



Un ejemplo práctico de almacén de datos en Sistemas DJ

A practical example of a data warehouse in DJ Systems

Um exemplo prático de um data warehouse em DJ Systems

Diego Javier Bastidas-Logroño ^I
diego.bastidas@esPOCH.edu.ec
<https://orcid.org/0000-0003-3924-7468>

Evelin Marisol Rosero-Ordóñez ^{II}
marisol.rosoero@esPOCH.edu.ec
<https://orcid.org/0000-0001-9024-7725>

Oscar Danilo Gaviláñez-Álvarez ^{III}
oscar.gavilanez@esPOCH.edu.ec
<https://orcid.org/0000-0002-7245-5640>

Correspondencia: diego.bastidas@esPOCH.edu.ec

Ciencias Técnicas y Aplicadas
Artículo de Investigación

* **Recibido:** 25 de marzo de 2022 * **Aceptado:** 20 de abril de 2022 * **Publicado:** 27 de mayo de 2022

- I. Ingeniero en Gerencia de Sistemas, Magíster en Tecnologías de la Información, Magíster en Gestión de Proyectos. Docente Investigador, Escuela Superior Politécnica de Chimborazo (ESPOCH), Riobamba, Ecuador.
- II. Ingeniera en Sistemas. Magister en Sistemas de Información Gerencial. Docente Investigador, Escuela Superior Politécnica de Chimborazo (ESPOCH), Riobamba, Ecuador.
- III. Ingeniero en Sistemas, Magíster en Interconectividad de Redes. Docente Investigador, Escuela Superior Politécnica de Chimborazo (ESPOCH), Riobamba, Ecuador.

Resumen

Este trabajo pretende analizar los aspectos pertinentes al Data Ware House, tomando como caso práctico el sistema desarrollado para la empresa SISTEMASDJ (www.sistemasdj.com). El realizar la implementación de una óptima aportando información que permitirá mejorar su análisis. La metodología está orientada a amortiguar el tedio que provoca seguir pasos sin comprender el porqué de la ejecución de una data ware house que posibilite a los usuarios de la empresa contar con mejores prestaciones e información y mejorar su trabajo. Hoy en día, el uso del DW en el mundo es un boom que nos proporciona información clave para la toma de decisiones, mejorar la calidad de las decisiones tomadas permitiendo mayor flexibilidad y rapidez en el acceso a la información.

Palabras clave: Análisis de datos; Almacén de datos

Abstract

This work tries to analyze the pertinent aspects to the Data Ware House, taking as a practical case the system developed for the company SISTEMASDJ (www.sistemasdj.com). Realization of the implementation of an optimal supply of information that improves its performance. The objective of the methodology is to cushion the tedium caused by following steps without understanding the reason for the execution of a data storage house that allows the users of the company to have better services and information and improve their work. Today, the use of DW in the world is a boom that provides us with key information for decision-making, improving the quality of decisions made, allowing greater flexibility and speed in accessing information..

Keywords: Data Analysis; Data Warehouse

Resumo

Este trabalho pretende analisar os aspectos relevantes do Data Ware House, tomando como caso práctico o sistema desenvolvido para a empresa SISTEMASDJ (www.sistemasdj.com). Realizar a implementação de um ótimo fornecimento de informações que irão melhorar sua análise. A metodologia visa amortecer o tédio causado por seguir etapas sem entender o motivo da execução de um data ware house que permita aos usuários da empresa ter melhores recursos e informações e melhorar seu trabalho. Hoje, o uso de DW no mundo é um boom que nos fornece informações

fundamentais para a tomada de decisões, melhora a qualidade das decisões tomadas, permitindo maior agilidade e agilidade no acesso às informações.

Palavras-chave: Análise de dados; Armazem de dados

Introducción

Este trabajo pretende analizar los aspectos pertinentes al Data Ware House, tomando como caso práctico el sistema desarrollado para la empresa de empeños SISTEMASDJ (www.sistemasdj.com). El realizar la implementación de una óptima aportando información que permitirá mejorar su performance. La metodología está orientada a amortiguar el tedio que provoca seguir pasos sin comprender el porqué de la ejecución de una data warehouse que posibilite a los usuarios de la empresa contar con mejores prestaciones e información y mejorar su trabajo. Hoy en día. Brinda una información fiable entre todos los departamentos de la empresa para que el Data Warehouse sea exitoso se necesita una limpieza continua, transformación e integración de los datos. Además, requiere sistemas, aplicaciones y almacenamiento específico. SISTEMASDJ una Pymes, en la cual se ha optado por seleccionar un conjunto significativo cercano a las necesidades de la misma. Según la investigación realizada, no existe ningún servicio igual o similar al propuesto que permita disponer de servicios de data warehouse, Cabe destacar que, de la inteligencia de negocios (business intelligence) se desprenden el análisis de big data y minería de datos. De esta forma, big data es una colección de conjuntos de datos muy grandes con una gran diversidad de tipos y de difícil procesamiento mediante plataformas de procesamiento de datos tradicionales. A su vez, su análisis se considera como un subproceso dentro del proceso general de obtención de ideas de una gran cantidad de datos. En consecuencia, el análisis de big data permite una toma de decisiones basada en evidencias e implementación de soluciones, puesto que las empresas necesitan procesos eficientes para convertir grandes volúmenes de datos rápidos y diversos en percepciones significativas. (Mukherjee, Kumar, & Goswami, 2019). Además, en el área de marketing de acuerdo a la captura datos abundantes sobre el comportamiento del consumidor en tiempo real y encuentra patrones ocultos en los datos; porque éstos brindan información conductual sobre los consumidores y los especialistas en marketing traducen esas ideas en la ventaja dentro del mercado. (Colin, 2012).

A continuación, se plantea el siguiente objetivo General: “Implementar la data ware house para

tener flexibilidad y rapidez en el acceso a la información en la empresa SISTEMASDJ”.

Una vez analizado el objetivo general y para dar cumplimiento al mismo se proponen los siguientes Objetivos Específicos:

- Estudio Preliminar del Data Ware House
- Identificar las características del Data Ware House
- Implementar el Data Ware House

Metodología

Diseño de investigación

Marco Metodológico

En este capítulo se presenta la metodología que permitió desarrollar el proyecto del data ware house en la empresa Sistemasdj.

Tipo De Investigación

Este proyecto puede catalogarse en su primera instancia como una investigación de tipo exploratoria ya que se realizará lo siguiente:

Se formula un problema: Dificultad al buscar la información sobre aplicaciones desarrolladas en tecnología y plataformas poco usadas en el País sobre data ware house. Se reúne información sobre el tema: se recopilará información de diversas fuentes bibliográficas y electrónicas para hacer un estudio detallado del tema. Se aclaran conceptos: es muy importante tener bien claros muchos conceptos de la tecnología que se va a aplicar antes de empezar a experimentar con la misma.

Una vez que la parte exploratoria haya concluido, se catalogará a la investigación como experimental por las siguientes razones:

Se provocará un fenómeno con el fin de probar algo: una vez que se sintetice la información usando Data Ware House, se establecerá las respectivas modificaciones, mejoras e implementación. Se requiere de un método lógico y sistemático para realizar experiencias científicas: el mejoramiento de la investigación académica se demostrará a partir de la experimentación con las diversas tecnologías y plataformas que Data Ware House brinda para implementarlas con open source.

Fuentes de Información

Se las ha clasificado en dos:

Fuentes Primarias

Fuentes Secundarias

Fuentes Primarias

Es necesario utilizar las siguientes técnicas de recopilación de información:

Observación: Esta técnica junto con la experiencia adquirida en prácticas pre profesionales anteriores será una importante fuente de información.

Fuentes Secundarias

Para el desarrollo del presente trabajo se recopilará la suficiente información escrita y en formato electrónico referente al tema, la que puede ser catalogada en las siguientes categorías:

Libros: libros formato digital de Data Warehouse y otras tecnologías de desarrollo de software en ambientes virtuales, manuales Data warehouse.

E-books: proporcionarán la información más actualizada acerca de Data Ware House

Revistas, artículos y documentos técnicos: información relacionada con Data Ware House.

Documentación en pantalla: ayudas y documentación principalmente sobre los paquetes y kits de desarrollo que sean utilizados para la ejecución de este trabajo.

Método de Investigación Exploratorio

Se reunirá toda la información sobre la implementación de Data Ware House, como sea posible.

Estructuras de dn Data Warehouse

La arquitectura de una data warehouse puede ser dividida en tres estructuras simplificadas: básica, básica con un área de ensayo y básica con área de ensayo y data marts. Con una estructura básica, sistemas operativos y archivos planos proporcionan datos en bruto que se almacenan junto con metadatos. (Mukherjee, Kumar, & Goswami, 2019) Los usuarios finales pueden acceder a ellos para su análisis, generación de informes y minería. Al añadir un área de ensayo que se puede colocar entre las fuentes de datos y el almacén, ésta proporciona un lugar donde los datos se pueden limpiar antes de entrar en el almacén. Es posible personalizar la arquitectura del almacén para diferentes grupos dentro de la organización. Se puede hacer agregando data marts, que son sistemas diseñados para una línea de negocio en particular. Se pueden tener data marts

separados para ventas, inventario y compras, por ejemplo, y los usuarios finales pueden acceder a datos de uno o de todos los departamentos del departamento. (Back, Goodman , & Hyde , 2013)

Resultados y Discusión

Análisis e Interpretación de Resultados

Dimensiones del Sistema

Figura.Nº1 Dimensiones de la empresa

| Procesos del negocio | fecha | skus | formas | clientes | provincias | categorias | tipos | ubicaciones | conceptos |
|---------------------------------------|-------|------|--------|----------|------------|------------|-------|-------------|-----------|
| Órdenes de compra | X | X | X | | | X | | X | |
| Entregas al centro de distribución | X | X | | | | X | | | |
| Inventario del centro de distribución | X | X | | | | X | | | |
| Entregas hacia el Almacén | X | X | | | | | | | |
| Inventario del Almacén | X | X | | X | X | | X | | |
| Despacho en Ventas | X | X | X | X | X | | X | X | X |
| TOTAL | 5 | 5 | 2 | 2 | 2 | 3 | 2 | 2 | 1 |

Fuente: Diego Javier Bastidas Logroño, Edwin Villamarín Rojas

Preguntas

- ¿Cuál es el problema del área, departamento?

No tener un sistema de data analítica que provea generar reportes sin estar consultando al área de sistemas.

- ¿Cómo hace el negocio para abordarlo ahora?

Utilizar por medio de este proyecto el desarrollo de data analítica por medio de pentaho server

- ¿Los datos están disponibles?

Se necesitaba siempre consultar al área de tics este tipo de reportes, vía email o llamada, ya que siempre se debía consultar a la base de datos esta área por medio de consultas SQL.

- ¿Cómo se podría usar los resultados?

Se requerida hacer un análisis de los productos tienen más salida y menos salida, para recibir

- ¿Quién es el custodio de los datos?

El área de tics

- ¿Quién es el dueño de los datos?

Gerencia

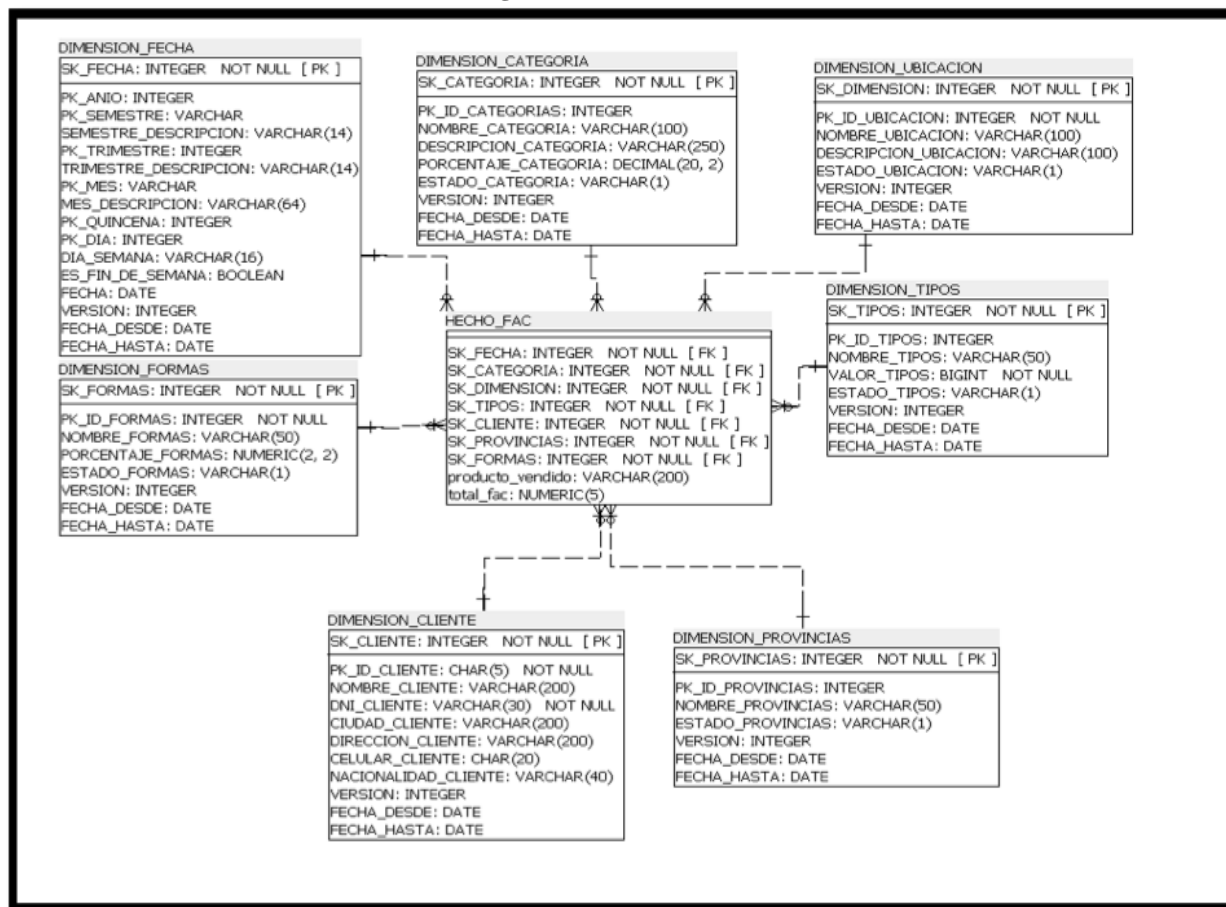
Hefesto

- Los objetivos y resultados esperados en cada fase se distinguen fácilmente y son sencillos de comprender.
- La piedra fundamental la constituyen los requerimientos de los usuarios, por lo cual, se adapta con facilidad y rapidez a los cambios del negocio.
- Reduce drásticamente la resistencia al cambio, ya que involucra a los usuarios finales en cada etapa para que tomen decisiones respecto al comportamiento y funciones del DW, y además expone resultados inmediatos.
- Utiliza modelos conceptuales y lógicos, los cuales son sencillos de interpretar y analizar. Es independiente del tipo de ciclo de vida que se emplee para contener la metodología. Es independiente del software/hardware que se utilicen para su implementación.
- Cuando se culmina con una fase, los resultados obtenidos se constituyen en la entrada de la fase siguiente.
- Se aplica en Data Warehouse Directorios de páginas blancas o amarillas

Mapeo

En el Data Source de la empresa analizada, el proceso de venta está representado por el siguiente Diagrama de Entidad Relación (representa la información a través de Entidades, Relaciones, Cardinalidades, Claves, Atributos y Jerarquías de generalización. (Moser, Bruner, & Day, 2017)

Imagen 1. Power Architect Diseño



Fuente: Diego Javier Bastidas Logroño, Edwin Villamarín Rojas

Resultados del Proyecto

El resultado principal del proyecto es una herramienta poderosa y flexible, dentro del marco de trabajo de una plataforma Data Analítica, capaz de mantener gestionar los datos para la empresa. En la tabla hecho utilizamos un inner join, que permite combinar registros de una o más tablas en una base de datos. En el Lenguaje de Consultas Estructurado hay tres tipos de JOIN: interno, externo y cruzado. El estándar ANSI Del SQL especifica cinco tipos de JOIN: INNER, LEFT OUTER, RIGHT OUTER, FULL OUTER y CROSS

```

selectconcat(substring(ve.ven_fechaactual,1,4),substring(ve.ven_fechaactual,6,2),
substring(ve.ven_fechaactual,9,2)) as sk_fecha, ve.ven_fechaactual,
date_format(ve.ven_fechaactual,"%Y") as pk_anio_ped, ca.cat_id, ub.ubi_id, ti.tip_id, pr.pro_id,
    
```



```
cl.cli_id, fo.for_id, dve.vde_identificacionbien, dve.vde_precioventa FROM ventas ve INNER JOIN ventadetalles dve on ve.ven_id=dve.ven_id INNER JOIN categorias ca on ca.cat_id = dve.cat_id INNER JOIN ubicaciones ub on ub.ubi_id = dve.ubi_id INNER JOIN tipos ti on ti.tip_id=ve.tip_id INNER JOIN clientes cl on cl.cli_id=ve.cli_id INNER JOIN provincias pr on pr.pro_id=cl.pro_id INNER JOIN formas fo on fo.for_id=ve.for_id
```

Conclusiones

- La implementación de data analítica en la actualidad en las empresas es necesario ya que sin esta la empresa no podría cumplir sus objetivos económicos a alcanzar ni de creación.
- Es relevante este tipo de estudio ya que se ven nuevas técnicas de análisis de datos, que para muchos son desconocidas hasta el momento.
- El mundo actual de la tecnología de la información se encuentra en el umbral de una revolución, los trabajos poco cualificados se están automatizando y pronto se volverán superfluos, el trabajo en campos tradicionales como el mantenimiento de aplicaciones, la gestión de la infraestructura y la programación manual se está trasladando a trabajos de alta calidad.

Referencias

1. Back, W., Goodman , N., & Hyde , J. (2013). *O'Reilly Media, Inc.* Recuperado el 2021, de Mondrian in Action: Open source business analytics: <https://learning.oreilly.com/library/view/mondrian-in-action/9781617290985/>
2. Colin, W. (2012). *O'Reilly Media, Inc.* Recuperado el 2021, de <https://learning.oreilly.com/library/view/information-visualization-3rd/9780123814647/xhtml/CHP001.html#CHP001tit1>
3. Moser, A., Bruner, J., & Day, B. (2017). *O'Reilly Media, Inc.* Recuperado el 2021, de Geospatial Data and Analysis: <https://learning.oreilly.com/library/view/geospatial-data-and/9781491984314/titlepage01.html>
4. Mukherjee, S., Kumar, A., & Goswami, S. (2019). *O'Reilly Media, Inc.* Recuperado el 2021, de Big Data Simplified: <https://learning.oreilly.com/library/view/big-data-simplified/9789353941505/chapter-01.html>

- a. ation-visualization-3rd/9780123814647/xhtml/CHP001.html#CHP001tit1
5. Mukherjee, S., Kumar, A., & Goswami, S. (2019). *O'Reilly Media, Inc.* Recuperado el 2021, de Big Data Simplified: <https://learning.oreilly.com/library/view/big-data-simplified/9789353941505/chapter-01.html>