

**SOLUCIÓN DE INTELIGENCIA DE NEGOCIO
PARA LA OBTENCIÓN DE INDICADORES EN
LA NÓMINA DE INSTRUCTORES DE LA
FUNDACIÓN MUSICAL SIMÓN BOLÍVAR**

Trabajo Especial de Grado
presentado ante la Ilustre
Universidad Central de Venezuela
por el Bachiller
Guillermo José Obando Marín C.I. V-14.018.279
para optar al título de
Licenciado en Computación

Prof. Franklin Sandoval

Caracas, 07 / 10 / 2016

Facultad de Ciencias
Escuela de Computación
Centro de Investigación de Sistemas de Información

ACTA

Quienes suscriben, miembros del Jurado designado por el Consejo de Escuela de Computación, para examinar el Trabajo Especial de Grado presentado por el bachiller Guillermo José Obando Marín portador de la Cédula de Identidad V-14.018.279, con el título: **“Solución de inteligencia de negocio para la obtención de indicadores en la nómina de instructores de la Fundación Musical Simón Bolívar.”**, a los fines de cumplir con el requisito legal para optar al título de Licenciado en Computación, dejan constancia de lo siguiente:

Leído el trabajo por cada uno de los Miembros del Jurado, se fijó el día 07 de octubre de 2016, a las 11:00 am, para que su autor lo defendiera en forma pública en la Sala PB-III de la Escuela de Computación, Facultad de Ciencias, Universidad Central de Venezuela, lo cual se realizó mediante una exposición oral de su contenido, y luego respondió satisfactoriamente a las preguntas que les fueron formuladas por el Jurado, todo ello conforme a lo dispuesto en la Ley de Universidades y demás normativas vigentes de la Universidad Central de Venezuela.

Finalizada la defensa pública del Trabajo Especial de Grado, el jurado decidió aprobarlo con la nota de 20 puntos.


En fe de lo cual se levanta la presente acta, en Caracas a los 07 días del mes de octubre del año 2016.



Prof. Franklin Sandoval (Tutor)



Profa. Mercy Ospina
Jurado



Prof. Franky Uzcátegui
Jurado

DEDICATORIA

Este trabajo especial de grado está dedicado primeramente a Dios Todopoderoso, por acompañarme en todo momento. Por la educación impartida y por ser ejemplo de mujer luchadora, por apoyarme en todo momento, acompañarme y guiarme durante toda mi vida, por procurar lo mejor para mí y apartarme de los inconvenientes, por quererme y amarme como lo haces, por esas y más razones, esta tesis está dedicada a ti MADRE.

A mi familia, por su cariño, afecto y por contar siempre con su apoyo. A mis amigos, que nunca dejaron de darme ánimos, que siempre estuvieron brindando el empuje en los momentos más difíciles, ayudando y apoyándome en todo momento.

Y por último a todos quienes estuvieron pendientes de mi desempeño académico.

AGRADECIMIENTOS

A la ilustre Universidad de Venezuela, por aceptarme en sus filas, convirtiéndome en un UCEVISTA de corazón. Por permitirme realizar actividades culturales, deportivas, académicas entre otras, que me han formado como persona íntegra, de valores y principios que solo nuestra U.C.V. puede brindar.

A mis Profesores, por su guía y enseñanzas a lo largo del camino que me han ayudado a transitar este camino académico cuya meta es ser Licenciado de la U.C.V.

A mi Tutor de tesis, Profesor Franklin Sandoval por la paciencia, empeño y dedicación al momento de guiarme, tanto para la elaboración del trabajo especial de grado, como en otros ámbitos. Por su apoyo en todo momento, el cual conseguí siempre con una sonrisa.

A mis preparadores, compañeros con ganas de ayudarnos, ya sea en un aula, en un laboratorio o incluso en los pasillos, futuros profesores que seguramente formaran a las nuevas generaciones de Ucevistas y profesionales excelentes.

A mis amigos y compañeros, por los buenos momentos, consejos y recomendaciones oportunos, además de la ayuda incondicional.

A todas aquellas personas que de alguna u otra manera estuvieron involucradas en esta etapa de mi vida.

Muchísimas gracias.



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

SOLUCIÓN DE INTELIGENCIA DE NEGOCIO PARA LA OBTENCIÓN DE INDICADORES EN LA NÓMINA DE INSTRUCTORES DE LA FUNDACIÓN MUSICAL SIMÓN BOLÍVAR

Autor: Guillermo José Obando Marín
C.I. V-14.018.279
guilleom@gmail.com

Tutor: Prof. Franklin Sandoval
franklin.sandoval@ciens.ucv.ve

Fecha: 07/10/2016

RESUMEN

Actualmente las instituciones conocen la importancia de explotar la información contenida o generada de los procesos inherentes a su actividad. Por tal razón mantener indicadores, se convierte en herramienta importante, al momento de tomar decisiones, que pudiesen cambiar o no procesos. El presente trabajo especial de grado, tiene como objetivo desarrollar una solución de inteligencia de negocio para la obtención de indicadores en la nómina de instructores de la Fundación Musical Simón Bolívar (Fundamusical), para atender las necesidades de la división de nómina de instructores, con la finalidad de apoyar a la toma de decisiones estratégicas relacionadas con ese tipo de recursos humanos. Para el desarrollo de esta solución, que estuvo fundamentada en la metodología propuesta por Ralph Kimball y utilizando el software libre "Pentaho" con sus diferentes componentes, se definieron los indicadores de gestión, que apoyan a Fundamusical en el análisis de la información del proceso de nómina. Se configuró el almacén de datos y estructuras necesarias para el procesamiento de la información relevante. También se realizaron las consultas sobre esta información y por último, se configuró el *dashboard* donde se muestran los resultados conseguidos. Finalmente se obtuvo, como principal conclusión, que para la alta y media gerencia de Fundamusical, el uso de un *dashboard* para visualizar los indicadores, constituye una nueva fortaleza, ya que la información está condensada en un solo sitio, es fácil de obtener, está ordenada, y de sencilla comprensión, sirviendo así, como herramienta útil al momento de tomar decisiones.

Palabras Claves: Recursos Humanos, Inteligencia de Negocio, Indicadores de Gestión de nómina, Metodología de Ralph Kimball, Pentaho.

INDICE GENERAL

Acta.....	ii
Dedicatoria.....	iii
Agradecimientos	iv
Indice General.....	v
Indice de Tablas y Figuras	vii
Introducción	1
CAPÍTULO 1 - EL PROBLEMA.....	3
Situación actual.	3
Planteamiento del problema.....	5
Solución Planteada.....	7
Objetivo General.....	8
Objetivos Específicos	8
Justificación de la investigación	9
Alcance.....	10
CAPÍTULO 2 - MARCO CONCEPTUAL	12
Conceptos básicos	12
OLAP y OLTP.....	18
Bases de datos	21
Sistema Manejador de Bases de Datos	23
Data Warehouse.....	24
Modelado dimensional.....	30
Inteligencia de negocios	43
CAPÍTULO 3 - MARCO APLICATIVO	56
Planificación del proyecto.....	58
Definición del alcance.....	59
Justificación de la solución.	61
Definición de Requerimientos de Negocio.....	61
Diseño técnico de la arquitectura.	64
Selección de productos e instalación.	65
Diseño del Modelo Dimensional.	66
Diseño Físico.....	70
Diseño y construcción de procesos ETL.	72
Especificación y desarrollo de aplicaciones analíticas.	75
Implementación.	84
Mantenimiento y crecimiento.....	84
CAPÍTULO 4 – CONCLUSIONES Y RECOMENDACIONES	86
Referencias Bibliograficas	88

INDICE DE TABLAS Y FIGURAS

<i>Figura 01. Componentes de la estructura para Business Intelligence.</i>	7
Figura 02. Transformación de datos en conocimiento.	11
Figura 03. Pirámide de jerarquía de los sistemas de información.....	16
Figura 04. Arquitectura de un Data Warehouse.....	27
Figura 05. La estructura básica en el modelo multidimensional: el Cubo.	30
Figura 06. Esquema e instancia de un cuadro de doble entrada	31
Figura 07. Dimensión sin jerarquías alternativas.	31
Figura 08. Dimensión con jerarquías alternativas.	32
Figura 09. Granularidad de la dimensión, con una jerarquía de niveles.	33
Figura 10. Tabla de Hechos - Ventas.	35
Figura 11. Tabla de Hechos – Indicador.	35
Figura 12. Dimensiones	35
Figura 13. Ejemplo de un modelo de datos en estrella.	37
Figura 14. Ejemplo de un modelo de datos en copo de nieve.	37
Figura 15. Evolución del Bussines Intelligence.....	42
Figura 16: Evolución de la Inteligencia de Negocios.....	43
Figura 17. Plataforma Business Intelligence utilizando Dashboard.	50
Figura 18. Ciclo de Vida Tareas de la metodología.	52
Figura 19 Arquitectura de la solución de Inteligencia de Negocio.....	59
Figura 20. Jerarquía de la información transaccional	62
Figura 21. Modelo Dimensional	64
Figura 22. Modelo Físico.	65
Figura 24. Archivos de las transformaciones del ETL.....	68
Figura 25. Transformación en Spoon.....	68
Figura 26. Secuencia de tranformaciones (Job).	69
Figura 27. Ejecución del Job.....	70
Figura 28. Configuración del cubo de información.....	71
Figura 29. Configuración del cubo de información medidas	71

Figura 30. Ejemplo de consulta MDX.....	72
<i>Figura 31 Configuración de Data Sources</i>	72
Figura 32 Configuración de Analysis	73
Figura 33. Créate new – Jpivot view	73
Figura 34. Configuración de JPivot View (Reportes).	74
Figura 35. Ejemplo de Reporte	74
Figura 36. Layout panel	75
Figura 37. Datasources panel	76
Figura 38. Component panel	77
Figura 39. Dashboard.	77
Tabla 01. Diferencias entre OLTP y OLAP	18
Tabla 02. Características de un almacén de datos	22
Tabla 03. Plan de trabajo.....	54

INTRODUCCIÓN

Para las organizaciones, conseguir oportunidades de crecimiento es tan importante como las prestaciones que generan o beneficios que brindan. Por tal razón, al momento de tomar decisiones que afectarán al futuro de esta, se ven en la necesidad de estudiar la información que está relacionada con el cambio que posiblemente está por adoptar. Esta información puede venir de investigaciones realizadas en el entorno de la organización.

Por otra parte, cada organización lleva a cabo transacciones relacionadas con su actividad, transacciones que generan grandes volúmenes de datos e información y por lo general son manejados por algún sistema de información para el procesamiento de la misma. Estos sistemas usualmente son una fuente de información para la toma de decisiones.

La información no llega sola, por lo que el esfuerzo de los ejecutivos al recabarla y conseguir que la información sea actualizada y veraz, es de vital importancia. También es importante la existencia de información histórica en este proceso, ya que con estas se pueden observar tendencias sostenidas en el tiempo o patrones que se pueden identificar en algún sector de la organización sobre el cual se tomara cierta decisión.

En la Fundación Musical Simón Bolívar, dentro de la dirección de recursos humanos, existe una división dedicada a la nómina de los instructores que trabajan en los núcleos. Ellos representan la estructura funcional, educativa, artística y administrativa que da forma al Sistema Nacional de Orquestas y Coros Juveniles e Infantiles de Venezuela.

Los instructores de la fundación son de gran importancia como ya se mencionó, entre otras cosas, ellos representan el centro del eje educativo y por esta razón, mantener su información se torna significativa para la toma de decisiones estratégicas. Como meta a mediano plazo, la Fundación tiene estimado llegar a mantener a un millón de niños, niñas y adolescentes como beneficiarios de su plan social.

Para el apoyo en el logro de esta meta, la información de los instructores debe estar actualizada, pues regularmente es solicitada a la división de

registro control y servicio al instructor, información relacionada a proyecciones de cantidad y tiempo para el logro de esta meta.

La inteligencia de negocio al estar definida como un proceso que explora y analiza información en un área de negocio, para descubrir tendencias o patrones, de los cuales derivar ideas y extraer conclusiones y que incluye la comunicación de los descubrimientos, representa una alternativa para el manejo de esta situación, ya que ayudará a la obtención de la información deseada y el esfuerzo para conseguirla oportunamente, analizarla y presentarla se verán disminuidos sustancialmente, creando una ventaja estratégica.

Además representar una mejora tecnológica, también constituye una ventaja al poder brindar a la media y alta gerencia de la institución, información pertinente, rápida y veraz para llevar a cabo la toma de decisiones, que afectará el buen desenvolvimiento de los procesos que contribuyan con logro de las metas propuestas.

El presente trabajo se estructura en 4 capítulos descritos de la siguiente manera:

En el Capítulo I – Descripción del problema, se expone la situación encontrada en la Fundación Musical Simón Bolívar y relevancia de la misma, además se marca cuál fue el objetivo general y los objetivos específicos que se deben cumplir con este trabajo especial de grado.

El Capítulo II – Marco conceptual, contempla los fundamentos que sustentan este trabajo especial de grado, como son la noción de sistemas de información, almacenes de datos e inteligencia de negocios.

El Capítulo III – Marco aplicativo, tiene como finalidad, exponer el proceso de aplicar la metodología seleccionada, haciendo una completa descripción de las actividades que fueron realizadas en cada fase del proceso.

Y por último el Capítulo IV – Conclusiones y recomendaciones, está compuesto por las conclusiones, recomendaciones, bibliografías y referencias bibliográficas y no bibliográficas, del presente trabajo de grado.

CAPÍTULO 1

EL PROBLEMA

En este capítulo se describe la situación actual en la que se encuentra la división de registro, control y servicios al instructor de la Fundación Musical Simón Bolívar (Fundamusical), en cuanto al manejo de información importante, para la toma de decisiones de índole laboral y su necesidad de manejar indicadores de gestión sobre la ejecución de la nómina. Además, contiene el objetivo general y los objetivos específicos, la justificación y el alcance de esta investigación.

Situación actual.

Actualmente, es común encontrar en pequeñas y medianas empresas del país, la utilización de sistemas de gestión empresarial del tipo ERP (por sus siglas en inglés, *Enterprise Resource Planning*), los cuales permiten apoyar los procesos operativos de una compañía. Estos sistemas generalmente incluyen las actividades de producción, distribución, facturación, nómina, contabilidad, entre otras, integrando diferentes funciones de la organización en una plataforma homogénea.

Como resultado de estas actividades operacionales, organizaciones de todo tipo poseen una inmensa cantidad de información residente en sus sistemas de procesamiento de transacciones, la cual aumenta en tamaño progresivamente, dificultando las labores de análisis, particularmente cuando las relaciones entre los datos se tornan más complejas. La mayoría de los sistemas administrativos modernos proporcionan reportes estándares preconfigurados que satisfacen los requerimientos básicos de presentación de la información. Estos informes proporcionan indicadores de los procesos

operativos y administrativos de manera estática, que deben ser consultados manualmente, de forma independiente. Sin embargo, el mercado cambiante exige mejores herramientas que permitan a los ejecutivos realizar un análisis exhaustivo de los datos disponibles para administrar sus negocios.

Pero hoy, la información es un factor de valor incalculable para el éxito empresarial; si esta no es orientada adecuadamente y no está disponible para su uso en el momento oportuno, puede perder todo valor ante el proceso de toma de decisiones. Esto hace evidente la necesidad de procesos y herramientas que faciliten el manejo e interpretación de los datos que arrojan los sistemas de operación y producción. De acuerdo a García y Murillo (2008), entre otros, coinciden con que

...en muchas empresas se presenta la necesidad de analizar la información para ser utilizada al momento de tomar decisiones. Para esto se valen de reportes, tablas, gráficos y cualquier otro, donde diferentes datos son relacionados dependiendo de la información que desean extraer para tomar la mejor acción a seguir sobre un tema en particular. (p.22).

Por lo tanto, las iniciativas corporativas deben estar orientadas a proveer a la gerencia información relevante para responder a los problemas del negocio. Las razones expuestas, muestran causas del por qué la información proporcionada por estos sistemas, no resulta fácilmente utilizable por los altos directivos para la toma de decisiones. Además, los niveles estratégicos de las organizaciones requieren tener conocimiento del entorno que les permita evaluar las condiciones del negocio y vaticinar futuros eventos que afecten sus intereses.

En este contexto, la Fundación Musical Simón Bolívar (Fundamusical), no es ajena a estas necesidades de gestión de la información y al enfocarse en su División de Registro, Control y Servicios al Instructor (DRCSI), la cual se dedica, dentro de la Dirección de Recursos Humanos, a llevar un control de la nómina de los instructores, quienes a diario tienen la responsabilidad de

formar a los niños, niñas y adolescentes en aproximadamente 500 núcleos del Sistema, a nivel nacional.

De manera mensual es requerida, desde diferentes direcciones a la DRCSI informes, dentro de los cuales se pueden mencionar, el costo de la nómina, cantidad de instructores, costos de incidencias, cantidades de instructores por núcleo, para ser validado el cumplimiento de metas impuestas o tomar medidas presupuestarias, entre otras. Como ejemplo concreto, solicitan conocer costo de la nómina de los instructores en uno o más meses determinados con la cantidad de instructores respectiva, clasificada por niveles (nivelación salarial de los instructores), núcleos y proyectos, para evaluar la factibilidad de un aumento salarial y el porcentaje del mismo.

La DRSCI elabora estos informes a partir de información reflejada en el cierre de las nóminas o consultando directamente la base de datos del sistema transaccional de nómina. Los analistas manejan grandes volúmenes de información y al relacionarlas es probable que se pierda exactitud, originando que este proceso de extracción de información sea susceptible a errores.

El proceso de análisis y presentación de la información solicitada, ya sea en un gráfico, cuadro o un correo electrónico, requiere tiempo adicional de los analistas encargados, lo que origina retrasos en los demás procesos de la división. Tomando en cuenta que el tiempo es un recurso que debe administrarse de manera precisa para el pago de la nómina de los instructores, debe mejorarse la forma de obtención de la información para ser analizada y presentada posteriormente.

Planteamiento del problema

Los procesos administrativos, contables y de nómina de Fundamusical, son llevados en un sistema transaccional (Sistema Integrado de Gestión Administrativa, SIGA), el cual mantiene una base de datos, la cual es la fuente

principal de información para la presentación de informes a la directiva para la toma de decisiones. La inexistencia de un almacén de datos que sostenga un modelo dimensional, para mantener información histórica de la nómina de los instructores y hacer que esta sea de fácil acceso para dicha toma de decisiones, representa una debilidad para Fundamusical. Como algunas características de esto se pueden enumerar:

- Retrasos en la obtención de información y elaboración de reportes.
- Utilización de tiempo extra, dedicado a atender solicitudes de información de otras direcciones.
- La información generalmente se entrega en cuadros, tablas o información plana y no es presentada con la ayuda de alguna herramienta analítica, tanto digital como física, para facilitar su comprensión.
- La información de los reportes puede ser manipulada y modificada, hasta llegar al usuario final, perdiendo veracidad en algunos casos.
- La extracción, análisis y presentación de los datos permanecen en el computador del analista que realizó esa labor.

Ya que los reportes se elaboran manualmente, no existe garantía de la veracidad y exactitud de la información presentada, aunado al retraso inherente en las labores del analista es sus demás responsabilidades. Generalmente, el analista que elabora el reporte, con la ayuda de una hoja de cálculo no conoce que decisión se tomará con la información que le están solicitando y por este motivo el gerente que recibe la información debe realizar un análisis adicional para extraer la información que para él, es relevante.

En resumen, el soporte para la toma de decisiones es escaso, ya que no existe una herramienta que permita gestionar la información histórica, considerándola en algunos casos, inexistente o inaccesible, originando que existan aspectos que al momento de tomar decisiones no sean tomados en

cuenta. Por ejemplo, es vital conocer con exactitud, la cantidad de instructores por cada nivel, para determinar proyecciones de gastos por concepto de nómina para meses venideros o en cuál núcleo existen más profesores que dictan menos de un número de horas determinadas.

Solución Planteada

Tomando en cuenta las debilidades y situaciones identificadas en los puntos anteriores, y luego de realizar el análisis de los requerimientos a la DRSCI, se propone proveer a los involucrados, tanto la DRSCI como las direcciones solicitantes, una solución de inteligencia de negocios, que soportará la toma de decisiones estratégicas, tomando la información del sistema transaccional que actualmente funciona.

Para esto se implementará una arquitectura orientada a Inteligencia de negocios (BI en inglés: *Business Intelligence*), como se muestra en la figura 01, que consta de un proceso de extracción, transformación y carga (en inglés: *Extract, Transform and Load, ETL*) de la información relevante, contenida en el sistema transaccional existente. El esquema relacional que estructura los datos existentes, servirá de base para alimentar un modelo multidimensional.

Además soportará el diseño de metadatos que permita la creación de reportes o consultas detalladas, que serán administrados por la Dirección de Tecnología en sus servidores, para la utilización de los usuarios interesados. Y por último proveer un ambiente donde los usuarios de las diferentes direcciones puedan realizar el análisis de la información requerida de manera rápida, precisa y disponible fácilmente.

Con lo anteriormente expuesto se logra automatizar el proceso de extracción, análisis y presentación de la información relevante de la nómina de instructores, ya que toda la información estará disponible en un solo lugar, lista para ser presentada al usuario final.

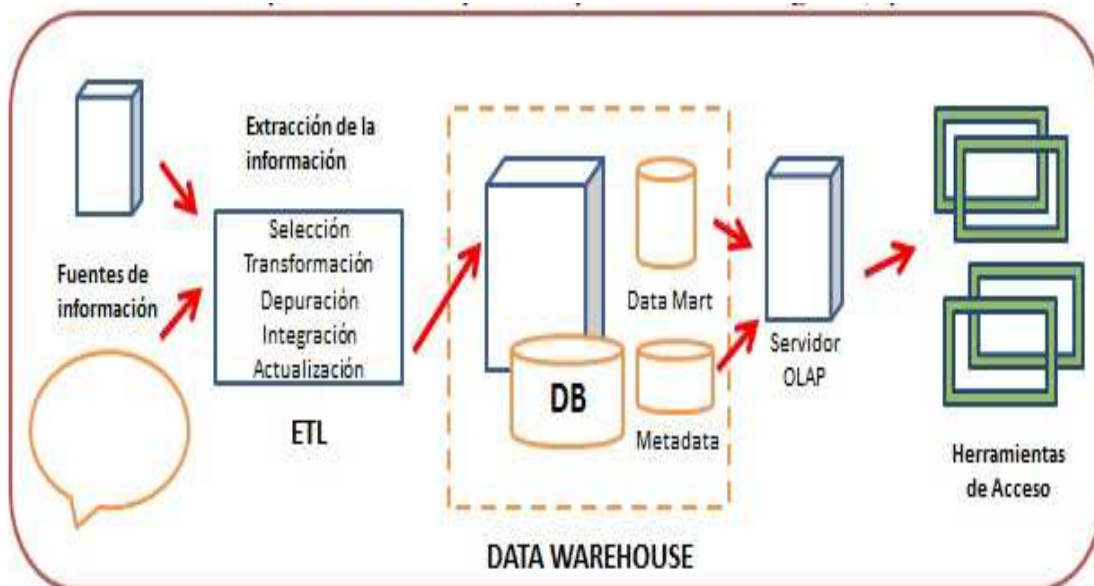


Figura 01. Componentes de la estructura para *Business Intelligence*.
Fuente: pensando en bi (2015).

En la imagen anterior se aprecian los componentes de la estructura de inteligencia de negocios, que servirá como guía para este trabajo especial de grado

Objetivo General

Desarrollar una solución de inteligencia de negocio para la obtención de indicadores de gestión en la nómina de instructores de la Fundación Musical Simón Bolívar.

Objetivos Específicos

- Identificar los requerimientos de información solicitada a la DRSCI, reflejando los indicadores sobre la nómina de instructores de la Fundación Musical Simón Bolívar.
- Identificar el modelo de datos del proceso de nómina para la construcción de una solución de BI.
 - Diseñar un modelo dimensional que cubra las necesidades de la DRSCI.

- Construir un almacén de datos, basado en el modelo dimensional diseñado.
- Construir el proceso de extracción, transformación y carga de la información de la nómina de instructores de Fundamusical.
- Elaborar un Portal Web Interactivo (*dashboard*) donde serán publicados los indicadores y reportes creados.
- Desarrollar un conjunto de pruebas de usabilidad, aceptación y encuestas correspondientes a la evaluación de los componentes de la solución.

Justificación de la investigación

Toda empresa, en el desarrollo de su actividad económica, debe tomar decisiones acertadas que le permitan posicionarse en un mercado altamente competitivo. Generalmente estas decisiones se toman con base en información que procede de fuentes como la intuición personal, el criterio empresarial, las opiniones de compañeros, o bien datos internos y externos de la organización. Sin embargo, las empresas deben disponer de información de calidad que permita conocer el comportamiento de la organización y monitorear los indicadores críticos de gestión, de manera que los ejecutivos dispongan de bases sólidas para generar una ventaja competitiva derivada del conocimiento obtenido de los datos operacionales provenientes de los diferentes procesos de la organización.

Es fundamental entonces, que las empresas dispongan de herramientas que permitan diagnosticar situaciones que se presenten durante la ejecución de las actividades del negocio, disponiendo de una mayor cantidad de información útil en el menor tiempo posible. Además, estas deben presentar los datos críticos de la organización de forma fácil e intuitiva, de tal manera que pueda ser manipulada por los gerentes y que el acceso a la información no requiera de la intervención de un experto en tecnología. Cada vez más, las organizaciones han notado que los beneficios provenientes de las inversiones

en soluciones ERP pueden aumentar dramáticamente con la extracción de información adicional, a partir de los datos del negocio almacenados en estos sistemas.

En esta medida, las tecnologías de BI han sido ampliamente utilizadas como un medio para obtener valor agregado de estos datos. La intención de este proyecto es proporcionar un marco de trabajo para la implementación de una solución de Inteligencia de Negocios que pueda ser utilizado por la Fundación Musical Simón Bolívar, o mejor conocida como Fundamusical, en su División de Registro, Control y Servicios al Instructor (DRCSI), división encargada de la generación y obtención de indicadores de gestión en la nómina de instructores.

Alcance

El alcance de este trabajo especial de grado, contempla una solución de inteligencia de negocios, la cual plantea como el propósito principal, que los usuarios de alta y media gerencia puedan tener acceso la información relevante para la toma de decisiones estratégicas. Se tomarán en cuenta diez (10) indicadores esenciales para esta investigación.

De igual forma se trabajará con los reportes necesarios para garantizar el buen funcionamiento de los indicadores creados, y se llevará a cabo la creación de un portal web para mostrar los resultados de los indicadores planteados. Posteriormente a esta adopción de inteligencia de negocios, podrán adicionarse más indicadores, según sean los requerimientos, lo cual queda fuera del alcance de este trabajo. Los indicadores a considerar son:

- Cantidad de Empleados por nómina.
- Cantidad de instructores por nivel.
- Cantidad de Cargos por nivel.
- Cantidad de instructores sin cuenta bancaria.
- Cantidad de instructores que dictan cases en más de un núcleo.

- Cantidad de instructores por cátedra.
- Cantidad de instructores que dictaron más de 160 horas al mes.
- Cantidad de instructores que dictaron más de 160 horas al mes, clasificados por estado.
- Cantidad total de instructores por estado.
- Cantidad de instructores con pagos extraordinarios (fuera del mes de nómina).

CAPÍTULO 2

MARCO CONCEPTUAL

En este capítulo se presenta el sustento teórico relacionados con el trabajo de investigación del desarrollo de una solución de inteligencia de negocio para la obtención de indicadores en la nómina de instructores de la Fundación Musical Simón Bolívar. El mismo se divide en 3 grandes secciones, las cuales se describen a continuación:

La primera sección, hace referencia a conceptos relacionados con sistemas de información y bases de datos. La segunda sección está formada por contenido referente a *datawarehouse* y modelado dimensional. Por último en la tercera sección se encuentran tópicos de inteligencia de negocios.

Conceptos básicos

Davenport y Prusak, (2000) afirman que

*“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”.* (p.20).

Para Beynon-Davies (2013) un dato

*“es uno o más símbolos que se usan para representar algo. La información son datos interpretados. La información son datos emplazados dentro de un contexto significativo. Por lo tanto, el uso del término *información* implica un grupo de gente realizando una interpretación.”* (p.27).

Chiavenato (2006.) define

“Información es un conjunto de datos con un significado, o sea, que reduce la incertidumbre o que aumenta el conocimiento de algo. En verdad, la información es un mensaje con significado en un determinado contexto, disponible para uso inmediato y que proporciona orientación a las acciones por el hecho de reducir el

margen de incertidumbre con respecto a nuestras decisiones". (p. 129).

Ferrell et al, (2004) complementa la definición anterior con

"...la información comprende los datos y conocimientos que se usan en la toma de decisiones". (p.11).

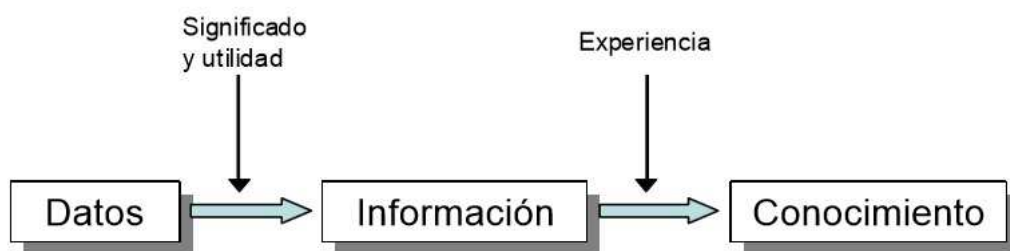


Figura 02. Transformación de datos en conocimiento.
Fuente: Suarez (2012).

Véase la Figura 02 para observar la relación entre dato, información y conocimiento.

Para Peralta (2006) **Sistemas** como

"...el modelo desarrollado para la aplicación, encargado de satisfacer el requerimiento planteado por el usuario, caracterizado por el empleo de recursos que realizan las tareas especificadas por un conjunto de procesos dedicados al tratamiento de información".

Bernal (2004) citando a Spaulding indica que

"El sistema es un conjunto de procedimientos interrelacionados que señalan el plan de acción para alcanzar objetivos básicos de una organización". (p.14).

Peña (2006) agrega que

"Los **Sistemas de Información** (SI) son un conjunto de procedimientos diseñados para el acopio, almacenamiento, procesamiento y explotación de información en apoyo a las

actividades que se realizan en una organización y que resultan de interés para el personal y el desarrollo del negocio”. (p.45).

Para Laudon y Laudon (2012) un Sistema de Información (SI)

“...es un conjunto de componentes interrelacionados que recolectan (o recuperan), procesan, almacenan y distribuyen información para apoyar los procesos de toma de decisiones y de control en una organización”. (p. 4).

Y Peralta (2006) complementa con que

“Un sistema de información es un conjunto de elementos que interactúan entre sí con un único fin, el de apoyar todas las actividades de una empresa o negocio. Teniendo en cuenta el equipo computacional que se necesita para que el sistema de información pueda operar y un grupo de personas que es el recurso humano que interactúa con el sistema de información”.

Según su uso, los sistemas de información, se pueden clasificar observando una jerarquía dentro de una empresa como: Sistemas Transaccionales, TPS o de gerencia operacional. Situados en la base de la pirámide jerárquica que se mostrará más adelante, Kendall y Kendall (2005). Indican que

“...son creados para procesar grandes cantidades de datos relacionadas con transacciones rutinarias de negocios, como las nóminas y los inventarios. Un TPS elimina la molestia que representa la realización de transacciones operativas necesarias y reduce el tiempo que una vez fue requerido para llevarlas a cabo de manera manual”.(p27).

Pérez et al (2015) menciona que

“...la principal características de los sistemas transaccionales, es que a través de estos suelen lograrse ahorros significativos de mano de obra, debido a que automatizan tareas operativas de la organización”.

Alfaro y Lenon (2013) Aporta las siguientes características:

- “Apoya las tareas a nivel operativo de la organización.
- *Intensivos en entradas y salidas de transacciones.*
- *Recolectores de información porque cargan grandes bases de información para su explotación posterior.*
- *Garantizan el registro adecuado de las transacciones en forma correcta y en tiempo razonable.*
- *Sirven como plataforma para los otros tipos de sistemas de información”.*(p.7).

Sistemas de apoyo o soporte a las decisiones.

Tramullas (1997) dice: los

“Sistemas de apoyo a las decisiones (SSD o DSS, en inglés, *Decision Support Systems*): ofrecen información personalizada para el usuario determinado, cuándo debe tomar una decisión ante situaciones particulares semiestructuradas o no estructuradas”. (p.18).

Según Keenan (1997) un SSD

“...debe tener integrados tres componentes: la interfaz, la base de datos y los modelos. Es importante aclarar que estos componentes deben estar integrados y que la sola colección de las herramientas de software pueden ayudar a tomar decisiones, pero no constituyen un SSD”.

Para Toro (2004), la interfaz es la parte del SSD que interactúa con el usuario, debe ser amigable y usualmente incluye componentes gráficos; en este último caso, es llamada interfaz gráfica del usuario (GUI, por sus siglas en inglés). Son características de la interfaz:

- Debe ser fácil de usar, intuitiva, lógica y lo suficientemente general para minimizar errores y optimizar el tiempo durante el manejo de datos y su interpretación.

- Debe permitir el control de la base de datos y de los modelos del sistema.

- Debe concentrarse en decisiones específicas y tener como objeto ayudar, más que remplazar, a los usuarios.

La base de datos es la colección de datos que organiza de tal manera que su contenido pueda ser fácilmente encontrado, manejado y actualizado. La base de datos del SSD suministra datos para ser usados por los modelos del sistema y permite al usuario una oportunidad directa de explorar los datos relevantes al problema. Los modelos del SSD se encargan de la evaluación, análisis y correlación de la información de las diferentes fuentes disponibles en la base de datos. La “arquitectura” del SSD se refiere a cómo los componentes del sistema están integrados y conectados”. (Toro, 2004, p.60).

Por su parte, Pérez et al (2015) complementan enumerando características, tales como:

–Apoyan la toma de decisiones, por su misma naturaleza, son repetitivos y soportan decisiones no estructuradas que no suelen repetirse. Este tipo de sistemas puede incluir la programación de la producción, compra de materiales, flujo de fondos, proyecciones financieras, modelos de simulación de negocios, modelos de inventarios, etcétera.

–Suelen introducirse después de la implantación de los sistemas transaccionales más relevantes de la empresa, porque estos últimos constituyen su plataforma de información.

Sistemas estratégicos o de información a ejecutivos (EIS).

Al tope de la pirámide jerárquica según González et al (2011) indica que

“...los EIS (*Executive Information System*), es un sistema de información para ejecutivos que tiene aplicación a cualquier parte de la empresa, permite que la información que es recogida en la figura de informes se encuentre online a todo momento para que sea aprovechada por todas las partes de la organización.

Es uno de los sistemas de información disponibles, solo que se enfoca al nivel de la decisión. Es como tomar una fotografía de nuestra empresa en cualquier momento que se quiera permitiendo

llegar hasta cierto nivel de detalle para lograr encontrar las anomalías en alguna parte de la organización”.(P.174).

Pérez et al. (2015) Afirman la siguiente característica principal:

“Su función primordial no es apoyar la automatización de procesos operativos ni proporcionar información para apoyar la toma de decisiones”.

Mas adelante el mismo autor expresa

“Ahora bien, los sistemas de gestión de información (EIS) permiten.

- Comprender la marcha de las organizaciones desde un enfoque analítico (donde se quiere estar), evaluador (donde se está) y creativo (donde se podría estar).
- Develar oportunidades que merezcan ser explotadas y contrarrestar amenazas.
- Establecer los factores que resulten críticos y las necesidades asociadas con el EIS.
- Estudiar el impacto de los EIS en la posición del negocio y buscar nuevas oportunidades.

En el caso del diseño de un EIS precisa de:

- Un análisis previo de las necesidades de información de la organización.
- Un diagnóstico de la situación.
- Una auditoría de información que permita conocer los recursos de información disponibles y los que faltan, para qué y quiénes lo utilizan, qué valor se le añade en su uso, entre otros”.



Figura 03. Pirámide de jerarquía de los sistemas de información.
Fuente: Montelongo(2016)

En la Figura 03 se observa la jerarquía de los tipos de sistemas de información antes mencionados.

OLAP y OLTP

Para Roussel (2006)

“El procesamiento analítico en línea, (OLAP en inglés, *On-line Analytical Process*) permite obtener acceso a datos organizados y agregados de orígenes de datos empresariales, organiza subconjuntos de datos con una estructura multidimensional de manera que represente un significado especial o responda a una pregunta en particular”.

Silberschatz et al. (2006) agrega que

“Estas herramientas soportan el análisis interactivo de la información de resumen, soportando muchas tareas de agrupación

de datos que no pueden realizarse empleando las facilidades básicas de agregación y agrupamiento”.

Rosado (2010) indica que:

“Tradicionalmente, este sistema se clasifica según las siguientes categorías:

- ROLAP. Implementación que almacena los datos en un motor relacional. Típicamente, los datos son detallados, evitando las agregaciones y las tablas se encuentran normalizadas.
- MOLAP. Esta implementación almacena los datos en una base de datos multidimensional. Para optimizar los tiempos de respuesta, el resumen de la información es usualmente calculado por adelantado.
- HOLAP (Hybrid OLAP). Almacena algunos datos en un motor relacional y otros en una base de datos multidimensional.” (p.3).

Según Christiansen y Gaete (2014)

“OLTP (Procesamiento de Transacciones En Línea) facilita y gestiona aplicaciones orientadas a transacciones, típicamente para ingreso de datos y retención de transacciones. Para OLTP se basan en la arquitectura cliente-servidor ya que suelen ser utilizados por empresas con una red distribuida.

Las Características de los sistemas OLTP son:

- Velocidad de procesamiento de datos muy rápido.
- Permite estructurar de manera organizada la información contenida.
- Los datos son de tipo operacionales.
- Su procesamiento es de cliente servidor y software de intermediación (*middleware*).
- Revela la actualidad de los procesos de negocios.

Técnicas

- Visualización de tablas.
- Estructuración de la información.
- Extracción y almacenamiento de información en archivos planos.
- Conformación de cubos planos de información.
- En el procesamiento se puede involucrar técnicas de inserción, Modificación y borrado de datos.

- Objetivos
- Reducción de documentación, automatización de procesos y mejoramiento en el flujo de información.
- Obtención y pre visualización de ingresos o gastos de una organización.
- Proporciona métodos eficaces para el mejoramiento y estabilidad de una organización.
- Ampliación de las bases de datos existentes.
- Ejecución de procesos individuales, se hacen de forma más rápida". (p.8).

Tabla 01. Diferencias entre OLTP y OLAP

Características	OLTP	OLAP
Objetivos Principales	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.
Orientación	Orientado a la aplicación, hace cumplir las reglas del negocio.	Orientado al sujeto, se define en base a lo que el analista necesita ver.
Vigencia de los datos	Los datos se usan a medida que se van produciendo y dejan de ser importantes a corto plazo.	Se guardan los datos actuales y los históricos para poder realizar análisis comparativos.
Granularidad de datos	Es dada por controles definidos por la organización, así como por normas legales importantes.	Estará dado por el tipo de análisis que se requiera analizar.
Organización	Es normalizado.	Se basa en estructuras jerárquicas desnormalizadas modeladas de acuerdo a como se analizaran los datos.
Cambios en los datos	Modifica sus datos en forma constante porque maneja las transacciones de la empresa.	No tiene como objetivo la presentación de los datos en línea ni tampoco modificar datos originales, solo consultarlos.
Acceso y manipulación de los 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 DWS tiene una carga y acceso masivo de datos, no se realizan inserts, updates o deletes. La validación de datos o después de la carga.
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 o sistemas de archivos maestros preexistentes en las mismas organizaciones.

Fuente: Tibiño (2014).

Existen diferencias notables entre los sistemas transaccionales (OLTP) y los sistemas analíticos (OLAP) que se presentan en la

Tabla 01 anteriormente presentada.

Bases de datos

Para Migel y Piattini (1999).

“Una base de datos es una colección o depósito de datos integrados, almacenados en soporte secundario (no volátil) y con redundancia controlada. Los datos, que han de ser compartidos por diferentes usuarios y aplicaciones, deben mantenerse independientes de ellos, y su definición (estructura de la base de datos) única y almacenada junto con los datos, se ha de apoyar en un modelo de datos, el cual ha de permitir captar las interrelaciones y de restricciones existentes en el mundo real.

Los procedimientos de actualización y recuperación, comunes y bien determinados, facilitarán la seguridad del conjunto de los datos.” (p.66)

Greiner (2015) indica que

“Una base de datos se puede definir como un conjunto de información relacionada que se encuentra agrupada o estructurada. Desde el punto de vista de la informática, la base de datos es un sistema formado por un conjunto de datos almacenados que permiten el acceso directo a ello y un conjunto de programas que manipulen ese conjunto de datos.” (p.3).

Continúa el mismo autor que

“Se puede decir que una base datos es:

- Un conjunto de información relacionada que se encuentra agrupada o estructurada. Un archivo por sí mismo no constituye una base de datos, sino más bien la forma en que está organizada la información, es la que da origen a la base de datos.
- Es la colección de datos aparentes usados por el sistema de aplicaciones de una determinada empresa.

- Una colección de datos organizada para dar servicio a muchas aplicaciones al mismo tiempo al combinar los datos de manera que aparezcan estar en una sola ubicación.” (p.3).

Según Moreno (2000) se puede decir que las principales características de una Base de Datos son las siguientes:

-“**Independencia de datos:** es el concepto que más ha ayudado a la rápida proliferación del desarrollo de Sistemas de Bases de Datos. La independencia de datos implica un divorcio entre programas y datos.

-**Integridad en la información:** la integridad de la base de datos se refiere a la validez y la consistencia de los datos almacenados. Es una de sus principales cualidades y tiene por objetivo que sólo se almacena la información correcta.

-**Redundancia mínima:** se trata de usar la base de datos como repositorio común de datos para distintas aplicaciones, evitando almacenar los mismos datos varias veces.

-**Globalización de la información:** permite a los diferentes usuarios considerar la información como un recurso corporativo que carece de dueños específicos.

-**Acceso concurrente por parte de múltiples usuarios:** la base de datos pertenece a la empresa y puede ser compartida por todos los usuarios que estén autorizados.

-**Consistencia de datos:** si existen dos o más archivos con la misma información, los cambios que se hagan a éstos deberán hacerse a todas las copias del archivo.

-**Consultas complejas optimizadas:** la optimización de consultas permite la rápida ejecución de las mismas.

–**Seguridad de acceso y auditoria:** se refiere al derecho de acceso a los datos contenidos en la base de datos por parte de personas y organismos. El sistema mantiene el control de acceso a la base de datos, con el objeto de saber qué o quién realizó una determinada modificación y en qué momento.

–**Acceso a través de lenguajes de programación estándar:** se refiere a la posibilidad ya mencionada de acceder a los datos de una base de datos mediante lenguajes de programación ajenos al sistema de base de datos propiamente dicho”.

Sistema Manejador de Bases de Datos

Un Sistema Manejador de Bases de Datos (SMBD) que según Elmasri y Navathe (2002) “es un sistema de software de propósito general que facilita los procesos de definición, construcción, manipulación y compartición de bases de *datos entre varios usuarios y aplicaciones*”. (p.20).

Por su parte, Date (2001) afirma que “Un sistema de bases de datos es básicamente un sistema computarizado cuya finalidad general es almacenar información y permitir a los usuarios recuperar y actualizar esa información con base a peticiones.” (p.5).

Según Silberschatz et al (2002) “...es un conjunto de programas, procedimientos, lenguajes, entre otros. Que suministra a los usuarios no informáticos, analistas, programadores, administradores de base de datos, los medios necesarios para describir, recuperar y manipular los datos almacenados en la base de datos, permitiendo mantener su integridad, confidencialidad y seguridad.

Entre las principales funciones de los SMBD se tienen:

–Almacenar, recuperar y actualizar los datos en el disco.

–Manejar la integridad y consistencia de los datos cuando estos son actualizados.

–Coordinar la concurrencia de acceso de los diferentes usuarios a la base de datos, garantizando que los datos no se dañen.

Un sistema manejador de base de datos es esencial para poder manipular y darle un adecuado funcionamiento a una base de datos”.(p.445).

Data Warehouse

Según Elmasri, y Navathe (2002) “un *data warehouse* o almacén de datos también es una colección de información unido con un sistema de soporte”. (p.851). Por su parte Inmon (1990) “un almacén de datos es una colección de datos orientada al sujeto, integrada, no volátil y de tiempo variable para el soporte de las decisiones de los directivos”. (p.35). Adicionalmente Gil y Berriel (2006) menciona que

“...como diferencia fundamental con las bases de datos tradicionales manejadas en las empresas, habitualmente basadas en las técnicas OLPT, reside en que los datos no son volátiles, con lo que se está refiriendo a bases de datos de una dimensión superior a las anteriores que permiten registrar los datos de modo permanente”. (p.4).

Continúa Inmon (1990) enumerando las siguientes características de un almacén de datos:

Tabla 02. Características de un almacén de datos.

Orientado a Temas	<i>La información se clasifica con base a los aspectos que son de interés para la organización, se busca centrar la información en las actividades básicas de la empresa (ventas, compras, producción, entre otros.) y no en los procesos (gestión de pedidos, facturación, inventario, entre otros.) diferenciándose estos en el nivel de detalle.</i>
Integrado	<i>El objetivo principal de la integración es reunir todos los datos de las diferentes fuentes y agruparlas en un modelo coherente. Los datos recolectados en el almacén de datos deben ser integrados en estructuras consistentes, es decir, se eliminan las inconsistencias existentes en las fuentes de datos (internas y externas) operacionales. La información suele estructurarse en varios niveles de detalle para adaptarse a las necesidades del usuario final.</i>
No Volátil	<i>La información que se encuentra en el almacén de datos no se modifica ni se elimina, es decir, sólo se agrega, por lo tanto la información es permanente. Esta característica elimina la necesidad de tener mecanismos de control de concurrencia y recuperación de datos.</i>
Variante en Tiempo	<i>Los datos recolectados en un almacén pertenecen a un determinado periodo de tiempo. Dichos datos deben actualizarse cada cierto tiempo permitiendo al usuario final visualizar las variaciones que ha tenido el negocio en el tiempo, identificar tendencias y realizar proyecciones con el fin de mejorar el rendimiento del negocio. Por ejemplo, se almacenan los productos vendidos en enero, febrero, marzo, entre otros. Luego es posible comparar y validar en que mes hubo mayores ganancias y cuáles fueron las estrategias a seguir para lograr dicha meta.”</i>

Fuente: Inmon (1990)

Datamart

Según Gil y Berriel (2006). un Datamart, “...es una versión limitada de un *Data Warehouse* que es establecido para un grupo de usuarios o un área y está disponible para un número pequeño de personas.

Los *Data Warehouses* se construyen para servir a las necesidades de información de una organización completa, mientras que los *Datamarts* sirven a las necesidades de información de un grupo o área específica de la misma.

Ambas son herramientas que incorporan las mismas funciones, si bien la capacidad de un *Data Warehouses* es superior en cuanto a que permite almacenar mayor volumen de datos que un *Datamart*, configurándose como

un almacén central para toda la información del negocio. Con ella, se logra reducir la redundancia de los datos almacenados en diferentes departamentos mientras que al mismo tiempo se facilita a muchos usuarios una amplia variedad de análisis.

Del mismo modo, las dos tecnologías ayudan a incrementar la seguridad de los datos almacenados y a la vez otorgan una gran operatividad a los usuarios de los SI, pero se diferencian básicamente no sólo en cuanto al tamaño sino también en el coste de implantación. En este sentido, un *Datamart* resulta menos costoso de construir que un proyecto de *Data Warehouse*; requiere menos ingeniería, menor coste del hardware y se necesitan menos recursos para desarrollarlo y mantenerlo.

Por este motivo, muchos expertos recomiendan comenzar con un *Datamart* como paso previo para vincular a la empresa con esta tecnología sin asumir el mayor riesgo que implica el alto coste de implantación de un *Data Warehouse*". (p.5).

Hernández (2005) contribuye con el concepto de *Datamart* "...es una extensión natural del *Data warehouse* y está enfocado a un departamento o área específica, como por ejemplo los departamentos de Finanzas o *Marketing*, permitiendo así un mejor control de la información que se está abarcando".(p.4).

ETL

Para Rivas (2012), "Los sistemas ETL (*Extract, Transform and Load* - Extracción transformación y carga) tienen la capacidad de extraer datos desde diferentes tipo de sistemas, transformarlo y cargarlos a un sistema destino que no necesariamente es el *data warehouse*. En la mayoría de los casos existen sistemas intermedios (*stage*) en los que es posible verificar los

datos extraídos por medio de diferentes procesos de calidad de datos (DQ), auditoría y control.

Finalmente, todos los datos transformados y consolidados son almacenados en un depósito de datos dimensional (8DDS) el cual es consultado por diferentes procesos para los que se dispuso el *data warehouse*".(p.13). Espinosa (2010) afirma que "las herramientas ETL, deberían de proporcionar, de forma general, las siguientes funcionalidades:

- *Control de la extracción de los datos y su automatización*, disminuyendo el tiempo empleado en el descubrimiento de procesos no documentados, minimizando el margen de error y permitiendo mayor flexibilidad.

- *Acceso a diferentes tecnologías*, haciendo un uso efectivo del hardware, software, datos y recursos humanos existentes.

- *Proporcionar la gestión integrada del Data Warehouse y los Datamarts existentes, integrando la extracción, transformación y carga para la construcción del Data Warehouse corporativo y de los Datamarts.*

- *Uso de la arquitectura de metadatos*, facilitando la definición de los objetos de negocio y las reglas de consolidación.

- *Acceso a una gran variedad de fuentes de datos diferentes.*

- *Manejo de excepciones.*

- *Planificación, logs, interfaces a schedulers de terceros*, que permitirán llevar una gestión de la planificación de todos los procesos necesarios para la carga del DW.

- *Interfaz independiente de hardware.*

- *Soporte en la explotación del Data Warehouse".*

Para Cibertec (2015), "La arquitectura de un *Data Warehouse* está conformado por las aplicaciones, los componentes funcionales del *Data*

Warehouse (adquisición, almacenamiento y acceso) y las infraestructuras (técnica y operativa).

Las aplicaciones analíticas.

Se pueden visualizar tres desde el punto de vista de procesos: los componentes funcionales que son parte del proceso son la adquisición de datos, el almacenamiento y el acceso por parte de usuarios finales.

- **Adquisición:** Consiste en recoger los datos útiles del sistema de producción. Se deben identificar los datos que sean necesarios para atender los requerimientos de información, luego planificar las extracciones con el fin de evitar saturación en la red, o afectar al sistema transaccional de producción. Los procesos de extracción deben estar sincronizados con la finalidad de garantizar la integridad de la información. Los problemas que surgen al hacer esta sincronización pueden ser muy complejos.

Después de extraer los datos del sistema transaccional, estos se deben “preparar” para adecuarlos a la forma del *Data Warehouse*. Esta “preparación” incluye la correspondencia de los formatos, la limpieza, la transformación y la agregación en muchos casos.

La carga es la última fase de la adquisición de datos, esta fase es particularmente importante sobre todo si se trata de volúmenes muy grandes.

- **Almacenamiento:** El componente básico del soporte del almacenamiento es el DBMS (*DataBase Manager System*). El DBMS o motor de base de datos debe tener las características que le permitan responder eficientemente a las exigencias de las consultas analíticas. Para lograrlo debe contar con diversos recursos como el paralelismo, la optimización del indexado con la finalidad de acelerar las consultas agregadas, ordenamientos y agrupaciones.

En relación con los tipos de datos, generalmente, se almacenan en formatos relacionales; sin embargo, frente a la gran cantidad de datos en forma de documentos, imágenes, audio y video, los DBMS están evolucionando en el sentido de permitir la gestión de estos tipos de datos. Esta evolución se ve reforzada aún más con la llegada de Internet.

- **Acceso:** El acceso al *Data Warehouse* se da mediante herramientas o aplicaciones de tipo Cliente/servidor o herramientas que pueden utilizarse desde el Web. Hay una gran variedad de herramientas en el mercado y el número de aplicaciones de acceso que se pueden desarrollar es también muy grande. Sea cual sea el tipo de herramienta, tendrá que adaptarse a las exigencias del usuario y su manera de trabajar.

En el mundo de la decisión, el análisis es también un proceso iterativo y los resultados de la consulta actual influyen a menudo en la consulta siguiente. Esto se puede resumir en la siguiente frase: “Dame lo que te pido y luego podré decirte lo que realmente quiero “.

- **Infraestructuras:** Para hacer frente a las necesidades de *Data Warehouse*, el papel de la informática es definir e integrar una arquitectura sobre la que implementará el *Data Warehouse*. Se deben considerar dos niveles de infraestructura en un *Data Warehouse*: la infraestructura técnica o conjunto de componentes materiales y programas, y la infraestructura operativa o conjunto de procedimientos y servicios para administrar los datos, gestionar los usuarios y utilizar el sistema.

Por un lado, la infraestructura técnica se compone de productos que implementan las tecnologías elegidas, integrados en un conjunto coherente y homogéneo. Por otro lado la infraestructura operativa se compone de todos los procesos que permiten, a partir de los datos de producción, crear y gestionar el *Data Warehouse*”. (p12).

Componentes DW/BI

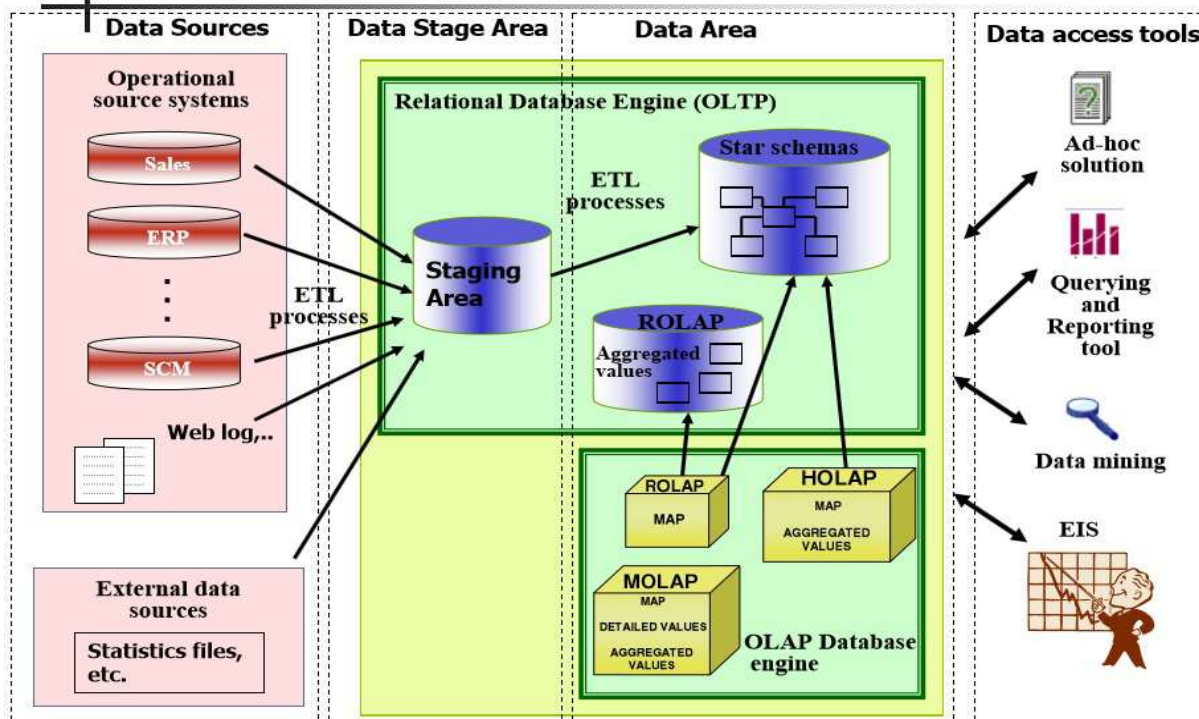


Figura 04. Arquitectura de un Data Warehouse.
Fuente Zorrilla (2010).

En la figura 04 se aprecia la arquitectura de DW/BI.

Modelado dimensional

Según Kimball et al (1998) "Se entiende por modelado dimensional como, una técnica de diseño lógico el cual busca presentar los datos de manera estándar e intuitiva y que permite un acceso de alto rendimiento"(p.5.1). También indica que: "...es una técnica de diseño de bases de datos destinadas a apoyar a las consultas de los usuarios finales en un almacén de datos. Está orientado a la comprensibilidad y rendimiento". Según él, "a pesar de que el modelo Entidad-Relación es de gran utilidad para la captura de transacciones, se debe evitar presentar al usuario final.

Por lo tanto, busca presentar la información de una manera estándar, sencilla y sobre todo intuitiva para los usuarios finales, además de permitir acceso rápido a la información por parte de los manejadores de bases de datos". (p.5.2).

Hecho

Rasmussen et al (2002) lo definen como "En el área de almacenes de datos se puede entender hecho como, una medida o valor de rendimiento (indicador) que representa un evento en una organización. Por lo general, estas medidas toman valores numéricos o agregados. ." (p.53). Los hechos agregados son hechos que derivan de funciones como la suma, promedio, máximo, mínimo, entre otros. Por ejemplo, se podría tener como hecho agregado el total de activos de una empresa, es decir, la suma de todos sus activos (circulantes, fijos, entre otros.)

Dimensión

Rasmussen et al (2002) indican que Se puede definir dimensión "...como, una entidad independiente en un modelo dimensional, el cual agrupa la información por categorías (clientes, productos, fechas, entre otros.) ". (p.54), y se utiliza como mecanismo de acceso a dicha información. Estas dimensiones, por lo general, suelen representar el contexto de un hecho, respondiendo a preguntas como quién participó, cuándo y dónde pasó, y su tipo

Mirabella (2011) aporta que "en los almacenes de datos estas dimensiones se utilizan para "rebanar y cortar" la información que en ella se encuentra, es decir, acceder a diferentes vistas agrupadas bajo ciertos criterios. Estas se caracterizan principalmente por proporcionar filtrado, agrupación y etiquetado de datos. Por otro lado, existen las llamadas claves

subrogadas que no son más que nuevas claves primarias que el almacén o bodega de datos, usa para poder relacionar las dimensiones con los hechos”. (p.51).

Jerarquía

Kimball y Ross (2002) dicen que

“cuando se hablan de cubos de información un concepto importante que siempre tiene que estar presente son las Jerarquías. Estas se pueden definir como una serie de relaciones, padre e hijo, que poseen los atributos de una dimensión. Además, el conjunto de una serie de atributos padres pueden generalizarse como los hijos de una relación padre que esté a un nivel superior. Entonces, se puede decir que cada elemento en una dimensión se puede resumir utilizando una jerarquía. Por ejemplo, día -> semana -> mes -> año (Fecha)”. (p.6.7).

Granularidad

Kimball y Ross (2002) dicen que

“El concepto de granularidad se refiere al nivel de detalle con que se almacenan los datos en una tabla de hechos o dimensiones. Esta tiene una relación estrecha con las jerarquías, ya que cuando se habla de jerarquías implícitamente se está definiendo la granularidad de los datos, en donde la granularidad más alta (grano grueso) empieza en el nivel superior de la jerarquía hasta llegar al nivel más bajo (grano fino) donde se encuentran los datos en detalle”. (p.5.12).

Cubo

Kimball y Ross (2002) afirman que

“...también conocidos como cubos de información, son estructuras de datos que permiten analizar de manera rápida los datos. Estos cubos permiten manipular y detallar los datos desde múltiples perspectivas que varían dependiendo de criterios construidos para su consulta. Estos cubos están compuestos principalmente por hechos, los cuales no son más que indicadores que representan algún evento de la organización, y dimensiones. Los hechos y dimensiones se obtienen a partir de tablas de hechos y tablas de dimensiones respectivamente.

Generalmente los metadatos de estos cubos son creados a partir de esquemas, como estrella o copo de nieve, de tablas pertenecientes a base de datos relacionales”. (p.7.9).

Deppen et al (2009)

“En una base de datos multidimensional la información se representa como matrices multidimensionales, cuadros de múltiples entradas o funciones de varias variables sobre conjuntos finitos denominadas “Cubos”.

Una instancia de un cubo, queda determinada por un conjunto de datos para cada eje y un conjunto de datos para la matriz”.

El esquema de un cubo queda determinado dando a conocer sus ejes con sus respectivas estructuras y la estructura de los datos que se presentan en cada celda de la matriz. Se asume que todas las posiciones de la matriz tienen datos con igual estructura (datos uniformes).

A los ejes se les llama Dimensiones, al dato que se presenta en la matriz se le llama Medida y a los elementos del producto cartesiano de los ejes se le llama Coordenadas.

La matriz definida puede ser dispersa.

Dimensiones con Jerarquías.

Normalmente las dimensiones se estructuran en jerarquías.

A cada nivel de una jerarquía se le llama Nivel de Agregación o simplemente Nivel. Una dimensión puede tener más de una jerarquía”. (p.16).

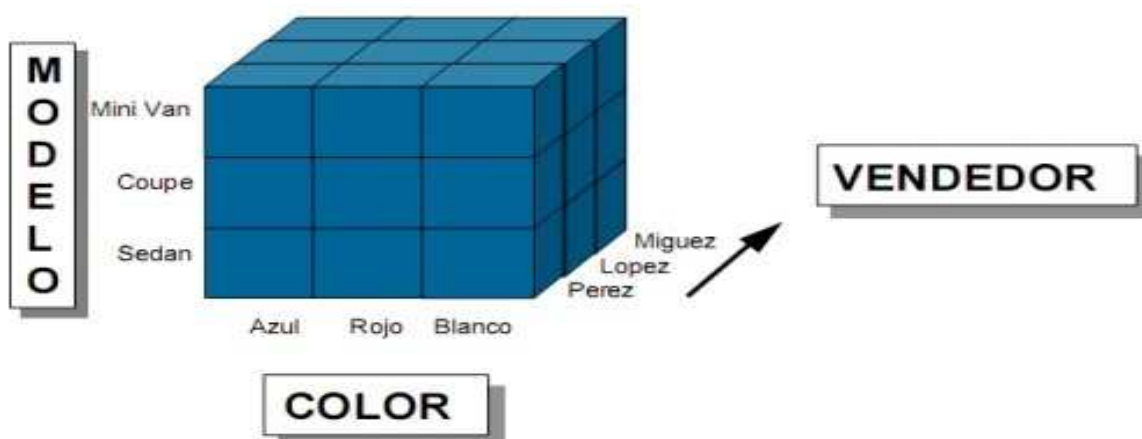


Figura 05. La estructura básica en el modelo multidimensional: el Cubo.
Fuente: Deppen et al (2009).

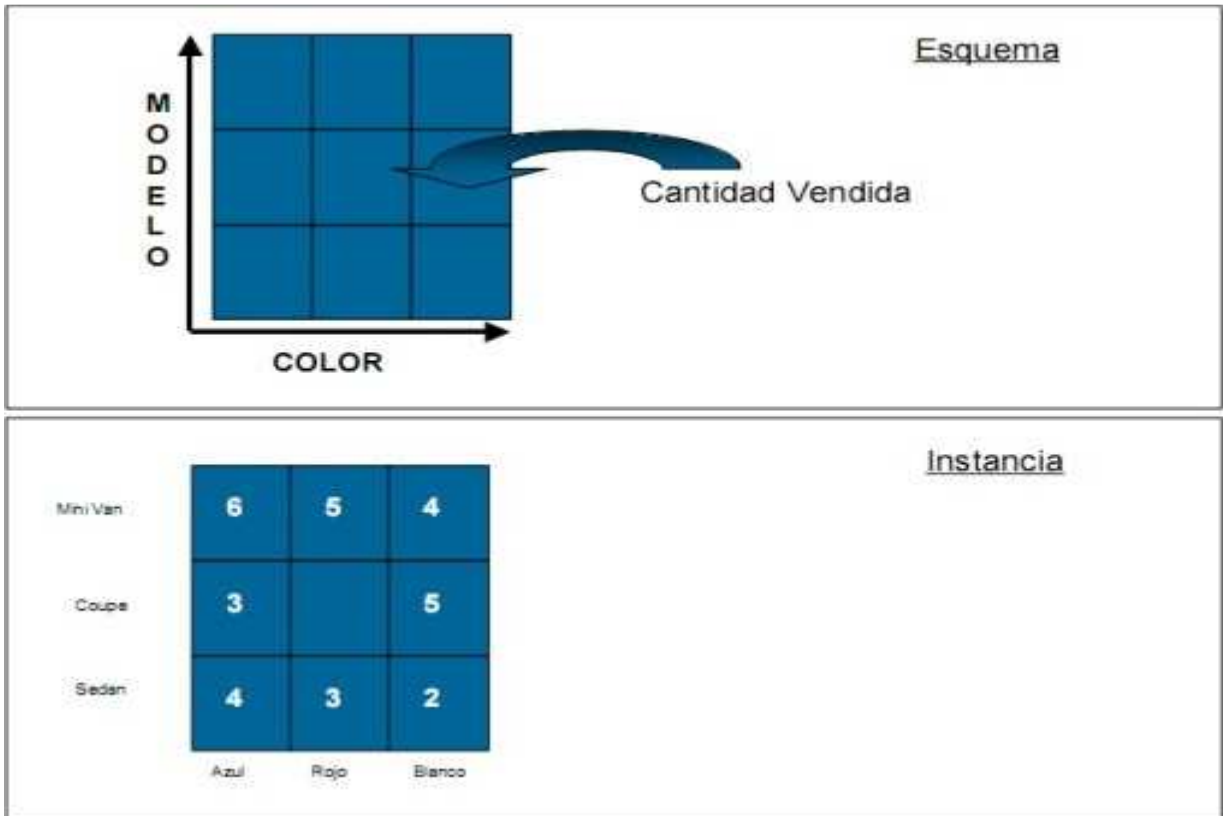


Figura 06. Esquema e instancia de un cuadro de doble entrada
Fuente: Deppen et al (2009).



Figura 07. Dimensión sin jerarquías alternativas.
Fuente: Deppen et al (2009).

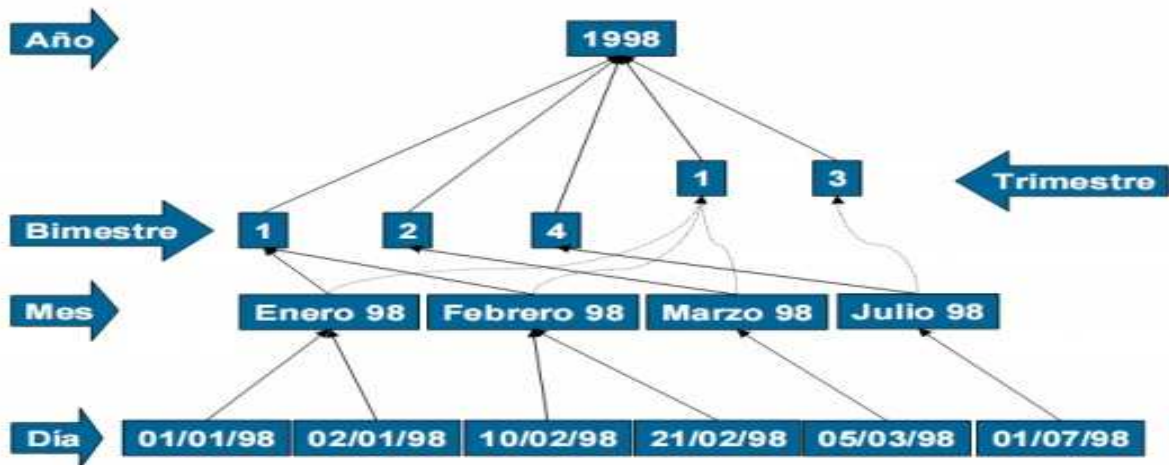


Figura 08. Dimensión con jerarquías alternativas.
Fuente: Deppen et al (2009).

En la Figura 05 se muestra un ejemplo de cubo. En la figura 06 para ver el esquema e instancia de un cuadro. En la Figura 08 para observar una dimensión con jerarquías alternativas y en contra posición ver Figura 07 para una dimensión sin jerarquías alternativas.

Tabla de Dimensiones

Mirabella (2011) afirma que en los almacenes de datos estas tablas "...son utilizadas principalmente para crear dimensiones estandarizadas, las cuales puedan ser compartidas a través de los diferentes almacenes de datos existentes en una organización. Esto permite garantizar consistencia e integración de los datos y reducir los tiempos de desarrollo.

Por lo tanto, se tiene que una tabla de dimensión es una tabla que contiene una clave primaria que la identifica y una serie de atributos descriptivos que se utilizan para filtrar y agrupar los datos bajo ciertos criterios. Los datos contenidos en estas dimensiones son utilizados por los

indicadores en la tabla de hechos para crear las vistas que serán mostradas a los usuarios finales.

Por otra parte, las dimensiones pueden ser medidas de diferentes maneras. Esto se logra determinando el nivel de detalle (granularidad) en que se van almacenar los datos de las dimensiones”. (p.53).

Por ejemplo, en la Figura 09. Se observa una dimensión zona geográfica con cinco niveles de jerarquía.

Continúa el mismo autor, haciendo referencia al ejemplo

“...en dicha dimensión, se tiene que el indicador será quien determine la granularidad. Cuanto más detallado sea, más fino será el nivel de granularidad (grano fino) y viceversa (grano grueso)”. (p.53).



Figura 09. Granularidad de la dimensión, con una jerarquía de cinco niveles.
Fuente: Mirabella (2011).

Se aprecia un ejemplo en la figura anterior de granularidad.

Tabla de Hechos

Kimball y Ross (2002). Indican que “Es una tabla que contiene una clave primaria que lo identifica, un grupo de claves foráneas asociadas a las dimensiones con las que se relaciona y un conjunto de indicadores que se utilizarán para representar los eventos a modelar. En otras palabras, lo que se

busca con las tablas de hechos es representar un proceso de negocio. Además, en ella se almacena información transaccional que varía en el tiempo”. (p.5.4).

Agrega Mirabella (2011):

“Por lo general, cuando se habla de tablas de hechos se suele hacer énfasis en los tipos de hechos que estas almacenan. Estos hechos se suelen clasificar en:

–**Aditivos:** Son aquellos que pueden ser resumidos a través de cualquiera de las dimensiones en la tabla de hechos.

Se observa (ver Figura 10) que monto es un hecho aditivo, ya que este se puede agregar usando cualquiera de las tres dimensiones, y tomar un valor significativo.

–Monto + Fecha X + Fecha Y = Total ventas entre las fechas X e Y.

–Monto + Tienda = Total ventas de la tienda Z.

–Monto + Producto = Total ventas del Producto W.

–**Semi-aditivos:** Son aquellos que pueden ser resumidos por algunas, pero no todas, las dimensiones en la tabla de hechos.

Se observa (ver Figura 11) que saldo es un hecho semi-aditivo, esto se debe a que si se junta con cuenta se podrá conocer el saldo de una cuenta Y. Sin embargo, no tendría sentido juntar saldo con fecha ya que no arrojaría ningún valor significativo.

–**No aditivos:** Son aquellos que no pueden ser resumidos por ninguna de las dimensiones presentes en la tabla de hechos.

Siguiendo con el ejemplo anterior (ver Figura 11), Prueba ácida es un hecho no aditivo ya que no tiene sentido juntarla con ninguna de las dimensiones. Nota: Prueba Ácida es un indicador que mide la capacidad inmediata que tiene una empresa para cancelar sus deudas a corto plazo.

Una de las características más importantes de las tablas de hechos es el nivel de detalle (granularidad) de la información que almacenan. Este nivel de detalle permite a los indicadores comparar valores específicos y también generalizarlos.

Por otra parte, las dimensiones asociadas a esta tabla de hecho son las que determinan el alcance que esta puede llegar a tener. Es decir, a partir de los datos contenidos en estas dimensiones, y agregando criterios de consulta, se construirán los indicadores o medidas existentes en dicha tabla (ver Figura 12). Por ejemplo, se podría tomar los montos de la dimensión Producto y crear un indicador que represente las ganancias de dicho producto.

Con esto se podrían analizar las ganancias de dicho producto en distintos meses del año.

Las tablas de hechos se pueden clasificar en:

–**Acumulativas:** Este tipo de tablas describe lo que ha pasado en un periodo de tiempo. Por ejemplo, las ganancias producidas por un producto por tienda por día. Los hechos en este tipo de tablas son más que todo aditivos.

–**Fotografías (Snapshot):** Este tipo de tablas describe el estado de las cosas en un periodo particular de tiempo. Los hechos representados en este tipo de tablas son semi-aditivos y no aditivos”. (p.54).

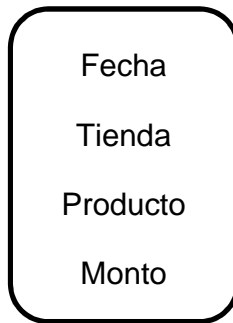


Figura 10. Tabla de Hechos - Ventas.
Fuente: Mirabella (2011).

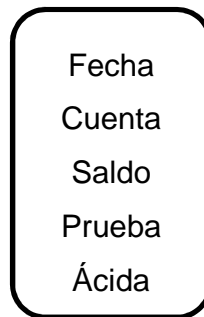


Figura 11. Tabla de Hechos – Indicador.
Fuente: Mirabella (2011)



Figura 12. Dimensiones
Fuente: Mirabella (2011).

Las figuras 10, 11 y 12 dan ejemplo de tablas de hechos.

Formas de Modelado

Mirabella (2011) aporta:

- **Esquema Estrella:** Es uno de los modelos más utilizados a la hora de diseñar almacenes de datos en la Inteligencia de Negocios (BI). Estos esquemas son optimizados con el fin de realizar consultas sobre grandes cantidades de datos, apoyar las consultas de los cubos OLAP, aplicaciones de Inteligencia de Negocios y analíticas.

Es llamado así debido a que la forma de su diagrama se asemeja a una estrella (Figura 13). En el centro de este modelo se encuentra la tabla de hechos, con sus respectivas claves foráneas y medidas, y en las puntas se encuentran las tablas de dimensiones que componen al hecho. Uno de los aspectos más notables de este modelo es que la tabla de hechos se encuentra en tercera forma normal (3FN), mientras que las dimensiones están des-normalizadas. Se caracteriza principalmente por:

- Modelar una jerarquía de dimensiones lineales.
 - Fácil modificación.
 - Simplicidad en la navegación.
 - Simular la vista que los usuarios finales observan.
- Ser esquemas físicos y no lógicos.

-**Esquema Copo de Nieve:** A diferencia del anterior, es un esquema en el cual las dimensiones son almacenadas en múltiples tablas de dimensiones, eliminando así la redundancia de datos (Figura 14). Por lo general, las tablas de dimensiones en este modelo están representadas por relaciones normalizadas, en tercera forma normal (3FN).

La composición de este esquema es parecida al estrella, en donde se tiene una tabla de hechos central y las tablas de dimensiones asociadas a esta, diferenciándose en que estas tablas de dimensiones, a su vez, tienen

otras dimensiones con las que se relacionan. De esta manera su diagrama esquematiza un copo de nieve.

Las aplicaciones de Inteligencia de Negocios que utilizan una arquitectura de procesamiento analítico en línea relacional tienden a mejorar su rendimiento con almacenes de datos que utilicen este esquema.

Una de las complicaciones de este modelo se presenta a la hora de extraer o consultar datos de las tablas. Esto es debido a que existen casos en los cuales hay que vincular muchas tablas, lo cual resulta bastante complejo de realizar y difícil de mantener". (p.57).

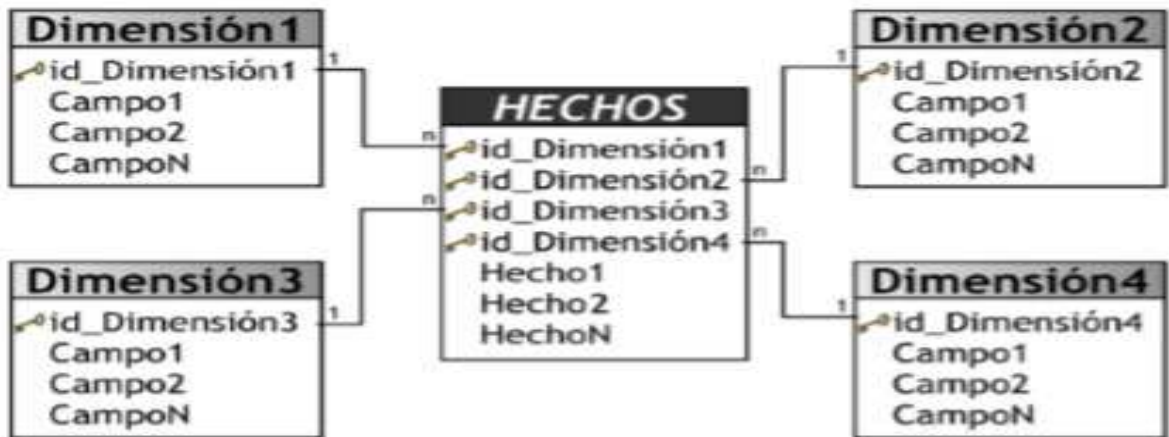


Figura 13. Ejemplo de un modelo de datos en estrella.
Fuente: Mirabella (2011).

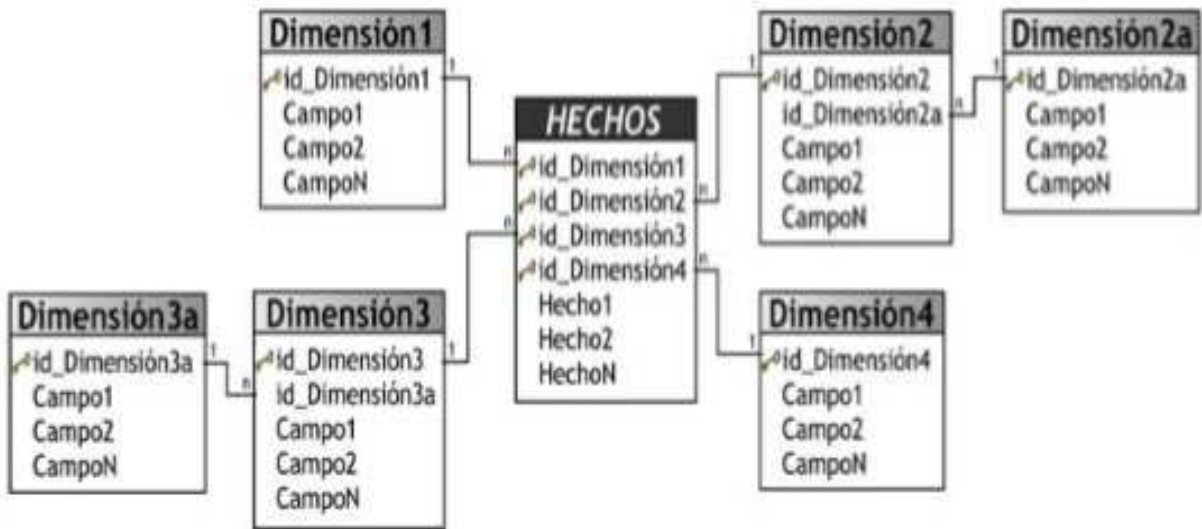


Figura 14. Ejemplo de un modelo de datos en copo de nieve.
Fuente: Mirabella (2011).

Las figuras 13 y 14 ejemplifican los modelos de datos descritos anteriormente.

Pasos para el diseño de un modelo dimensional

Mirabella (2011) indica los siguientes pasos para el diseño de un modelo dimensional:

“Paso 1. Determine el área del negocio y sus procesos: El primer paso para el diseño es determinar los procesos de negocio o el evento que será medido con este modelo. El análisis del negocio proporcionará al diseñador el conocimiento del área a modelar y así atender debidamente los requerimientos del usuario final. Para identificar estos procesos se utilizan generalmente técnicas como entrevistas, mesas de trabajo, encuestas, entre otras.

Paso 2. Determinar la Granularidad Una vez que el proceso de negocio es identificado, el equipo de diseño debe declarar el grano de la tabla de hechos. Es imprescindible determinar lo que representa una fila en la tabla de hechos del modelo dimensional del negocio propuesto.

Paso 3. Identifique las Dimensiones Luego de determinar la granularidad, identificar las dimensiones es una labor sencilla. El grano en sí, a menudo determinará un número inicial de dimensiones que se complementaran con el resto que sean necesarias.

Paso 4. Identificar los hechos El paso final en el proceso de diseño es seleccionar cuidadosamente los hechos o indicadores que son aplicables al proceso de negocio. Cada hecho debe ser fiel al nivel de granularidad establecida en el paso número 2, no se mezclan indicadores de otro periodo o nivel de detalle que no coincida con el grado declarado”.(p.91).

Inteligencia de negocios

Loshin (2003). Afirma que

“La inteligencia de negocios son procesos, tecnologías y herramientas necesarias para transformar datos en información, información en conocimiento, y conocimiento en planes de negocios rentables.”

Para Cano (2007),

“Inteligencia de negocio es un proceso interactivo para explorar y analizar información estructurada sobre un área (normalmente almacenada en un almacén de datos), para descubrir tendencias o patrones, a partir de los cuales derivar ideas y extraer conclusiones. El proceso de inteligencia de negocio incluye la comunicación de los descubrimientos y efectuar los cambios. Las áreas incluyen clientes, proveedores, productos, servicios y competidores”. (p.13).

Castrillo, A. y Dugarte (2011) enumeran las características de una solución BI:

- “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 provenga, para así dar la sensación 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 y/o *datamarts*.
- Vistas gráficas e interactivas para la presentación de información analítica: A través de cuadros de mandos integrales y 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”. (p.26).

Beneficios

Mirabella (2011) lista los beneficios de una solución de Inteligencia de Negocios:

- “Ayuda a la alta gerencia a planear y obtener pronósticos, presentando una descripción clara de los procesos de negocio existentes en la organización.

- Control de costos, soportado por la solución de Inteligencia de Negocios implantada en la organización, y que es de uso común para todos los departamentos, de esta manera facilita el manejo de los distintos procesos de negocio existentes.

- Mejora la colaboración y calidad de las decisiones, dado que la información se comparte en todos los niveles de la organización.

- Proporciona una visión más profunda del negocio, a través de herramientas especializadas que facilitan el análisis de los datos, búsqueda de comportamientos y tendencias en las ventas, estudio de la evolución del mercado y cambios en el consumo o producción, como por ejemplo: cuadros de mando integrales, informes, minería de datos, almacenamiento analítico y tableros de instrumentos.

- Optimiza el rendimiento de los sistemas, a través de transformaciones en los sistemas transaccionales y analíticos, para lograr una mejora en las consultas de alto nivel.

- Realiza seguimientos de los planes estratégicos, a través de los cuadros de mando integrales permite crear, manejar y monitorear los objetivos estratégicos que se han propuestos en los planes, de esta manera, se logran conseguir a tiempo posibles desviaciones y corregirlas.

- Obtiene el valor real de las aplicaciones de gestión, mediante el acceso a la información que tienen todos los empleados de la organización, permitiendo que esta se presente de forma actualizada, adecuada e integrada”. (p.40).

Rasmussen et al (2002) agrega que además de los beneficios ya listados, una solución de Inteligencia de Negocios permite a una organización obtener varios tipos de información, como:

- “Comparativa, comparar las ventas de la fecha actual con fechas anteriores.

- De tendencias, cuáles productos se vendieron más en cuáles regiones.

- Veraz, conocer si las ventas cumplieron con los objetivos propuestos.

- Rentable, saber si vale la pena seguir invirtiendo en cuáles producto y cuáles servicios”. (p.24).

Narra Hernández (2015) que la inteligencia de negocios “...no es un tema nuevo es algo que se viene tratando desde los años 50, cuando Hans Peter Luhn científico de IBM, escribió un artículo donde hablaba de un sistema dedicado a la Inteligencia de Negocios. En dicho artículo describe la Inteligencia como la habilidad para percibir la interrelación de los hechos presentados de manera tal que orienten la acción hacia una meta deseada.

Para los años 60, se trabajaba con la información almacenada en archivadores, existían miles de carpetas, con datos de todo tipo, y en muchas ocasiones se perdía información importante, manejar esta información para aquel entonces era una tarea compleja. Para esta década, con la aparición del computador, también surgieron las bases de datos creadas por Edgar Frank Codd lo cual cambió el concepto de almacenar la información en carpetas físicas para hacerlo ahora en el computador.

A principio de los años 70, varias empresas empezaron a crear aplicaciones empresariales empleando esas bases de datos que muchas empresas ya habían comenzado a utilizar. Además, las aplicaciones permitieron realizar “*dataentry*” en los sistemas aumentando la información disponible pero no tenían un fácil y rápido acceso a esa información.

Luego para la de década de los 80 por primera vez aparece el término *Data Warehouse* desarrollado por Ralph Kimball y Bill Inmon y los primeros sistemas (SAP) que podían generar reportes para el usuario, pero seguía siendo algo difícil de utilizar por parte de los usuarios finales y surgió un nuevo problema, potentes sistemas de bases de datos sin aplicativos que permitieran una fácil explotación de las mismas.

Luego, en 1989, Howard Dresner utiliza el término Inteligencia de Negocios para describir que son un conjunto de métodos y procesos para mejorar la toma de decisiones de las organizaciones con el uso de sistemas de apoyo basado en hechos.

No es hasta finales de la década de los 90s cuando el uso de las herramientas de Inteligencia de Negocios comienza su popularización. Existían múltiples aplicaciones para acceder a la información y el uso de los almacenes de datos estaba centrado en el soporte de los datos.

Finalmente, a partir de los años 2000s y hasta la actualidad, las aplicaciones de Inteligencia de Negocios se han consolidado en pocas

plataformas como Oracle, SAP, IBM y Microsoft. Además, el uso de los almacenes de datos se utiliza para proveer datos y facilitar su análisis más que para el soporte de los mismos”. (p.40).

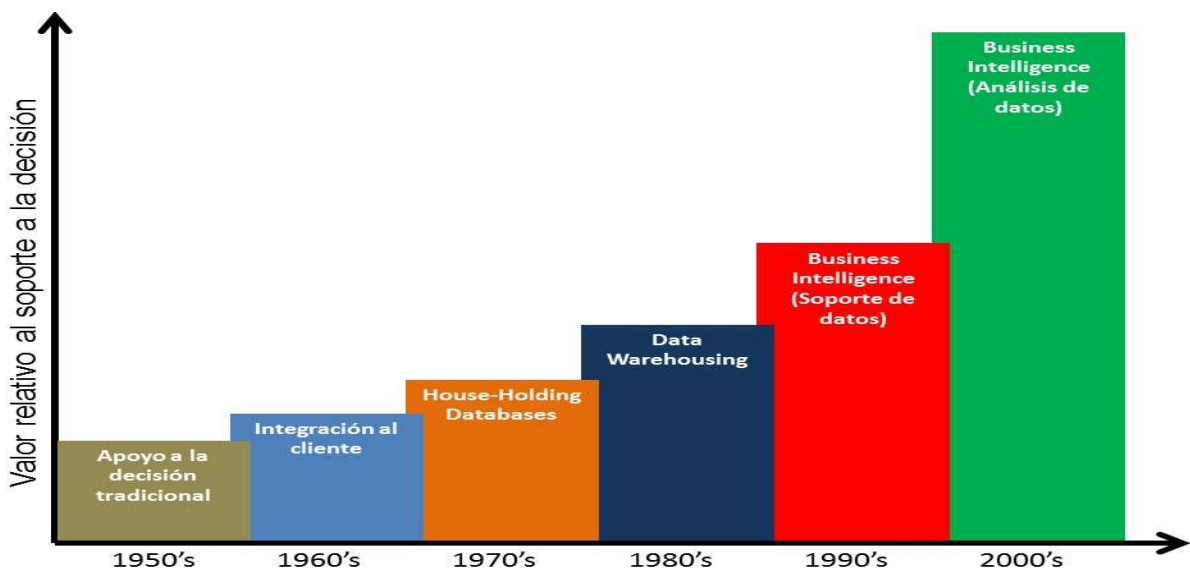


Figura 15. Evolución del *Business Intelligence*.
Fuente: Medina et al (2009)

En la figura 15 se aprecia un gráfico con aspectos de la evolución de BI.

Para Vásquez (2013) “Los sistemas de inteligencia de negocios han evolucionado a la fecha, incorporando nuevas y poderosas funcionalidades de análisis y presentación de información para distintos niveles de usuarios y áreas en una empresa.

La oferta actual de soluciones de inteligencia de negocios incluye reportes y tableros de mando, a través de los cuáles se puede responder a interrogantes tales como ¿Qué pasó? y ¿Qué está pasando?, y además agrega funcionalidades analíticas mejoradas para realizar consultas en línea y ejecutar modelos estadísticos de simulación y optimización que ayudan a

responder a las preguntas ¿Por qué pasó? y ¿Qué pasará? Ver figura 16”. (p.63).

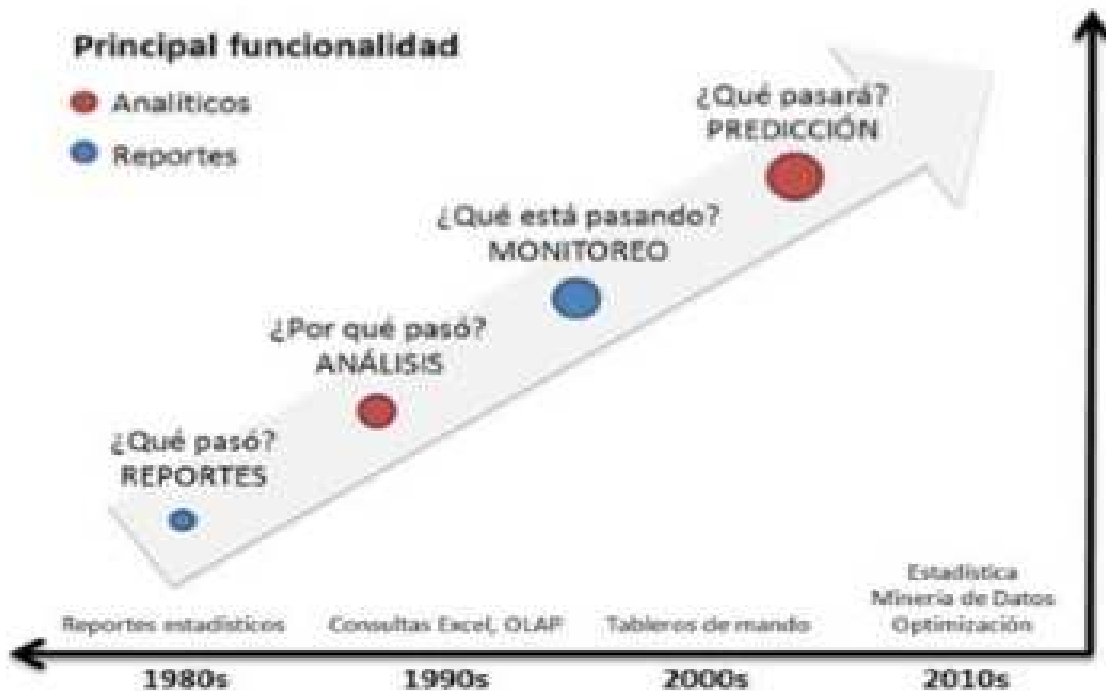


Figura 16. Evolución de la Inteligencia de Negocios.
Fuente: Cano (2007).

La figura anterior muestra, desde el punto de vista de las respuestas que ofrece la evolución de la inteligencia de negocios.

Herramientas tecnológicas a usar.

Para Moreno (2012) "... Inteligencia de Negocios ofrece una infraestructura de herramientas de análisis e informes integrados, con un motor de *workflow* de procesos de negocio. La plataforma será capaz de ejecutar las reglas de negocio necesarias, expresadas en forma de procesos y actividades, así como también de presentar y entregar la información

específica en el momento adecuado. Su modelo de ingresos parece estar orientado a los servicios de soporte, formación y consultoría.

En la Web se presenta una organización por productos: Reportes, Análisis, Cuadros de Mando, Minería de datos y se hace mención específica al *workflow* como una de las capacidades BI claves de la plataforma.

Características

Moreno (2012) enumera las siguientes características

“Es multiplataforma, tanto a nivel de cliente como a nivel del servidor no hay dependencia, utiliza lenguaje interpretado.

–Está construido en torno al servidor de aplicaciones J2EE JBoss y Jboss Portal, Tomcat y es portable en cualquier servidor de aplicaciones como glassfish, habilitando que toda la información sea accesible mediante un browser en la intranet de la empresa.

–Pentaho presenta informes en los formatos habituales (Html, Excel, PDF, entre otros) mediante JfreeReport u otras plataformas como BIRT o JasperReports. Para la generación de PDFs utilizan el conocido Apache FOP. Asimismo incorpora la librería JPivot, la cual permite visualizar tablas OLAP a través de un browser y realizar las operaciones típicas de análisis OLAP.

–Se cuenta con áreas de reporte llamado *Pentaho Reporting* en el cual existen tres productos con diferentes enfoques y dirigidos a diferentes tipos de usuarios:

–*Pentaho Report Designer*. Editor basado en eclipse con capacidad de personalización de informes a las necesidades de negocio destinado a desarrolladores, esta incluye asistente para facilitar las configuración de propiedades.

–*Pentaho Report Design Wizard*. Herramienta de diseño de informe que facilita el trabajo y permite a los usuarios obtener resultados de forma

inmediata. Está destinada a usuarios con pocos conocimientos técnicos.

–Web ad-hoc *reporting*: Es similar a la herramienta anterior pero vía Web. Extiende la capacidad

–Para generar gráficos a los usuarios finales para la creación de informes a partir de plantillas pre configuradas y siguiendo un asistente de creación. Apoyan en JFreeChart, una librería para generar los gráficos más comunes (2D, 3D, barras, líneas series temporales, entre otros), interfaces para acceder a diferentes fuentes de datos, exportación a PNG, JPEG y PDF, soporte para *servlets*, *JSPs*, *applets* y aplicaciones clientes.

–Los Cuadros de Mandos (*Dashboards*) son un desarrollo propio de Pentaho, recogen información de todos los componentes de la plataforma incluyendo aplicaciones externas, *feeds* RSS y páginas Web. Incluyen gestión y filtrado del contenido, seguridad basada en roles y *drill down*. Pueden ser integrados en terceras aplicaciones, en portales o dentro de la plataforma Pentaho. En la figura 17 se puede observar el portal de la plataforma BI Pentaho usando *dashboards*.

–Los servicios Web constituyen una característica esencial de Pentaho. Las acciones, que son las tareas más sencillas que constituyen una solución de Pentaho, pueden publicarse como *WebServices*. Se utiliza como motor de *WebServices* Apache Axis, quedando los servicios descritos en el lenguaje de definición de servicios Web WSDL.

–Proporciona amplia capacidad de preparación de informes de clases y muy flexibles, desde simples informes en un sitio Web de gran formato a los informes para aplicaciones, como la presentación de informes financieros.

–Capacidad de conectarse directamente a las fuentes de datos o para conectar a través de una capa de metadatos centralizados. Se cuenta con una amplia gama de formatos incluyendo Adobe PDF, HTML, Microsoft Excel, RTF o texto plano, informes con gran contenido que incluyen tablas, gráficos,

plantillas, galerías, formato para cabeceras, pies de página, entre otros.

–Es compatible con amplia fuente de datos, compatible con conectividad de base de datos de Java (JDBC) 2.0 o posterior, soporta bases de datos propietarias incluyendo Oracle, IBM, DB2, Microsoft SQL Server, NCR entre otras. Soporta populares bases de datos de código abierto incluyendo MySQL, PostgreSQL, EnterpriseDB, entre otras.

–Se cuenta con el Pentaho *Data Integration* que permite una poderosa extracción, transformación y carga (ETL), ya que muchas organizaciones tienen información disponible en aplicaciones y base de datos separados. Pentaho *Data Integration* abre, limpia e integra esta información y la pone en manos del usuario, provee una sola versión de todos los recursos de información”. (p.46).

Arquitectura

Moreno (2012) con respecto a la arquitectura afirma que la plataforma de *Business Intelligence* Pentaho cuenta con la siguiente arquitectura:

–“Plataforma ciento por ciento J2EE asegurando la escalabilidad, integración y portabilidad.

–Puede correr en servidores compatibles con J2EE como JBoss AS, WebSphere, Tomcat, WebLogic y Oracle AS.

–Base de datos, vía JDBC, IBM DB2, Microsoft SQL Server, MySQL, Oracle, PostgreSQL, NCR Teradata, Firebird, etc.

–No hay dependencia del sistema operativo, utiliza lenguaje interpretado Java.

–Lenguajes de programación Java, Javascript, JSP, XSL (XSLT/XPath/XSL-FO).

–Interfaz de desarrollo basado en Java SWT, Eclipse, Web-based.

–Repositorio de datos basado en XML.

–Todos los componentes están expuestos vía *Web Services* para facilitar la integración con arquitecturas orientadas a servicios (SOA).

El servidor de Pentaho contiene los motores y componentes para la presentación de informes, análisis, reglas de negocio, correo electrónico, notificaciones de escritorio y flujo de trabajo. Estos componentes se integran en conjunto a fin de que puedan ser utilizados para resolver un problema de Inteligencia de Negocio.

Dentro de una solución de la conducta, la interoperabilidad, y la interacción del usuario de cada subsistema se define por una colección de documentos de definición de soluciones y estos son documentos XML que contienen definiciones de las actividades que se ejecutan como parte de los procesos, en la demanda o llamado por los servicios Web. Estas actividades incluyen las definiciones de: fuentes de datos, consultas, plantillas de informes, las normas de notificación, las normas de negocios, tableros de control, análisis y puntos de vista.

El servidor de Pentaho incluye la infraestructura que proporciona la administración avanzada del sistema. Esto incluye la supervisión del sistema SMNP (*Simple Network Management Protocol*), los servicios, informes de uso, soporte de servicios Web, herramientas de validación de configuración y herramientas de diagnóstico, también incluye una herramienta para la Integración de Aplicaciones Empresariales (EAI, por sus siglas en inglés) y ETL". (p.49).

Componentes

Moreno (2012) en relación a los componentes indica que "La plataforma Pentaho *Business Intelligence Open Source* incluye todos aquellos componentes, que se pueden encontrar en las soluciones de Inteligencia de Negocios propietarias más avanzadas, dichos componentes son: reportes,

análisis, cuadros de mandos, flujo de trabajo, minería de dato, integración de datos, sistema de autenticación para accesos a varios sistemas, auditoría de uso y rendimiento, planificador, notificador, seguridad y perfiles.

- **Reporte (*Reporting*)** : Pentaho *Reporting* es una solución basada en el proyecto JFreeReport que permite generar informes ágiles y de gran capacidad. Éste permite la distribución de los resultados del análisis en múltiples formatos, todos los informes incluyen la opción de imprimir o exportar a formato PDF, XLS, HTML y texto. Pentaho *Reporting* permite también la programación de tareas y ejecución automática de informes con una determinada periodicidad. Soporta amplia fuente de datos relacionales, OLAP y XML basados en fuentes de datos.

- **Análisis (*Analysis*)**: Pentaho Analysis suministra a los usuarios un sistema avanzado de análisis de información. Con uso de las tablas dinámicas (*pivot tables, crosstabs*) generadas por Mondrian y JPivot, el usuario puede navegar por los datos ajustando la visión de los datos y los filtros de visualización añadiendo o quitando los campos de agregación.

- Los datos pueden ser representados en una forma de SVG o Flash, cuadros de mando o también integrados con los sistemas de minería de datos y los portales Web (*portlets*). Además, con el Microsoft Excel *Analysis Services* se puede analizar los datos dinámicos en Microsoft Excel (usando la conexión a OLAP server Mondrian).

- **Cuadros de Mandos (*Dashboards*)** ; Todos los componentes del módulo Pentaho *Reporting* y Pentaho *Analysis* pueden formar parte de un cuadro de mando. En Pentaho *Dashboards* es muy fácil incorporar una gran variedad en tipos de gráficos, tablas y velocímetros e integrarlos con los *Portlets* JSP, en donde podrá visualizar informes, gráficos y análisis OLAP.

- **Minería de Datos (*Data Mining*)**: Pentaho incorpora la tecnología Weka, es un conjunto integral de herramientas para aprendizaje automático y

minería de datos. Su amplio conjunto de clasificación, regresión, reglas de asociación y algoritmos de agrupación se puede utilizar para ayudarle a comprender mejor la empresa y también ser aprovechado para mejorar el rendimiento futuro a través de análisis predictivo. Como principales características se tienen:

- Motor de *Data Mining* provee de un set de algoritmos de aprendizaje del proyecto Weka incluyendo *clustering*, segmentación, árboles de decisión, redes neuronales y análisis de componentes principales, entre otros.
- Integración de Pentaho *Data Integration*, automatización de procesos de transformación de datos al formato que requiere el motor de *Data Mining*.
- Los algoritmos pueden ser aplicados directamente a un set de datos o invocados desde código Java
- Provee filtros para normalización, re-muestreo, selección de atributos, transformación y combinación de atributos
- Modelos para predicción de variables nominales o numéricas
- Herramientas gráficas de diseño para pre-procesamiento de datos, regresión, *clustering*, reglas de asociación.

– **Integración de Datos:** Pentaho *Data Integration* (PDI, también llamada Kettle) es el componente de Pentaho responsable de la extracción, transformación y carga (ETL) de procesos. Aunque las herramientas ETL son los más usados en entornos de almacenes de datos, Kettle también se puede utilizar para otros fines:

- Migración de datos entre las aplicaciones o bases de datos.
- Exportación de datos desde bases de datos para archivos planos.

- Carga de datos de forma masiva en bases de datos.
- Limpieza de datos.

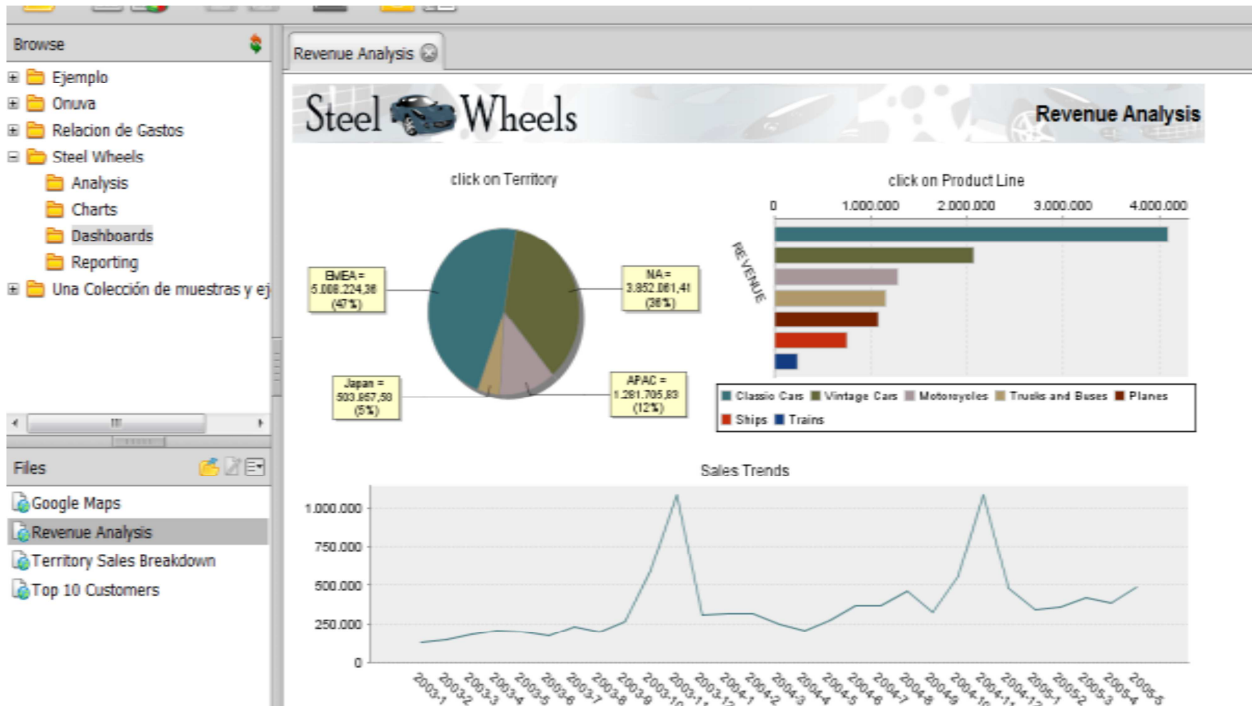


Figura 17. Plataforma *Business Intelligence* utilizando *Dashboard*.
Fuente: Pentaho, 2005.

En la figura anterior se muestra un ejemplo de *Dashboard* creado en Pentaho.

CAPÍTULO 3

MARCO APLICATIVO

En este capítulo se presentarán las fases que llevarán a la implementación de la solución propuesta, con BI, para la obtención de indicadores en la nómina de instructores de la Fundación Musical Simón Bolívar (Fundamusical). Para esto, se utilizará el método de Kimball, aprovechando la velocidad y la orientación ascendente del método.

La metodología ascendente es un enfoque, propuesto por Ralph Kimball en 1998, denominado Ciclo de Vida Dimensional del Negocio (*Business Dimensional Lifecycle*) y se utiliza a la hora de diseñar y crear almacenes de datos. La idea principal de este enfoque es conectar todas las bodegas de datos a un almacén mediante una arquitectura en bus, la cual contiene todos los elementos comunes que pueden utilizar dichas bodegas. De tal forma se tiene, por así decirlo, un almacén virtual que representa la unión de todas las bodegas, las cuales podrían estar localizadas en diferentes lugares, compartiendo los recursos que tienen en común.

Este ciclo de vida del proyecto está basado en cuatro principios básicos:

- **Enfocarse en el proceso del negocio:** Hay que centrarse en el levantamiento de requerimientos y realizar un análisis agudo de las necesidades por parte de la organización.
- **Construir una infraestructura de información adecuada:** Busca el diseño de un repositorio de información integrado, de alto rendimiento y fácil uso que satisfaga las necesidades del cliente.
- **Realizar entregas periódicas:** Crear el data warehouse de forma incremental estableciendo prioridades en los requerimientos de la

organización, este principio se asemeja a las metodologías ágiles de construcción de software.

- **Ofrecer la solución completa:** Proporcionar todos los elementos necesarios para cumplir con las exigencias del cliente, es decir, construir un DW sólido y confiable, brindar herramientas de consultas ad hoc, aplicaciones para informes y análisis avanzado, capacitación, soporte, sitio web y documentación.

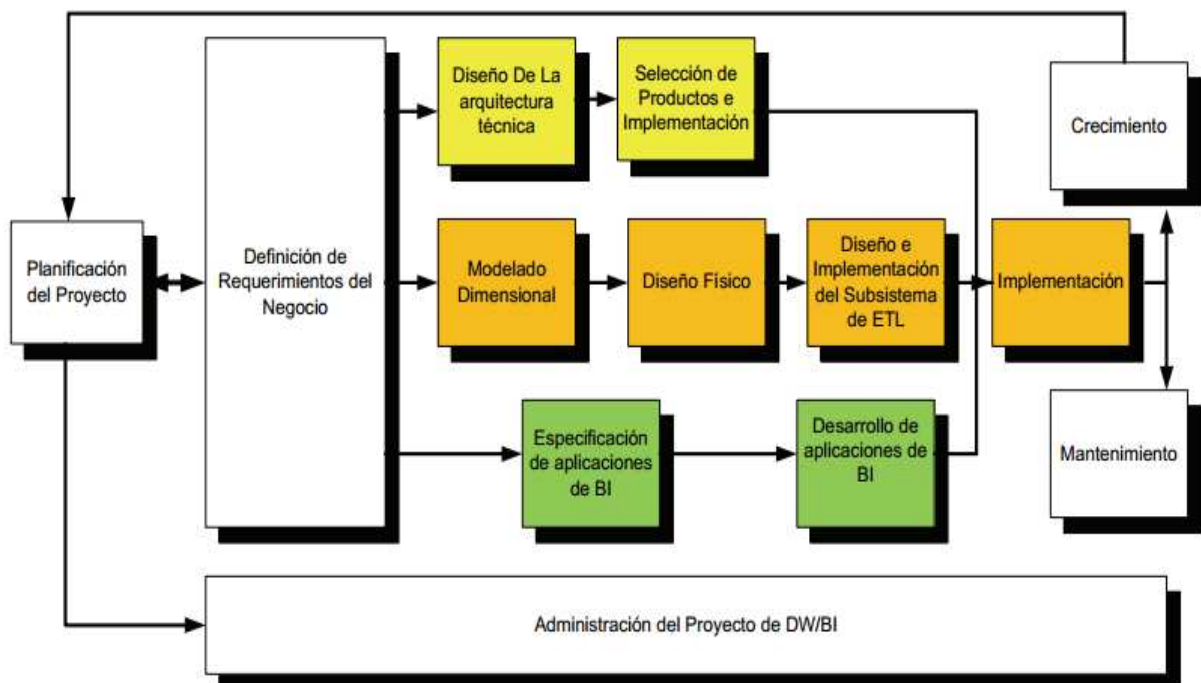


Figura 18. Ciclo de Vida Tareas de la metodología.
Fuente: Kimball y Ross (2002)

Una de las principales ventajas que tiene la implementación de esta metodología en una solución de Inteligencia de Negocios es su simplicidad. Partiendo de un modelo dimensional en estrella, se puede construir en un par de semanas una primera solución que responda a las necesidades básicas del negocio. Luego, dado que las soluciones de Inteligencia de Negocios

suelen ser iterativas, se pueden ir añadiendo más requerimientos a medida de que el proyecto lo vaya necesitando.

Las fases establecidas en dicha metodología utilizada para el desarrollo son:

1. Planificación del proyecto.
2. Definición del alcance
3. Justificación
4. Definición de requerimientos del negocio.
5. Diseño de la arquitectura técnica.
6. Selección de Productos e instalación.
7. Diseño del Modelo dimensional.
8. Diseño físico.
9. Diseño y construcción de procesos ETL.
10. Especificación y desarrollo de aplicaciones analíticas.
11. Implementación.
12. Mantenimiento y crecimiento.

A continuación se describen con detalle los procesos que se realizarán en cada una de las fases que se llevarán a cabo para la implementación de la solución.

Planificación del proyecto.

Kimball y Ross (2002) aseguran que “Esta fase está constituida por la identificación y manejo del problema actual como punto inicial se debe definir el proyecto, elaborar un plan de trabajo y gestionar la puesta en marcha del mismo definiendo y respetando el alcance establecido”. (p.23).

En este caso, se desarrolla una solución de Inteligencia de Negocios útil para la Fundación Musical Simón Bolívar, con la cual les permitirá tener una herramienta que facilite el análisis, a través de reportes, de la información

obtenida en el proceso de nómina de los instructores que laboran en dicha institución.

Tabla 03. Plan de trabajo.

Fase.	Tiempo estimado para la culminación.
Planificación del proyecto.	1 semana.
Definición del alcance	1 semana.
Justificación	1 semana.
Definición de requerimientos del negocio.	2 semanas.
Diseño de la arquitectura técnica.	1 semana.
Selección de Productos e instalación.	1 semana.
Diseño del Modelo dimensional.	1 semana.
Diseño físico.	1 semana.
Diseño y construcción de procesos ETL.	4 semanas.
Especificación y desarrollo de aplicaciones analíticas.	4 semanas.
Implementación.	1 semana.
Mantenimiento y crecimiento	Continuo.

Fuente: Elaboración Propia

Para lograr esto, se planea seguir el plan trabajo, constituido por una serie de fases que en conjunto totalizan el tiempo de ejecución de la solución propuesta. Ver tabla 03.

Definición del alcance.

Kimball y Ross (2002) acota que “Al momento de comenzar con el desarrollo de una solución, a pesar de la disposición y/o capacidad que pueda tener una organización, es necesario establecer límites en el proyecto inicial.

Estos límites deben establecerse buscando llegar a un punto de equilibrio entre los requerimientos de la organización y la disponibilidad y

posibilidades del grupo desarrollador, esto permite que el desarrollo de la solución sea factible en cuanto a objetivos y tiempo”. (p.23).

El presente Trabajo Especial de Grado está basado en el desarrollo de una solución de inteligencia de negocio para la obtención de indicadores en la nómina de instructores de la Fundación Musical Simón Bolívar. Se escogieron diez (10) indicadores de negocio en la nómina de instructores, que permita a la gerencia media y alta, de la institución, obtener reportes que faciliten el análisis de los resultados de los datos obtenidos del proceso de nómina y de igual forma fue mostrado en un portal web (*dashboard*) para su fácil acceso y comprensión.

Para Fundamusical, es importante poder identificar, el proceso de nómina, información relevante de la nómina de instructores, por lo cual se tomaran en cuenta los siguientes puntos de interés:

- Observar el crecimiento de la plantilla de instructores en nómina, comparado con el resto de los trabajadores.
- Identificar rangos de mayor incidencia, con respecto al nivel académico de los instructores, con la finalidad de tomar decisiones estratégicas.
- En relación a la cantidad de instructores, este indicador provee información de la cantidad cargos ocupados por instructores.
- Identificar cuantos instructores no tienen cuenta del Banco de Venezuela para planificar jornadas de apertura de cuentas masivas.
- Identificar oportunidades de incorporación de instructores en las cátedras dictadas reduciendo costos de traslado y maximizando la presencia de instructores en los núcleos.
- Vigilar que la plantilla de profesores no exceda a la estipulada para cada cátedra.
- Identificar la existencia de instructores que impartan más de 160 horas mensuales, para tomar medidas y evitar sanciones contempladas en la ley del trabajo.

- Identificar si existen instructores que dicten más de 160 horas mensuales para tomar medidas y evitar sanciones contempladas en la ley del trabajo conociendo su ubicación para tomar medidas correctivas.
- Vigilar que la plantilla de profesores no exceda a la estipulada para cada estado.
- Observar la existencia de pagos retrasados, con respecto al mes de pago.

Justificación.

Kimball y Ross (2002) indica que “Al justificar la solución es necesario tomar en cuenta los beneficios y costos asociados al desarrollo e implementación”. (p.2.3).

Este proyecto ha sido desarrollado en el contexto del Trabajo Especial de Grado presentando ante la ilustre Universidad Central de Venezuela, Facultad de Ciencias, Escuela de Computación, bajo la tutoría del profesor Franklin Sandoval, para optar al Título de Licenciado en Computación, por parte del bachiller Guillermo Obando Marín como único integrante del equipo desarrollador, con la finalidad de brindarle a la Fundamusical una solución de inteligencia de negocio para la obtención de indicadores en la nómina de instructores, con la cual podrán tomar decisiones, tanto financieras como operativas para mejorar su gestión en el departamento de Nómina.

Definición de Requerimientos de Negocio.

Kimball y Ross (2002) dicen que “A pesar de que se pueda conocer someramente el área del negocio, es necesario realizar entrevistas, encuestas y reuniones con el cliente para así verificar las necesidades que busca solucionar con inteligencia de negocios. De esta manera, se garantiza la construcción y utilización correcta del *Data Warehouse*.”

Los requerimientos obtenidos han de documentarse, dado a que no sólo servirá para certificar que se contempló todo lo que el cliente necesita, sino que también servirá para que el grupo de desarrollo conozca las funcionalidades que ha de tener el sistema, lo que se necesita para cumplir con los requerimientos especificados, los usuarios involucrados en el proyecto, el léxico que se emplea, entre otras cosas". (p.2.3).

Se realizaron entrevistas al personal de medio y alto nivel de gerencia, al igual que a los analistas responsables de suministrar la información a través informes, identificando los siguientes indicadores más relevantes a considerar, los cuales son:

- Cantidad de Empleados por nómina mensual.
 - Objetivo: Observar el crecimiento de la plantilla de instructores en nómina comparado con el resto de los trabajadores.
 - Valor esperado: No hay un valor especificado.
- Cantidad de instructores por nivel mensual.
 - Objetivo: Identificar rangos de mayor incidencia, con respecto al nivel académico de los instructores, con la finalidad de tomar decisiones estratégicas.
 - Valor esperado: No hay un valor especificado.
- Cantidad de Cargos por nivel mensual.
 - Objetivo: En relación a la cantidad de instructores, este indicador provee información de la cantidad cargos ocupados por instructores.
 - Valor esperado: No hay un valor especificado.
- Cantidad de instructores sin cuenta bancaria mensual.
 - Objetivo: Identificar cuántos instructores no tienen cuenta del Banco de Venezuela para planificar jornadas de apertura de cuentas masivas.

- Valor esperado: cero (0).
- Cantidad de instructores que dictan clases en más de un núcleo mensual.
 - Objetivo: Identificar oportunidades de incorporación de instructores en las cátedras dictadas reduciendo costos de traslado y maximizando la presencia de docentes en los núcleos.
 - Valor esperado: cero (0).
- Cantidad de instructores por cátedra mensual.
 - Objetivo: Vigilar que la plantilla de profesores no exceda a la estipulada para cada cátedra.
 - Valor esperado: dependerá de cada cátedra.
- Cantidad de instructores que dictaron más de 160 horas al mes.
 - Objetivo: Identificar si existen instructores que dicten más de 160 horas mensuales para tomar medidas y evitar sanciones contempladas en la ley del trabajo.
 - Valor esperado: cero (0).
- Cantidad de instructores que dictaron más de 160 horas al mes, clasificados por estado.
 - Objetivo: Identificar si existen instructores que dicten más de 160 horas mensuales para tomar medidas y evitar sanciones contempladas en la ley del trabajo, conociendo su ubicación para tomar medidas correctivas.
 - Valor esperado: cero (0).
- Cantidad total de instructores por estado mensual.
 - Objetivo: Vigilar que la plantilla de profesores no exceda a la estipulada para cada estado.
 - Valor esperado: dependerá de cada estado.

- Cantidad de instructores con pagos extraordinarios mensual (fuera del mes de nómina).
 - Objetivo: Observar la existencia de pagos retrasados, con respecto al mes de pago.
 - Valor esperado: cero (0).

Diseño técnico de la arquitectura.

Kimball y Ross (2002) “Ya que, el diseño de la arquitectura servirá como esquema referencial que integrará los elementos tecnológicos, también contribuirá a minimizar imprevistos en el desarrollo e implementación de la solución”. (p.2.3).

A continuación en la siguiente figura se representa la arquitectura de la solución de inteligencia de negocio a propuesta.

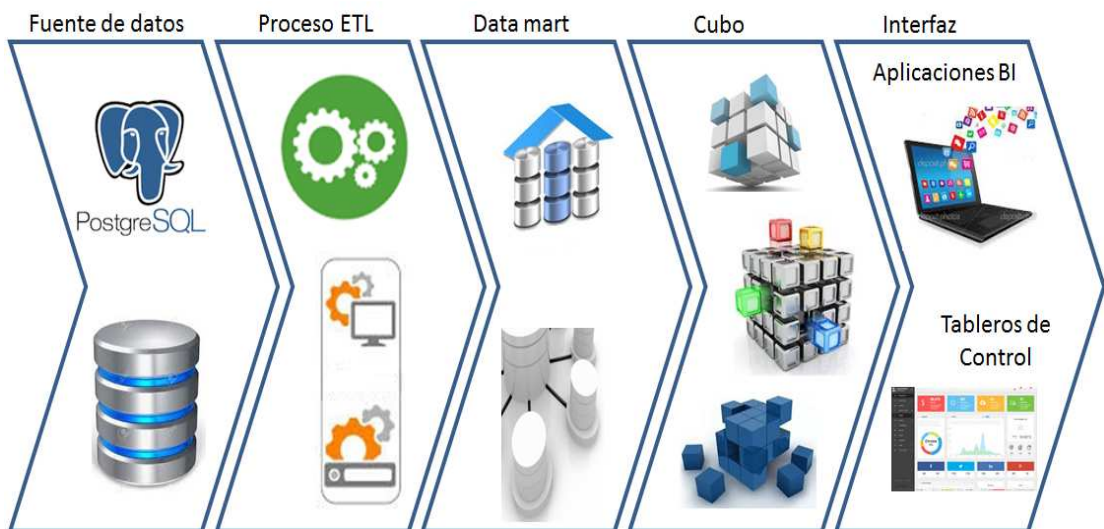


Figura 19 Arquitectura de la solución de Inteligencia de Negocio.
Fuente: Elaboración propia.

Esta arquitectura está conformada por cinco componentes:

El primer componente, denominado “Fuente de datos” corresponde al repositorio de datos que se debe considerar para la elaboración de la solución de inteligencia de negocios. En su totalidad los datos están contenidos en una base de datos Postgres la cual es alimentada por el sistema transaccional

(SIGA), que actualmente se mantienen en la institución y con el cual se lleva a cabo el proceso de nómina de los instructores.

El segundo componente, “Proceso ETL”, mantiene una base de datos temporal, donde se realizará periódicamente la extracción de la fuente de datos. La transformación que estandariza los datos es realizada en este proceso y la carga de los mismos para que ser utilizados en la solución de inteligencia de negocios.

El “*Datamart*” es el componente donde se podrán consultar y extraer los datos relevantes para esta perspectiva del negocio. Este será alimentado periódicamente por el componente anterior, como se mencionó previamente.

“Cubos de información”, es donde se manipularán y detallarán los datos desde las perspectivas deseadas que dependerán de los criterios de consulta.

En cuanto al último, “Interfaz”, es referente a la capa de presentación. Son las herramientas de acceso a los datos para que los usuarios puedan visualizar y analizar acerca de los indicadores y reportes de esta solución propuesta de inteligencia de negocios. La elaboración de un *dashboard* donde se presentaran los indicadores de forma organizada y fácil de visualizar en un solo lugar.

Selección de productos e instalación.

Kimball y Ross (2002) informa que “En esta fase se decide cuáles herramientas serán usadas para el desarrollo de la solución (motor de la BD, herramientas de ETL, plataforma de BI, entre otros). Luego de la instalación se debe probar el correcto funcionamiento de las herramientas en todos los ambientes disponibles”. (p.2.3).

Ya que la Fundación Musical Simón Bolívar es una institución que depende del Estado, debe apearse a regulaciones de uso de software libre, por esta razón, para la solución de inteligencia de negocio se selecciona

Pentaho BI Platform 5.4, la cual provee un conjunto de herramientas y componentes que facilitan las etapas de desarrollo, prueba e implementación.

Estos componentes son:

- *Business Analytics Platform*, es el servidor de Pentaho, donde correrán todas las demás aplicaciones.
- PDI - Spoon es el *Data Integration* de Pantaho. Es la aplicación donde se llevará a cabo el proceso ETL.
- En el *Schema Workbench*, se configuran los cubos de información que se utilizarán el esta solución de inteligencia de negocios.
- *JPivot View* es donde se configuran y muestran los reportes extraídos de consultas OLAP.
- *CDE Dashboard*, donde se configuran los componentes necesarios para presentar los resultados en un ambiente web, para ser accedido por los usuarios finales, de manera fácil y sencilla.

El sistema manejador de bases de datos, Postgres se utilizará por conveniencia, ya que es el actual sistema manejador de bases de datos utilizado por la institución, lo que otorga beneficios de administración, pues ya existe una infraestructura preparada para este manejador.

Diseño del Modelo Dimensional.

Kimball y Ross (2002) dice que “El modelo dimensional es el diseño físico que dará forma a las antiguas fuentes de datos en la estructura final del *Data Warehouse*, a través de una técnica que busca la presentación de los datos en un marco de trabajo estándar que es intuitivo y permite un acceso de alto desempeño”. (p.23).

Para la elaboración del diseño del modelo dimensional se plantean los siguientes pasos propuestos por Kimball que llevaran a identificar los elementos que conformarán los hechos medibles y dimensiones con los

cuales se definirá el modelo dimensional de la solución de inteligencia de negocio.

Definir el proceso de negocio.

Se aplicará la solución de inteligencia de negocio en el proceso de nómina de los instructores de la Fundación Musical Simón Bolívar, en la cual mensualmente se realiza el pago a sus trabajadores, calculando la cantidad de horas dictadas en una cátedra (instrumento) en algún núcleo situado en poblaciones diversas de la geografía nacional.

Definir del nivel de granularidad.

Kimball (1998) asevera que “Luego de conocer el proceso de negocio en el paso anterior, es el momento de definir la granularidad, es decir, el nivel de detalle al que se desea almacenar la información del proceso a modelar”.

Para determinar el nivel de granularidad o nivel de detalle de la información que se va a almacenar, se debe jerarquizar en niveles de información y posteriormente se identifica el grado de granularidad requerida. Se debe responder “¿De qué forma se representa una entrada, en la tabla de hechos? ”. (p.23).



Figura 20. Jerarquía de la información transaccional

Fuente: Elaboración propia

Se puede decir entonces que, para cada fecha de nómina, se puede detallar, en que núcleo, bajo qué cátedra, cuál instructor impartió clases, convirtiéndose esto en el grano de nuestra solución.

Determinar las dimensiones.

Kimball y Ross (2002) indica que “Las tablas de dimensiones son aquellas que le dan contexto a los hechos, por tanto juegan un papel importante en el almacén de datos. El poder del almacén de datos es directamente proporcional a la calidad de la profundidad lograda por los atributos de las dimensiones”. (p.23).

Luego de analizar en qué forma se pueden agrupar los datos, de manera que sea conveniente para la realización de esta solución de inteligencia de negocios, se determinó que las dimensiones a considerar son las siguientes:

- Dimensión Tiempo: En esta dimensión estarán los datos relacionados a la fecha de nómina, representada en diferentes formatos. Ej. año, semestre, trimestre, “día-mes-año”.
- Dimensión Núcleo: La dimensión Núcleo contiene la información de la ubicación del núcleo, tales como estado, ciudad, el nombre del núcleo y el programa al cual pertenece.
- Dimensión Empleado: esta dimensión mantiene los nombres de los empleados, el identificador de los empleados, el tipo de nómina a la cual pertenece y otros datos de interés relevantes de los empleados.
- Dimensión Cátedra: la información referente a la cátedra o instrumento del cual imparte clases el instructor, como nombre de la cátedra o su código de identificación está en la dimensión cátedra.
- Dimensión Nómina: el identificador de la nómina administrativa o de instructores así como su nombre, se mantiene en la dimensión nómina.
- Dimensión Cuenta: esta dimensión sirve para identificar presencia o ausencia de cuenta bancaria vinculada a un empleado.
- Dimensión Concepto: los distintos conceptos de nómina asociados los empleados relevantes, estarán en la dimensión concepto.
- Dimensión Nivel: Los instructores están clasificados por niveles, esta información permanecerá en esta dimensión.

- Dimensión Tipo de Empleado: existen tipos de empleados diferentes, en la dimensión tipo de empleado, se guardan estos tipos.

Identificación de los hechos.

Kimball y Ross (2002) dicen que “Es la tabla principal del almacén de datos, en ella se almacena el desempeño medible del negocio. Una fila en la tabla de hecho corresponde a una medida por lo tanto todas las medidas en la tabla, de hecho deben tener el mismo nivel de granularidad”. (p.23).

Igualmente se obtienen los hechos medibles, para la solución:

- Cantidad de Horas: refiere a la cantidad de horas dictadas por un instructor en un núcleo que imparte clases en una cátedra específica, en una fecha particular.
- Monto de remuneración: Es el monto de la remuneración obtenida por la cantidad de horas dictadas.

Con los elementos Anteriores se obtiene el modelo dimensional, que se ilustra en la figura 21. Una vez definidas las perspectivas a utilizar y los hechos a medir en la solución propuesta, se procede a identificar por cada dimensión los atributos y las jerarquías que se consideraron para poder cumplir con los objetivos planteados, lo que da paso al siguiente punto.



Figura 21. Modelo Dimensional Fuente: Elaboración Propia

El diseño dimensional de esta solución se ve claramente en la figura 21.

Diseño Físico.

Kimball y Ross (2002) aseguran que “En esta fase de la metodología se traslada el modelo dimensional diseñado en la fase anterior a un diseño físico, es decir, se implementa físicamente el almacén de datos que soportará la solución de inteligencia de negocios”. (p.2.3).

El diseño físico, es la estructura que mantiene físicamente el modelo dimensional del almacén de datos. La siguiente figura (Figura 22) ilustra como es el diseño físico de la solución de inteligencia de negocio para la Fundación Musical Simón Bolívar.

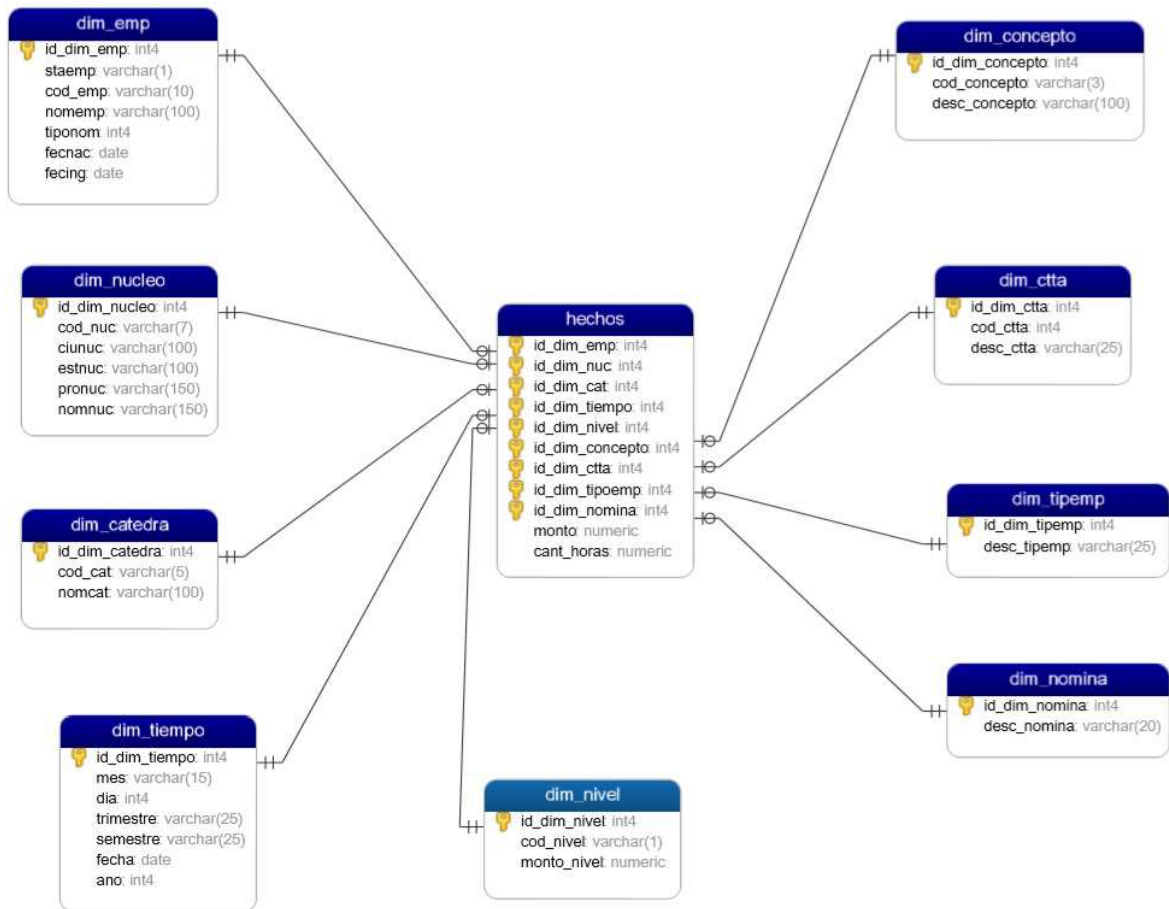


Figura 22. Modelo Físico. Fuente: Elaboración Propia

El modelo físico se desarrolla en postgres, utilizando pgAdmin III, el cual brinda el ambiente necesario para este fin. Para representar cada dimensión se crea una tabla, donde se podrán insertar registros característicos de cada dimensión, así como también para la tabla de hechos, ver Figura 23.

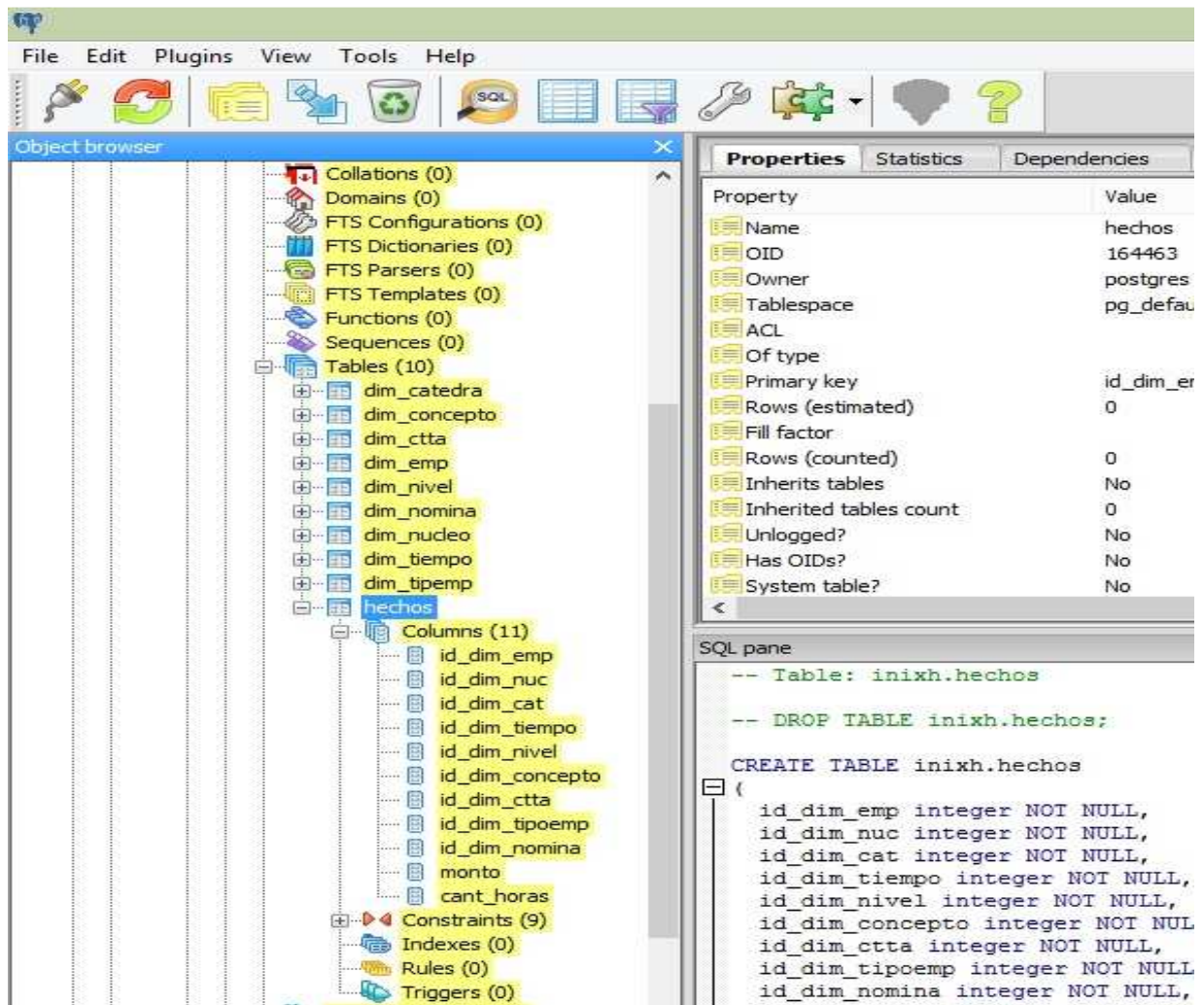


Figura 23. Modelo Físico en PgAdmin III.
Fuente: Elaboración Propia

Diseño y construcción de procesos ETL.

Ya creado el almacén físico en postgres, se deben diseñar e implementar los procesos ETL, que extraerán, transformarán y cargarán la información de la fuente de datos, de la bases de datos del SIGA (sistema integral de gestión administrativa) utilizado por Fundamusical, al almacén de datos creado en el paso anterior, para tener la información relevante que podrá ser utilizada por la solución de inteligencia de negocio.

Para la elaboración de los procesos ETL se utilizará PDI-Spoon (Pentaho *Data Integration*-Spoon) el cual es la herramienta que provee Pentaho. En Spoon se configuran, las conexiones a las bases de datos postgres, tanto la fuente de datos, donde se realizaran las consultas para obtener los datos necesarios, como las tablas que conforman el almacén de datos, donde se albergará la información que se utilizará para la solución de inteligencia de negocio.

Posteriormente se crean las “transformaciones” donde se realiza el procesamiento de los datos de la fuente de datos, para insertarlos luego de transformarlos en las tablas del almacén de datos. Y por último elabora un “Job” que ejecuta secuencialmente todas las transformaciones de las diferentes tablas del almacén de datos.

Cada transformación, como el Job se almacenan físicamente en un repositorio, el cual es una carpeta en particular, para poder ejecutarlos de manera ordenada tantas veces como sea necesaria, una carga inicial y posteriormente las cargas sucesivas que irá actualizando a través del tiempo la información que se utilizará en nuestra solución, ver Figura 24.

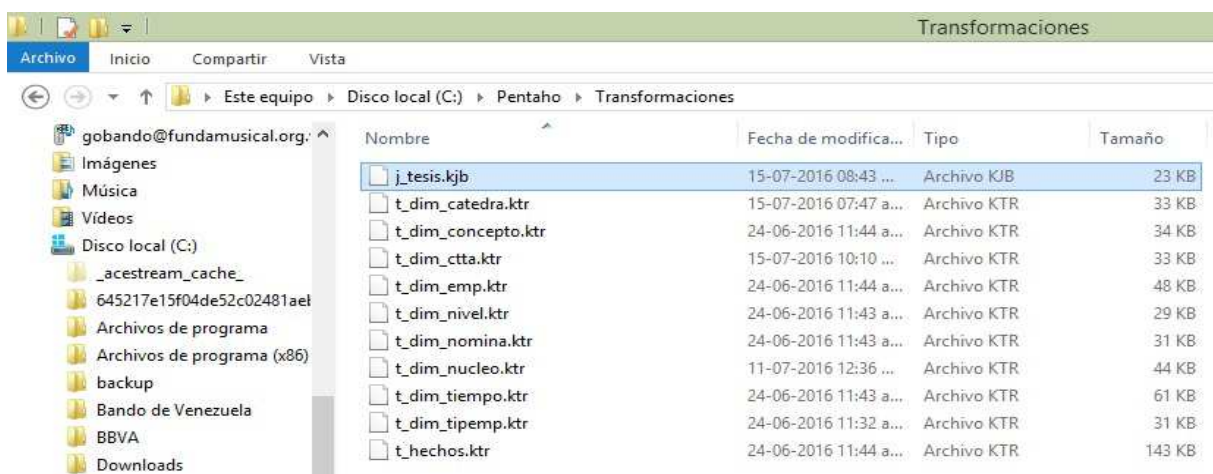


Figura 24. Archivos de las transformaciones del ETL.
Fuente: Elaboración Propia.

La siguiente figura, ilustra la configuración de una transformación. Esta refleja de forma similar a todas, cómo utilizando las funciones que provee Spoon, se manipulan los datos.

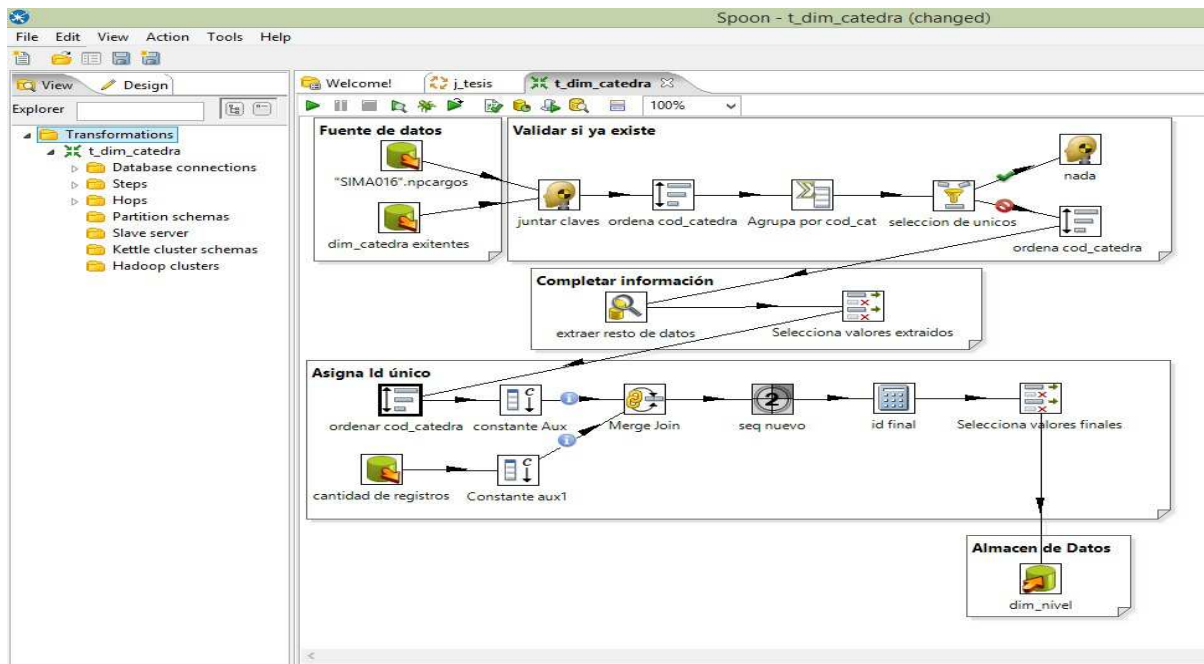


Figura 25. Transformación en Spoon.
Fuente: Elaboración Propia.

En la figura anterior, se muestra como en la fuente de datos se extraen los datos relevantes, luego se validan si existen en el almacén, posteriormente se extrae la información, que al transformarla, completará los datos que se cargaran a la tabla del almacén de datos, luego se le asigna el id único para esa dimensión y por último se realiza la carga en la tabla correspondiente.

Luego de tener las transformaciones, se elabora el Job, el cual ejecutará, de manera secuencial, cada transformación y se observan flechas verdes que representan el flujo de ejecución. Comienza con “STAR” seguidamente de las transformaciones de las distintas tablas de dimensiones, siendo la última en ejecutarse, la transformación que llenará la tabla de

hechos. Su ejecución se guarda para el final, debido a que las tablas que contienen la información de las dimensiones deben estar llenas y actualizadas pues esta información es la utilizada por la tabla de hechos.

Finalmente el paso con el cual concluye la ejecución es “*SUCCESS*”, indicando la correcta finalización de la ejecución. La siguiente figura refleja la secuencia de transformaciones en el Job.

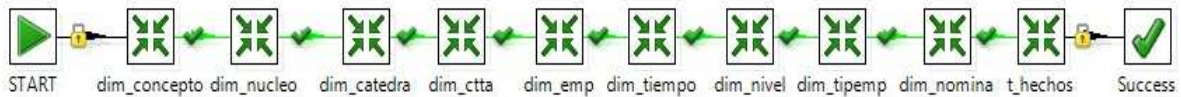


Figura 26. Secuencia de transformaciones (Job).
Fuente: Elaboración Propia.

Job / Job Entry	Comment	Result	Reason	Filename	Nr	Log date
START	Start of job execution		start		0	2016/07/11 12:38:45
START	Job execution finished	Success				2016/07/11 12:38:45
dim_concepto	Start of job execution		Followed unconditional link	file:///C:/Users/Guillermo/Desktop...	1	2016/07/11 12:38:53
dim_concepto	Job execution finished	Success				2016/07/11 12:38:53
dim_nucleo	Start of job execution		Followed link after success	file:///C:/Users/Guillermo/Desktop...	2	2016/07/11 12:38:55
dim_nucleo	Job execution finished	Success				2016/07/11 12:38:55
dim_catedra	Start of job execution		Followed link after success	file:///C:/Users/Guillermo/Desktop...	3	2016/07/11 12:38:56
dim_catedra	Job execution finished	Success				2016/07/11 12:38:56
dim_ctta	Start of job execution		Followed link after success	file:///C:/Users/Guillermo/Desktop...	4	2016/07/11 12:38:57
dim_ctta	Job execution finished	Success				2016/07/11 12:38:57
dim_emp	Start of job execution		Followed link after success	file:///C:/Users/Guillermo/Desktop...	5	2016/07/11 12:39:10
dim_emp	Job execution finished	Success				2016/07/11 12:39:10
dim_tiempo	Start of job execution		Followed link after success	file:///C:/Users/Guillermo/Desktop...	6	2016/07/11 12:39:25
dim_tiempo	Job execution finished	Success				2016/07/11 12:39:25
dim_nivel	Start of job execution		Followed link after success	file:///C:/Users/Guillermo/Desktop...	7	2016/07/11 12:39:26
dim_nivel	Job execution finished	Success				2016/07/11 12:39:26
dim_tipemp	Start of job execution		Followed link after success	file:///C:/Users/Guillermo/Desktop...	8	2016/07/11 12:39:27
dim_tipemp	Job execution finished	Success				2016/07/11 12:39:27
dim_nomina	Start of job execution		Followed link after success	file:///C:/Users/Guillermo/Desktop...	9	2016/07/11 12:39:29
dim_nomina	Job execution finished	Success				2016/07/11 12:39:29
t_hechos	Start of job execution		Followed link after success	file:///C:/Users/Guillermo/Desktop...	10	2016/07/11 12:47:10
t_hechos	Job execution finished	Success				2016/07/11 12:47:10
Success	Start of job execution		Followed unconditional link			2016/07/11 12:47:10
Success	Job execution finished	Success				2016/07/11 12:47:10

Figura 27. Ejecución del Job.
Fuente: Elaboración Propia.

Especificación y desarrollo de aplicaciones analíticas.

Cubo

Con *Schema Workbench*, se configura el cubo de información, cada dimensión está compuesta por jerarquías, niveles de jerarquías y tablas. En la figura siguiente (Figura 28) se puede observar las dimensiones, con su jerarquía, el nivel y tabla respectiva. Al final se ven las medidas, por ejemplo Cantidad de Empleados, Cantidad de Núcleos.

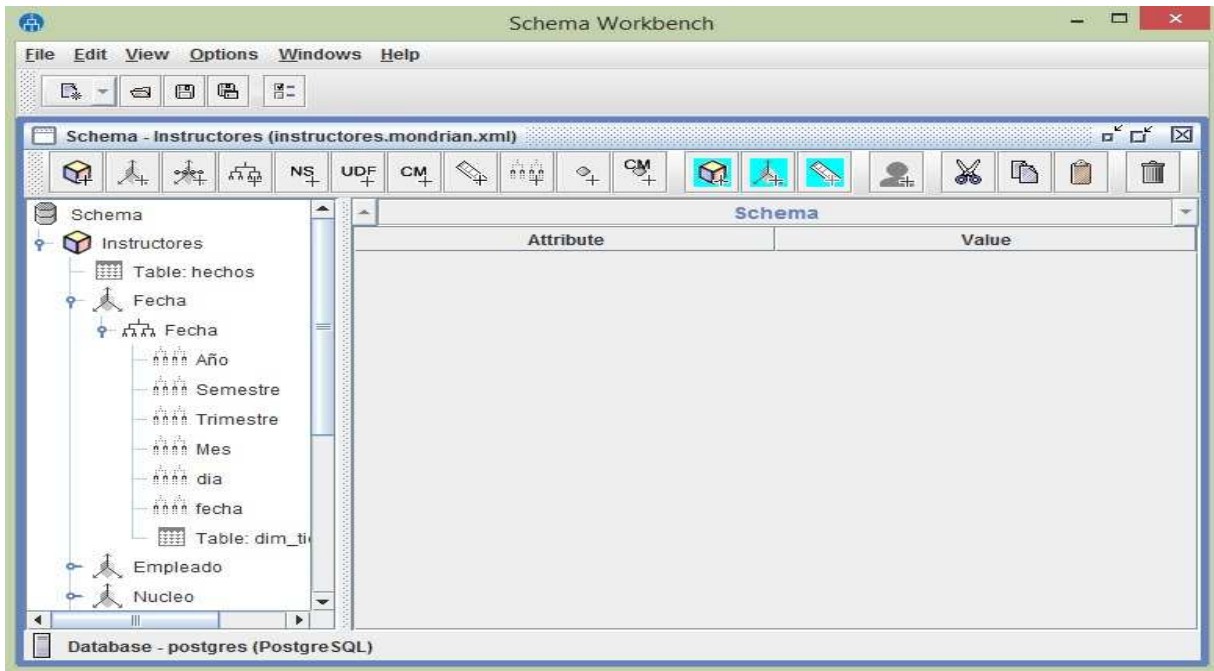


Figura 28. Configuración del cubo de información.
Fuente: Elaboración Propia.

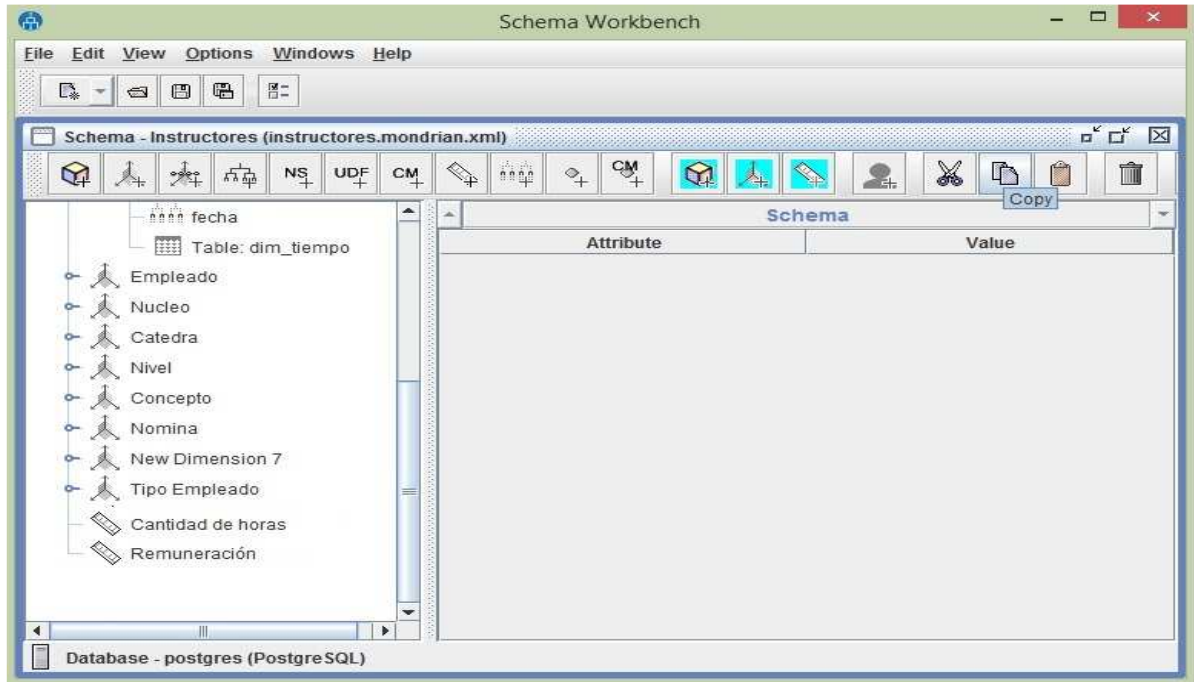


Figura 29. Configuración del cubo de información, medidas.
Fuente: Elaboración Propia.

Reportes

Los reportes son la forma de visualizar los resultados. Se realizan las consultas OLAP sobre los datos, tomando diferentes perspectivas de los mismos, para estas consultas dimensionales se utiliza MDX (*Multidimensional Expressions* o en español, expresiones multidimensionales). Las consultas MDX de los reportes creados se ejemplifican a continuación en la siguiente figura.

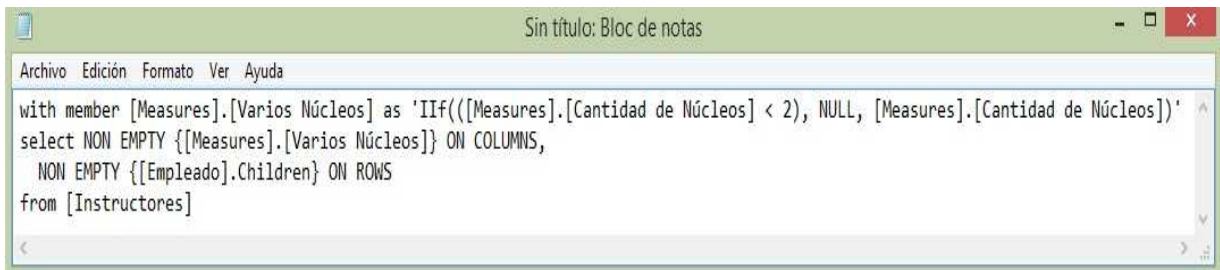


Figura 30. Ejemplo de consulta MDX. Fuente: Elaboración Propia.

Para realizar estas consultas se utilizó Pentaho *Business Analytics*, en el cual se deben configurar, la sección de “*Manage Data Sources*” la fuente de datos e importar el cubo con “*Import analysis*”.

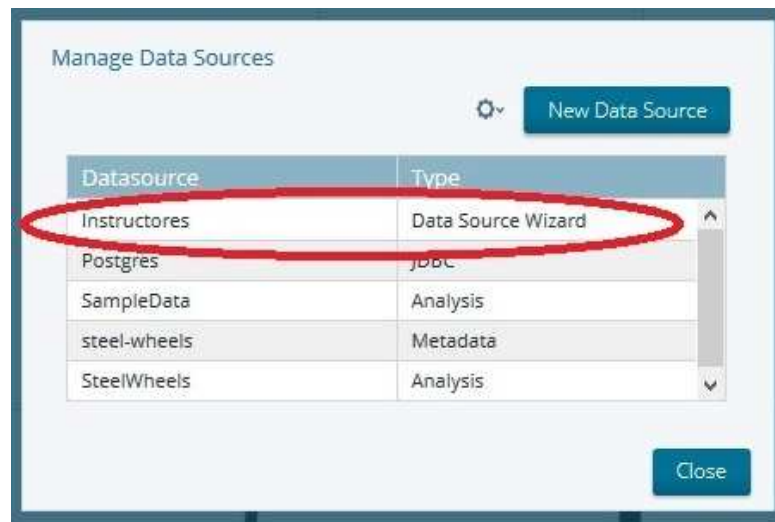


Figura 31 Configuración de *Data Sources* Fuente: Elaboración Propia.

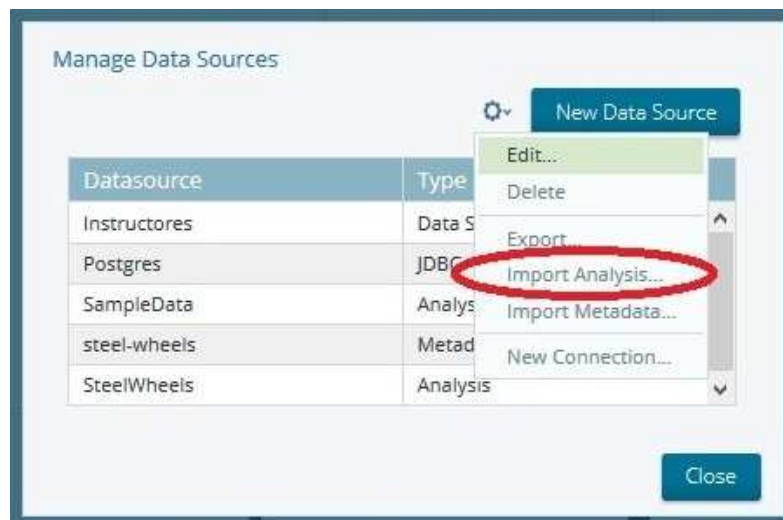


Figura 32 Configuración de *Analysis* Fuente: Elaboración Propia.

Posterior a esto se deben configurar los reportes utilizando las consultas, creando nuevas vistas Jpivot en la sección de “*créate new*” – “*Jpivot view*”.



Figura 33. “Créate new” – “Jpivot view”
Fuente: Elaboración Propia.

Los componentes que tiene el panel Jpivot View para configurar los reportes son variados, los más utilizados son el botón “MDX Query Editor” donde se puede incluir o modificar la consulta MDX del reporte deseado y el botón del “navegador OLAP”, con el cual se puede configurar la consulta, simplemente seleccionando los elemento deseados. En la figura 34, se ilustra lo anterior.

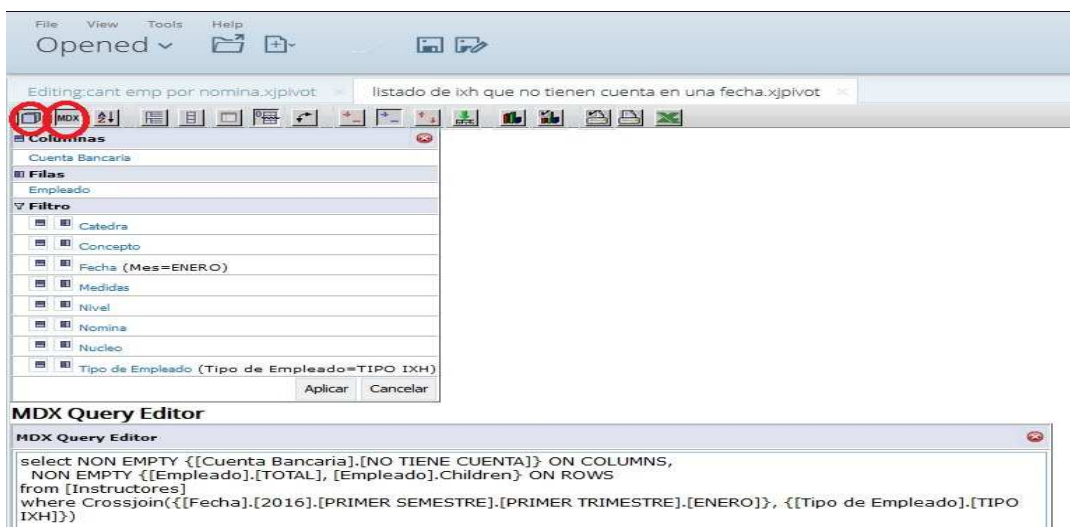


Figura 34. Configuración de JPivot View (Reportes).

Fuente: Elaboración Propia.

Catedra	Medidas			
	Cantidad de Empleados			
	Fecha			
	TOTAL	2014	2015	2016
TOTAL	8013	4871	7371	5802
CATEDRA DE KINDER MUSICAL	126	85	121	101
INSTRUCTOR CATEDRA DE ARPA CLASICA	8	5	7	6
INSTRUCTOR CATEDRA DE ARPA LLANERA	57	38	55	46
INSTRUCTOR CATEDRA DE BANDOLA	32	13	29	20

Figura 35. Ejemplo de Reporte
Fuente: Elaboración Propia.

Luego de configurar los reportes, se da paso a la creación del **Dashboard** que se muestra en un portal web, haciendo uso de tablas, gráficos y otros componentes que permiten manipular los datos de los indicadores solicitados por la Fundación Musical Simón Bolívar. Con *Pentaho Business Analytics*, se configuran los componentes necesarios para visualizar estos indicadores accediendo a “create new” – “CDE Dashboard”. En este módulo, se deben configurar tres secciones.

La primera en configurar es la sección “Layout Panel”, donde se organiza la forma de cómo se presentarán los componentes del portal web. En la segunda sección “*Datasources Panel*”, se configuran las consultas a las fuentes de datos, de dónde se extraen los datos a mostrar. Por último en la sección de “*Components Panel*” se configuran los componentes que muestra el resultado de las consultas. Estos componentes pueden ser, tablas y gráficos, por ejemplo. En la siguiente imagen se observa la configuración del “*Layout panel*” de la solución propuesta.

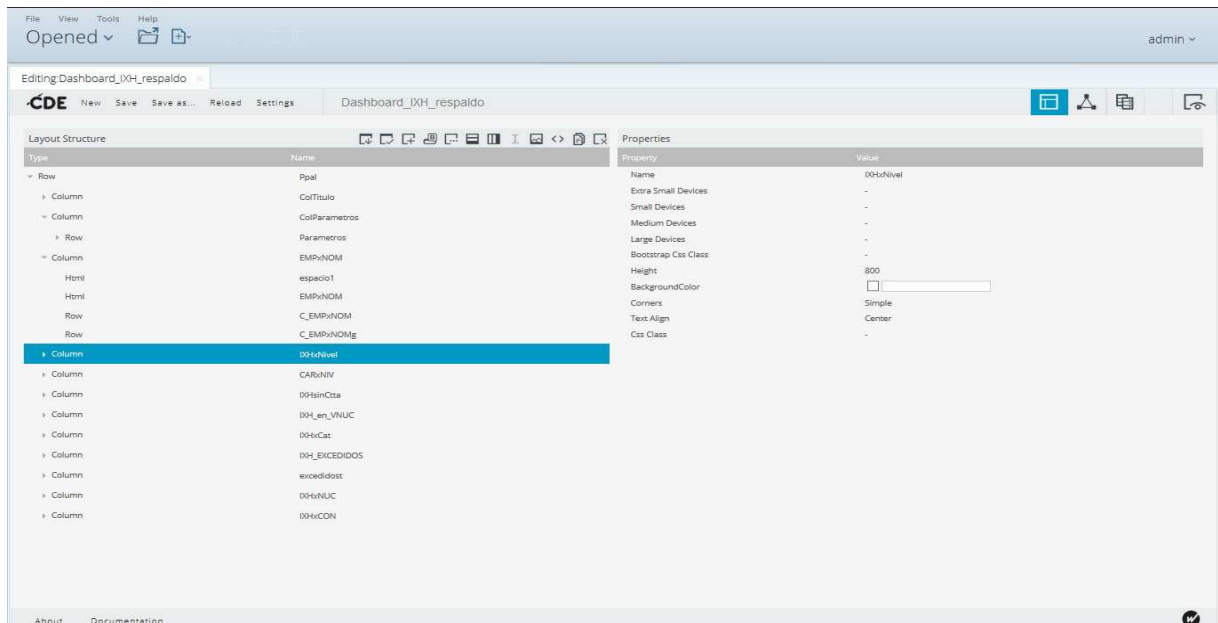


Figura 36. Layout panel Fuente: Elaboración Propia.

Seguidamente se ilustra la configuración del “*Datasources Panel*”, con las distintas consultas MDX y SQL necesarias para extraer del cubo los datos requeridos.

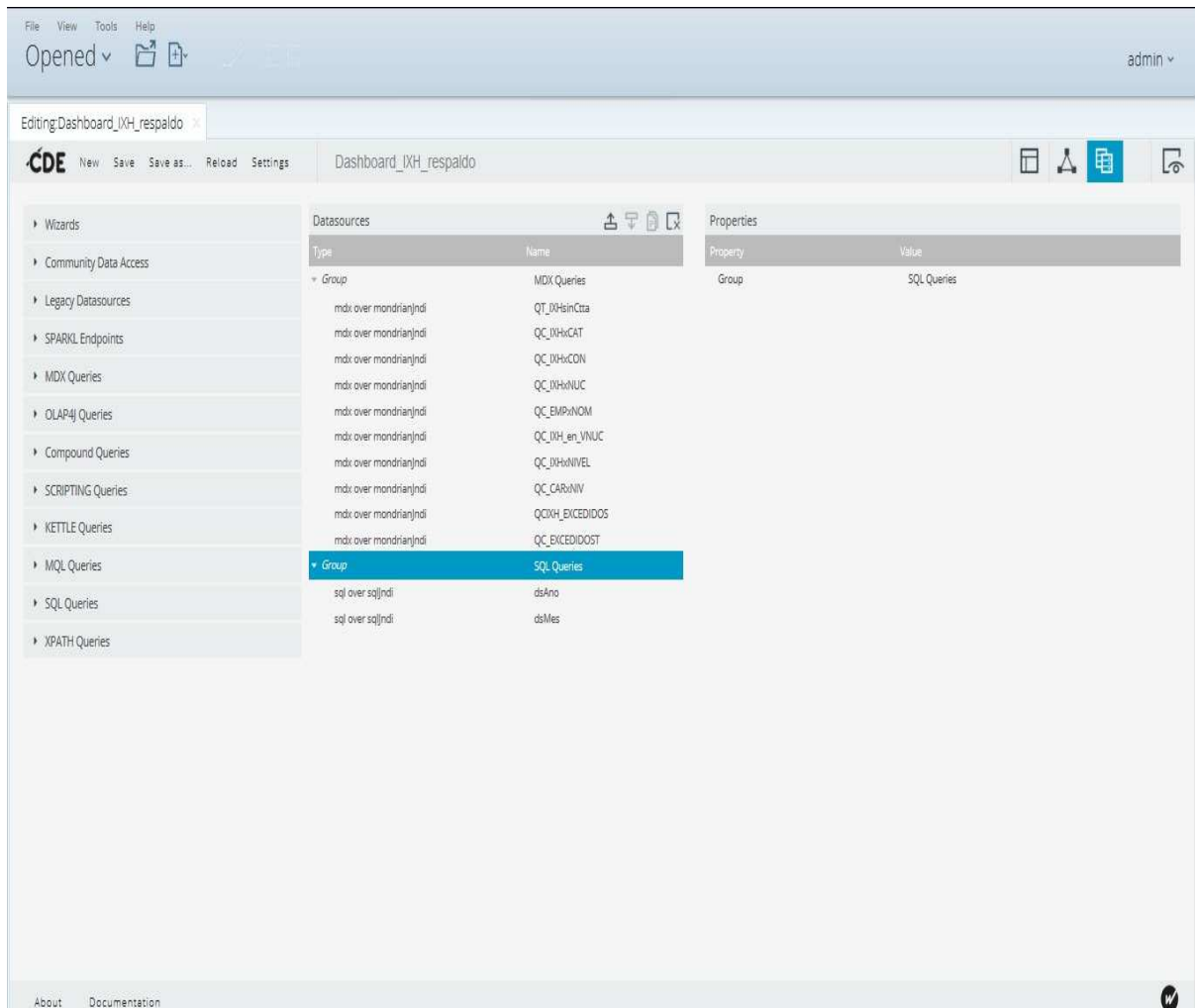


Figura 37. Datasources panel Fuente: Elaboración Propia.

Y por último se muestra en la siguiente imagen la configuración de los componentes en “*Components Panel*”, los cuales son tablas, gráficos y otros, necesarios para mostrar la información de forma comprensible y manipulable, por el usuario de esta solución de inteligencia de negocios.

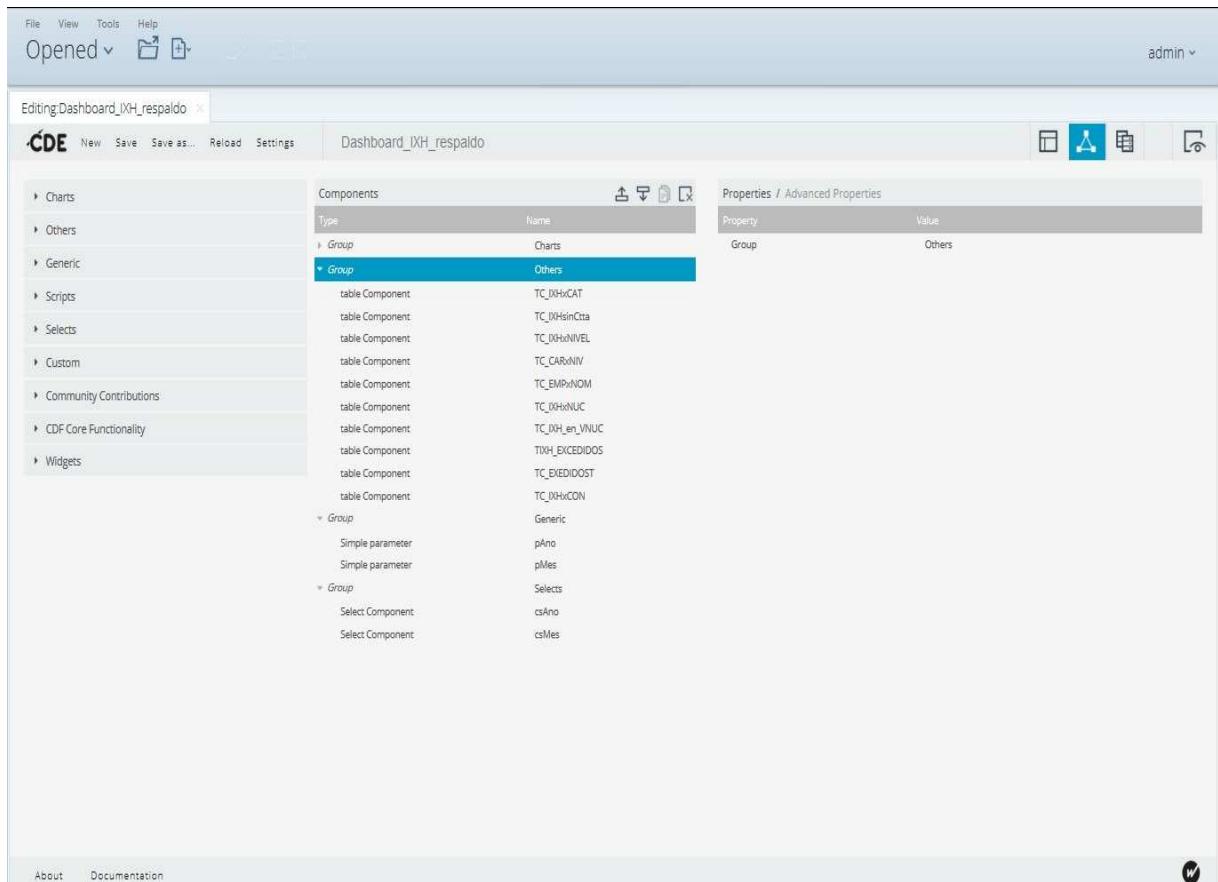


Figura 38. Component panel. Fuente: Elaboración Propia.

Al configurar todas las secciones correctamente se obtiene el *Dashboard* que, tal y como se puede apreciar en la Figura 39, se puede ejecutar desde cualquier navegador web.

Indicadores de la nómina de instructores de Fundamusical Simón Bolívar

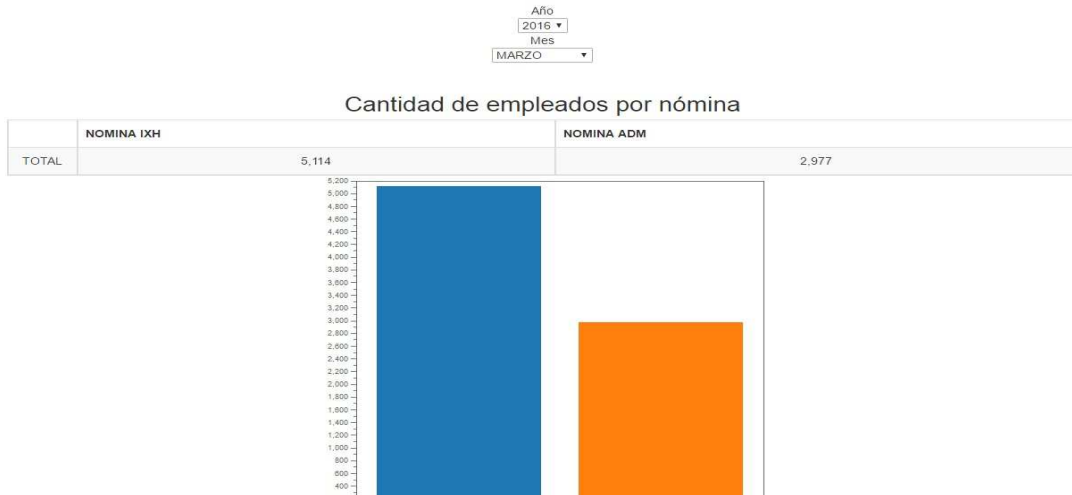


Figura 39. Dashboard. Fuente: Elaboración Propia.

En la figura 39 se observa el *Dashboard* ya terminado.

Implementación.

Para la implementación de esta solución de inteligencia de negocios, se deben instalar en un servidor, destinado por el personal de la Fundamusical Simón Bolívar, las aplicaciones utilizadas anteriormente, migrar los archivos correspondientes a cada paso descrito anteriormente, ya que la suma de todos, conforman la solución de inteligencia de negocios completa.

Para acceder a este portal web se debe autenticar con un usuario y contraseña, los cuales se crean en la sección “*Administration*” – “*Users & Roles*”. Luego de esto, se debe informar a las personas que utilizarán la solución, sus respectivos usuarios y contraseñas, así como el URL para acceder al *dashboard* para poder utilizarlo dentro de la red de la institución.

Mantenimiento y crecimiento.

Una vez completado el desarrollo de la solución, se realizan pruebas para certificar que la información resultante a través de los reportes e indicadores que se muestran en el *dashboard*, es consistente con los valores que se encuentran en el sistema transaccional de Fundamusical Simón

Bolívar. Luego de las pruebas, se mantiene el monitoreo de la solución con el fin de corregir los posibles errores que se puedan presentar una vez se encuentre la solución en ambiente de producción y/o para mejorar el desempeño o visualización de los indicadores.

Mensualmente deberá ejecutarse el Job de transformaciones en Spoon, para mantener la información actualizada y por tanto los indicadores que fueron revisados por la gerencia. Se analizan nuevos requerimientos en la solución de inteligencia de negocio desarrollada con el fin de solicitar el crecimiento de la misma, lo que implicaría otro proyecto de desarrollo.

CAPÍTULO 4

CONCLUSIONES Y RECOMENDACIONES

Dentro del desarrollo de este proyecto de grado, se hizo énfasis en la investigación de todo lo relacionado con la inteligencia de negocios, a modo de conceptualizar un estado del arte para la familiarización y toma de conocimientos en las áreas a abarcar, con la finalidad de desarrollar una solución de inteligencia de negocio para la obtención de indicadores en la nómina de instructores de la Fundación Musical Simón Bolívar, se lograron cumplir distintas actividades que llevarían a cabo este fin. Como primera actividad se identificaron los requerimientos de Fundamusical y se acordaron cuales indicadores serían tomados en cuenta para esta solución.

Posteriormente se elaboró un modelo de datos para mantener la información relevante a la solución propuesta. Seguidamente se diseñó el modelo dimensional para cubrir las necesidades de Fundamusical. Este modelo dimensional sería la base para la construcción del almacén de datos correspondiente. Los datos relevantes fueron extraídos del sistema transaccional que mantiene la Fundación, se transformaron de manera conveniente y fueron cargados en las tablas intermedias, para ser utilizadas en el almacén de datos.

Ya con las dimensiones y tabla de hechos llena con la información de la nómina de instructores, se configuró el cubo de información del cual se extrae la información para calcular los indicadores, que mediante consultas MDX puede presentarse en reportes y finalmente en un portal web (dashboard) que se configuró con el fin de mostrar la totalidad de los indicadores involucrados de una manera ordenada, cómoda y sencilla.

Sin duda, dentro de todo este proyecto, se ha mostrado y comprendido la importancia que posee la implementación de las herramientas de inteligencia de negocios en las empresas, sus dificultades, costos y beneficios, ya que así pueden lograr una ventaja competitiva frente al resto, poder tener un conocimiento real de su empresa, poder analizar información a tiempo y poder tomar las decisiones adecuadas, ayudado por la suite de BI seleccionada.

Los resultados fueron validados por el personal de la institución, para garantizar el correcto procesamiento de los datos y corroborar la validez de los instrumentos generados en este trabajo especial de grado, los cuales representan la solución completa de inteligencia de negocios para la obtención de indicadores en la nómina de instructores de la Fundación Musical Simón Bolívar.

Todo lo anteriormente descrito, constituye una nueva fortaleza para la Fundación Musical Simón Bolívar, ya que la información está condensada en un solo sitio, es fácil de obtener, está ordenada, y de sencilla comprensión, sirviendo así, como herramienta útil al momento de tomar decisiones.

Para finalizar, cabe decir que el uso de la Inteligencia de Negocios está tomando cada vez más fuerza dentro de las entidades y/o empresas, ya sea para cambiar sus estrategias, explotar su información o generar conocimiento importante. Todo ello se ve simplificado para el usuario, gracias al uso de las herramientas como las utilizadas en este proyecto de grado.

Se recomienda continuar con el estudio de nuevos indicadores para enriquecer esta solución, lo que contribuiría al soporte de toma de decisiones.

Para garantizar el crecimiento y hacer uso de esta nueva fortaleza en el resto de la institución, sería recomendable considerar la posibilidad de extender esta solución a otras áreas de Fundamusical, con lo cual poder tener una visión más amplia, al momento de tomar decisiones estratégicas.

REFERENCIAS BIBLIOGRAFICAS

- Alfaro, P. y Lenon, C. (2013). **Sistema de gestión de expedientes clínicos en el área de archivo del hospital escuela Alejandro Davila Bolaños de la ciudad de Managua**. Tesis para optar al título de Ingeniero en Sistemas y Tecnologías de la Información, presentada ante la Universidad Centroamericana, Managua, Nicaragua.
- Bernal, M. (2004) **Contabilidad, sistema y gerencia**. Editorial CEC, SA
- Beynon-Davies, P. (2013) **Introducción a la Informática en las Organizaciones** Barcelona. Reverte.
- Cano, J. (2007) Business Intelligence: **Competir con Información**. Banesto, Fundación Cultur
- Castrillo, A. y Dugarte, D. (2011). **Desarrollo de una solución de inteligencia de negocio para el apoyo a la toma de decisiones en el área ventas de empresas del sector salud**. Tesis para optar al título de licenciado en Computación presentada ante la Universidad Central de Venezuela, Caracas, Venezuela.
- Chiavenato I. (2006) **Introducción a la teoría general de la administración**, McGraw-Hill Interamericana.
- Christiansen, R. y Gaete E., (2014). **Sistemas de tipo OLTP**, Material de apoyo del Curso Lean Thinking entrega valiosas herramientas de negocios en INACAP Sede Arica, Chile.
- Cibertec (2015). **Inteligencia de negocios, Manual de Teoría**. Escuela de Tecnologías de la Información de CIBERTEC, Escuela Comunitaria en Miraflores, Perú
- Date, C. (2001). **Introducción a los sistemas de bases de datos**. México. Pearson Educación. 7ma. Edición.
- Davenport, T. H. & Prusak, L. (2000). **Working knowledge: How organizations manage what they know**. Boston, MA: Harvard Business School Press.
- Deppen, N., Ricca, P. y Trías, D. (2009). **Visión global del proceso de construcción de datawarehouse**. Proyecto de grado presentado ante la Universidad de la República, Montevideo.
- Elmasri, R. y Navathe S. (2002). **Fundamentos de los Sistemas de Bases de Datos**. Addison-Wesley. (3ª edición).
- Espinosa, E., (2010) **Herramientas ETL. ¿Que son, para que valen?. Productos más conocidos**. ETL's Open Source | respinosamilla blog.
Recuperado de:

<http://www.dataprix.com/blogs/respinosamilla/herramientas-etl-que-son-para-que-valen-productos-mas-conocidos-etl-s-open-sour>. Fecha de consulta mayo 2016.

- Ferrell, Hirt, Adriaenséns, Flores, Ramos (2004). **Introducción a los negocios en un mundo cambiante**. McGraw Hill, 4ta edición.
- García G. y Murillo C. (2008). **Estudio de herramientas business intelligence para la implementación de un sistema de información general en la unidad de planificación de la ESPOCH**. Tesis para optar al título de Ingeniero en Sistemas Informáticos, presentada ante la Escuela superior politécnica de Chimborazo. Riobamba. Ecuador.
- Gil, A. y Berriel, R. (2006). **Aplicación de las tecnologías datawarehouse en el contexto de la empresa turística de alojamiento hotelero**. Trabajo especial presentado ante la Universidad de Las Palmas de Gran Canaria, España
- González, F., Gil, H., León, D., Suárez I., (2011) **Descripción general: sistemas de información para ejecutivos (EIS)**. 2º Congreso SOCOTE. Soporte del Conocimiento con la Tecnología. Valencia España.
- Greiner, L. (2015). **Bases de datos**. Manual de Teoría. Universidad de Belgrado.
- Hernández, J. (2014). **Sistema de inteligencia de negocio para la obtención de indicadores en el área de ventas**. Tesis para optar al título de licenciada en Computación, presentada ante la Universidad Central de Venezuela, Caracas, Venezuela.
- Hernández. R, (2005). **Herramientas del Marketing: Data Warehousing, Tecnología necesaria para el Comercio Internacional**. Revista CENIC Ciencias Biológicas, Vol. 36, No. Especial, 2005.
- Inmon, W. (1990). **Building the Data Warehouse**. QED Technical Publishing Group
- Keenan, P., (1997). **Geographic Information Systems their contribution to the IS mainstream. Presented at Association for Information Systems**. 1997 Americas Conference. Indianapolis, USA
- Kendall K. y Kendall J. (2005). **Análisis y diseño de sistemas**, México: Pearson Educación.
- Kimball, R., & Ross, M. (2002). **The Data Warehouse Toolkit (2da ed.)**. Wiley Computer Publishing.
- Laudon K y Laudon J (2012). **Sistemas de información gerencial (22º edición)**, México. Pearson
- Loshin, D. (2003). **Business Intelligence: The Savvy Manager's Guide**. San Francisco: Morgan Kaufmann Publishers

- Medina, A., Campos, C. y Guerrero, R. (2009). **Business Inteligence, El soporte de decisiones en la empresa “Casa Marzam S.A. de C.V.** Tesis para obtener el título de Licenciado en Administración Industrial ante el Instituto Politécnico Nacional. Mexico. D.F.
- Migel, A. y Piattini, M. (1999). **Fundamentos y modelos de Bases de Datos.** RA-MA, 2da edición.
- Mirabella, P. (2011). **Desarrollo de una solución de Inteligencia de Negocios para la obtención de indicadores e informes legales asociados a la Contabilidad.** Tesis para optar al título de licenciado en Computación presentada ante la Universidad Central de Venezuela, Caracas, Venezuela.
- Montelogo (2016).**¿Qué son los EIS?**. Artículo recuperado de <http://www.timogo.com.mx/articulos/EIS.pdf>. Fecha de consulta mayo 2016.
- Moreno, A. (2000). **Diseño e implementación de un lexicón computacional para lexicografía y traducción automática.** Universidad de Málaga. Recuperado de <http://elies.rediris.es/elies9/>. Fecha de consulta mayo 2016.
- Moreno, F. (2012). **Implementación de Indicadores de Gestión para monitorear el rendimiento del proceso de Revisión Editorial, utilizando una plataforma de Inteligencia de Negocios, desarrollada en Software Libre.** Tesis para optar al título de licenciado en Computación presentada ante la Universidad Central de Venezuela, Caracas, Venezuela.
- Para Rivas, J. (2012). **Reingeniería de los procesos de extracción, transformación y carga de la plataforma prepago de Digitel.** Tesis para optar al título de ingeniero en Computación presentada ante la Universidad Simón Bolívar. Caracas, Venezuela.
- Pensandoenbi (2015). **Una historia de Implementación. Tercera Parte ~ Pensando en BI.** Recuperado de <http://pensandoenbi.blogspot.com/2013/05/una-historia-de-implementacion-tercera.html>. Fecha de consulta junio 2016
- Peña, A. (2006) **Tecnologías de la Información: Su alineamiento al Negocio de las Organizaciones,** Tesis para optar al título de doctorado en ciencias de la computación ante el Instituto Politécnico Nacional. México.
- Peralta, M. (2006). **Sistemas de Información** McGraw Hill, Madrid, España
- Pérez, M., Contreras, Y. y Rivero, S., (2015). **Características de los sistemas de información que permiten la gestión oportuna de la información y el conocimiento institucional 2000,** Editorial Ciencias Médicas.
- Rasmussen, N., Goldy, P., & Solli, P. (2002). **Financial Business Intelligence.** John Wiley & Sons, inc.

- Rosado, A., (2010). **Business intelligence: state of the art. Scientia et Technica** Año XVI, No 44, Abril de 2010. Universidad Tecnológica de Pereira. ISSN 0122-1701.
- Roussel, G. (2006). **Decision support systems serving the company: the secrets to a successful project**, 2006. Recuperado de http://www.symtrax.com/en/WhitePaper/EN_WhitePaper_SolutionsBI_SQ.pdf. Fecha de consulta junio 2016.
- Silberschatz, A., Korth, H. y Sudarshan, S. (2006). **Fundamentos de Bases de Datos** (4ª edición). McGraw Hill, Madrid, España, 2006.
- Suarez, C. (2012). **Los Sistemas de Información en la Empresa – ERP** McGraw Hill, Madrid, España
- Tibiño, S. (2014). **Sistema de Inteligencia de Negocios para Apoyar la Toma de Decisiones del Proceso Encuesta Simple**. Tesis para optar al título de ingeniero en Computación presentada ante la Universidad Simón Bolívar. Caracas, Venezuela.
- Toro, F. (2004). **Sistemas de soporte a decisiones para la creación de modelos numéricos hidroneumáticos**. Revista EIA, ISSN 1794-1237 Número 2 p. 53-65. Agosto 2004 Escuela de Ingeniería de Antioquia, Medellín. Colombia
- Tramullas, J.,(1997). **Los sistemas de informa clon: una reflexión sobre información, sistema y documentación**. Revista General de Información y Documentación, Vol. 7, n.0 1. Servicio de Publicaciones Universidad Complutense. Madrid, 1997
- Vásquez, F. (2013). **Desarrollo de indicadores de gestión para el seguimiento de la Planificación de Proyectos utilizando una plataforma de Inteligencia de Negocios**. Tesis para optar al título de ingeniero en Computación presentada ante la Universidad Simón Bolívar. Caracas, Venezuela.
- Zorrilla, M. (2010). **Introducción al Business Intelligence**. Guía Teórica Universidad de Cantabria.