

## ANÁLISIS E IMPLEMENTACIÓN DE PENTAHO EN EL CONTROL DE DATOS

### ANALYSIS AND IMPLEMENTATION OF PENTAHO IN DATA CONTROL

AUTORES: Paul Javier Álvarez Sagubay<sup>1</sup>

Chystiam Vinicio Toapaxi Acosta<sup>2</sup>

Manuel Fabricio Reyes Wagnio<sup>3</sup>

Tania Jeesenia Peralta Guaraca<sup>4</sup>

DIRECCIÓN PARA CORRESPONDENCIA: [paul.alvarezs@ug.edu.ec](mailto:paul.alvarezs@ug.edu.ec)

Fecha de recepción: 27-11-2017

Fecha de aceptación: 11-01-2018

#### RESUMEN

El manejo de los datos en su proporción, abarca de manera desmedida las estrategias de negocios dentro del entorno actual, las proyecciones administrativas que se enfocan son las necesidades del cliente que se tiene en cuenta al momento de emplear una herramienta que pueda medir los datos para su utilidad y beneficio. La administración estudia las organizaciones y las técnicas que se van ejecutar en el planteamiento del flujo de información, recursos y usos, según los objetivos propuestos en una estructura que obtenga máxima productividad haciendo de este, un proceso sistemático. Pentaho ofrece un rendimiento eficaz dentro de un mercado competitivo, que ofrece en su funcionalidad la implementación en una organización, el cual permite interactuar a través de los datos la toma de decisiones que ayuden en el crecimiento, que estable una evaluación, en la ponderación de un resultado. La optimización en la funcionalidad de una herramienta facilita los procesos como un factor estratégico para identificar oportunidades para generar estrategias e implementar la información en la toma de decisiones, para dar soluciones en el mercado de negocios, para definir el control de gastos de recursos y el monitoreo, análisis y administración de la información. En conclusión se debe establecer un seguimiento del mercado, en las medidas de las necesidades que se desea cubrir en los clientes o los requerimientos que exige el mercado para tener la capacidad de dirigir una empresa y se encamine en la dirección indicada.

**PALABRAS CLAVE:** Indicadores, Optimización, Pentaho, Inteligencia de Negocios.

---

<sup>1</sup> Máster en Administración de Empresas, Ingeniero en Sistemas Computacionales. Universidad de Guayaquil. Ecuador.

<sup>2</sup> Máster en Administración de Empresas con mención en Sistemas de Información Empresarial, Ingeniero en Sistemas Computacionales. Ecuador. E-mail: [ctoapaxi@armada.mil.ec](mailto:ctoapaxi@armada.mil.ec)

<sup>3</sup> Máster en Administración de Empresas, Ingeniero en Sistemas Computacionales. Universidad de Guayaquil. Ecuador. E-mail: [manuel.reyesw@ug.edu.ec](mailto:manuel.reyesw@ug.edu.ec)

<sup>4</sup> Máster en Gerencia y Docencia en Educación Superior, Ingeniero en Sistemas Computacionales. Universidad de Guayaquil. Ecuador. E-mail: [tania.peraltag@ug.edu.ec](mailto:tania.peraltag@ug.edu.ec)

## ABSTRACT

The handling of the data in its proportion, covers in an excessive way the business strategies within the current environment, the administrative projections that focus are the needs of the client that is taken into account when using a tool that can measure the data for its utility and benefit. The administration studies the organizations and techniques that will be executed in the planning of the flow of information, resources and uses, according to the proposed objectives in a structure that obtains maximum productivity making this a systematic process. Pentaho offers efficient performance within a competitive market, which offers in its functionality the implementation in an organization, which allows interacting through the data making decisions that help in growth, which establishes an evaluation, in the weighting of a result. Optimization in the functionality of a tool facilitates processes as a strategic factor to identify opportunities to generate strategies and implement information in decision making, to provide solutions in the business market, to define the control of resource expenditures and the monitoring, analysis and information management. In conclusion, a market monitoring should be established, in the measures of the needs that are to be covered by the clients or the requirements that the market demands in order to have the capacity to run a company and be directed to the indicated direction.

**KEYWORDS:** Indicators; Optimization; Pentaho; Business Intelligence.

## INTRODUCCIÓN

En la actualidad el ambiente socioeconómico es competitivo y complejo, en el que la crisis económica ha puesto presión sobre los líderes que manejan las diferentes empresas de productos, obligan a que las estrategias en la implementación del mercado sea más competitiva, se mantiene Una lucha dentro de un sector del mercado en el que se desenvuelven los métodos tradicionales, en las toma de decisiones se han vuelto relativamente ineficaces, ya que los responsables de dirigir las actividades de las empresas e instituciones se enfrentan a situaciones cada vez más complicadas y dinámicas, que requieren de soluciones creativas y prácticas apoyadas en una base cuantitativa sólida.(Hernández, Villalva et al. 2016)

Las empresas se preocupan constantemente por tomar decisiones que les permitan optimizar sus resultados de manera que se obtenga el mayor beneficio. La evidente dificultad de tomar estas decisiones ha hecho que el hombre se dirija a la búsqueda de herramientas o métodos que le permita lograrlo, de acuerdo a los recursos disponibles y a los objetivos que persigue. (Hernández, Aguilar et al. 2017)

Si comparamos épocas pasadas los informes de datos se generaban a través de reportes operativos, con la inversión del recurso humano y tiempo, en la presentación de resultados podemos ver cuánto hemos evolucionado en la forma como se maneja los datos de información, maniobrando de una manera sistemática la utilidad requerida, en la integración de datos optimizando la continuidad dentro de un proceso actual, ya que da un enfoque de proyección vigente en los negocios, que mejora la productividad.

El cambio en el hábito varia el consumo que posibilita llevar adelante la gestión de información que integra y automatiza muchas de las prácticas de negocio asociadas con los aspectos operativos o productivos de una empresa, a través de esta plataforma

llamada Pentaho. El cual responde a estas preguntas: ¿Qué nuevo producto tendría aceptación en la compra? ¿Los productos con menor rotación serian? ¿Cuánto y cuales productos se necesitan producir? ¿El financiamiento de crédito está acorde a la ganancia que se espera obtener? ¿Se mantiene una cartera de crédito efectivo?

Resultados y discusión:

### PENTAHO- Herramienta de Business Intelligence

Pentaho es un conjunto de técnicas que permite integrar, depurar, transformar y consolidar los datos de una empresa, ya que un datawarehouse es un repositorio de datos que proporciona una visión global, común e integrada de los datos de la organización (2010).

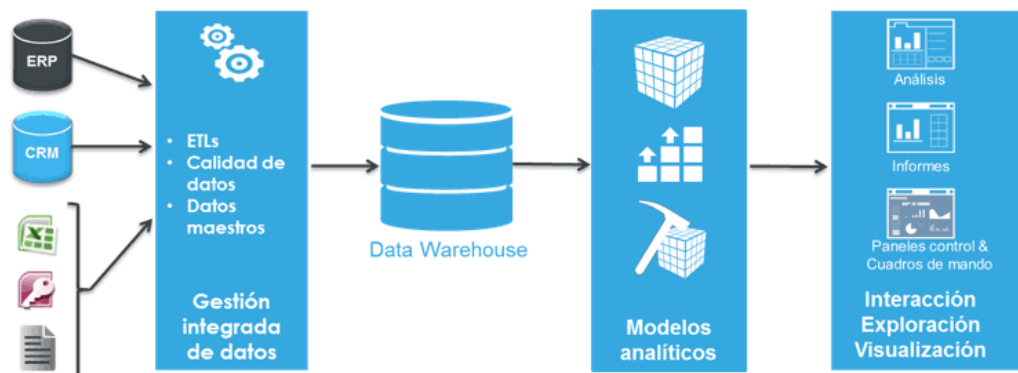


Grafico 1. Planificación de una solución de BI

Fuente: Business-Intelligence-esquema

### Inteligencia a partir de los datos

- Es un proceso que comienza con preguntas, y las respuestas son resultados de más preguntas o de subsecuentes interacciones del proceso.
- La información se convierte en Inteligente cuando se procesa y se analiza, para lo cual definimos los dos ejes sobre los cuales vamos a trabajar, las dimensiones que son las posibles, las categorizaciones de nuestra información y las métricas que son las variables cuantitativas de nuestro proceso operacional.
- El proceso de recolección son datos en crudo, en el cual se integra a un formato utilizable para el análisis, que forma una nueva base de datos.
- Finalmente tenemos el análisis y la producción de los Indicadores de Gestión.

- Y el resultado final es la producción de respuestas “inteligentes”, en nuestro propio contexto.

Hay gran cantidad de empresas que se han beneficiado en muchos niveles de esta nueva tecnología, y se enfoca dentro de un macro de la situación, en su análisis en cada unidad estratégica de negocio, es decir, les ha permitido detectar el desempeño de cada producto para los diferentes mercados en los que funcionan. La información también les permite comparar la rentabilidad de diversos productos a través de los ciclos de inventarios y de margen.



Gráfico No. 2 - Pentaho

Fuente: Business-Intelligence-esquema

Los sistemas de Inteligencia de Negocios le pueden ayudar a la compañía a reducir los ciclos de desarrollo de productos, agilizar operaciones, afinar campañas de marketing y mejorar relaciones con clientes y proveedores, todo lo cual significa menores costos y mayores márgenes de utilidad.

#### Características del Producto “Pentaho”

- Posee Tecnología Cloud
- Nos brinda una arquitectura abierta y basada en estándares, lo que permite una fácil integración con la infraestructura existente.
- Soporta grandes volúmenes de almacenamiento (BigData).
- El proceso es en línea, se refleja de forma automática en el usuario.
- Análisis visual e interactivo para una investigación en la información.
- Proporciona servicio de Dashboard, es decir, gráficos que están unidos al cubo de información e indicadores. No tiene costo extra, está incluido en el paquete.

- Permite hacer indicadores de gestión, cuyas métricas son establecidas por la alta gerencia.
- Trae su propio motor de reportes BIRT (reportes operativos)

Si define que la integración de datos de Pentaho crea un ambiente accesible e inteligente el cual prepara y modifica datos, a través de una plataforma cloud. El almacenamiento de los datos puede ser visualizado en la web. Empodera a la organización en el flujo de datos, con un sistema de análisis y modelos predictivos, en la toma de decisiones significativas, que mejora el proceso de productividad.

Algunas organizaciones no cuentan con información inmediata sobre sus unidades estratégicas de negocios y para poder evaluar el desarrollo de sus productos utilizan el manejo de su cartera de clientes, cobranzas, utilizando bases de datos en el cual se arroja información que es manipulada para poder realizar gráficos e interpretar los resultados para su debida toma de decisiones, que los impulse hacia el logro de sus objetivos organizacionales.

La funcionalidad de Pentaho brinda reportes interactivos, integra los datos de manera continua en su diseño, análisis y soporte que ofrece la seguridad en la información, ya que cubre la fisura de los datos.



Gráfico No. 2 - Estructura del sistema Pentaho

Fuente: Business-Intelligence (2014)

La estructura del sistema Pentaho respaldará el desempeño de las actividades y funciones que realiza el personal perteneciente a diversas áreas y niveles jerárquicos. Además, se encargará de brindar las facilidades del proceso que son necesarias para contribuir a la carga, operación y explotación de la aplicación desde su particular interés.

El ambiente de implementación de Pentaho está basado en Java. Eso hace que Pentaho sea una solución muy flexible para cubrir una amplia gama de necesidades empresariales, los módulos de la plataforma Pentaho BI incluyen: Reporting, Business Analytics, Dashboards, Cloud Business Analytics, Data Integration.

En el siguiente cuadro comparamos las ventajas de las diversas tecnologías de Inteligencia de Negocios que ocupan un gran sector del mercado Ecuatoriano, tomando como referencia las características más importantes.

## ANÁLISIS E IMPLEMENTACIÓN DE PENTAHO EN EL CONTROL DE DATOS

Cuadro No.1 Comparación de productos para selección del software

Fuente: Datos alcanzados en el estudio

|   | Cognos | Pentaho | Analysis Services | Business Object |
|---|--------|---------|-------------------|-----------------|
| Acceso Vía Web  | √      | √       |                   | √               |
| Informes Parametrizados                               |        | √       |                   | √               |
| Scheduling  | √      | √       | √                 |                 |
| Acceso a Fuentes de Datos relacionales                | √      | √       | √                 | √               |
| Acceso a Fuentes de Datos Heterogéneas                |        | √       |                   |                 |
| Permite crear reportes en su visor                    | √      |         |                   |                 |
| Uso intuitivo del cubo                                | √      |         |                   |                 |
| Integración con aplicaciones y Portales               |        | √       |                   |                 |
| Definición modular de informes                        |        | √       |                   |                 |
| Entorno de diseño gráfico amigable                    | √      | √       | √                 | √               |
| Capacidad de uso de Templates                         | √      |         |                   | √               |
| Puede ser embebible                                   |        | √       |                   |                 |
| Pocos Recursos  |        | √       | √                 |                 |
| Limitación en creación de cubo                        | √      |         |                   |                 |
| Alto grado de Procesamiento en creación del cubo      | √      |         |                   |                 |
| Portabilidad  | √      | √       |                   | √               |
| Escalabilidad   |        | √       |                   |                 |
| Integración   |        | √       |                   |                 |
| Fácil de Entender                                     | √      | √       | √                 | √               |
| Análisis Ad Hoc                                       |        | √       |                   | √               |
| Optimización para respuesta interactiva               |        | √       |                   |                 |
| Creación de Kpi                                       |        | √       |                   | √               |
| Drill a reportes de soporte                           |        | √       |                   | √               |
| Alertas basadas en reglas de negocio                  |        |         |                   | √               |
| Multiplataforma                                       | √      | √       |                   | √               |
| Posee herramienta de extracción propia                |        | √       |                   | √               |
| Uso de tecnologías estándar:<br>Java, Xml, JavaScript |        | √       |                   |                 |
| Posee Motor de Datamining                             |        | √       |                   |                 |
| Posee administrador de seguridad                      | √      | √       | √                 | √               |
| Posee administrador de Reportes                       | √      | √       |                   | √               |
| Auditoría de uso y rendimiento                        |        | √       |                   |                 |
| Alto Costo de Inversión                               | √      |         | √                 | √               |

## Materiales y métodos

Para que los datos de origen, reflejen las necesidades que debemos conocer para la construcción se utilizó la metodología HEFESTO, que puede ser utilizada en cualquier ciclo de vida que cumpla con la condición antes declarada. Con el fin de que se llegue a una total comprensión de cada paso o etapa, se acompañará con la ejecución dentro de

una empresa, para demostrar los resultados que se deben obtener y las valoraciones obtenidas en el proceso.

### *Objetivo al aplicar la Metodología*

Mejorar el flujo de información dentro de una empresa, para que pueda llevar a cabo un análisis profundo de la gestión operativa, con la finalidad de poder tomar de forma óptima decisiones que ayuden al crecimiento de la misma.

### *Desarrollo de la aplicación de una metodología*

Puede adaptarse muy bien a cualquier ciclo de vida de desarrollo de software, con la salvedad de que para algunas fases en particular, las acciones que se han de realizar serán muy diferentes. Lo que se debe tener muy en cuenta, es no entrar en la utilización de metodologías que requieran fases extensas de reunión de requerimientos y análisis, fases de desarrollo monolítico que conlleve demasiado tiempo y fases de despliegue muy largas. Lo que se busca, es entregar una primera implementación que satisfaga una parte de las necesidades, para demostrar las ventajas del DataWareHouse y motivar a los usuarios.

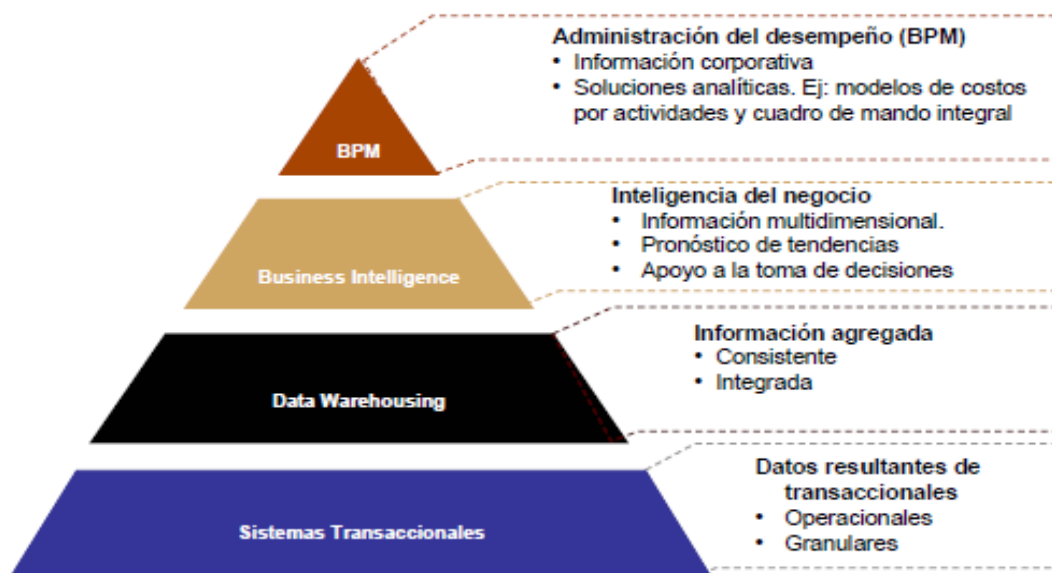


Gráfico No.4 – Estructuración de la Información

Fuente: Business-Intelligence (2014)

Como se puede apreciar, se comienza con la recolección de las necesidades de información en los usuarios y se obtienen las preguntas claves del negocio.

Luego, se deben identificar los indicadores resultantes de los interrogativos y sus respectivas perspectivas de análisis, mediante las cuales se construirá el modelo conceptual de datos del DataWareHouse.

En un estudio reciente, Luan & Jiang (2014) implementan un DataWareHouse con el objetivo de ayudar en la gestión de la colección. Todos los autores reconocen en



general que los sistemas tradicionales de biblioteca no son suficientes para un adecuado soporte en la toma de decisiones, y que una eficiente gestión de la biblioteca necesita de la creación de sistemas y bases de datos especializados que integren eficientemente toda la información existente en una biblioteca.

Y se define la selección y aplicación en la Metodología Hefestos, el cual precisa el tipo de esquema que se implementará.

HEFESTO es una metodología propia, cuya propuesta está fundamentada en una muy amplia investigación, comparación de metodologías existentes, experiencias propias en procesos de confección de almacenes de datos. Cabe destacar que HEFESTO está en continua evolución, y se han tenido en cuenta, como gran valor agregado, todos los feedbacks que han aportado quienes han utilizado esta metodología en diversos países y con diversos fines.(2010)



Gráfico No.5– Metodología Hefesto. Pasos

Fuente: Business-Intelligence (2014)

Como se puede apreciar en el gráfico, se comienza con la recolección de las necesidades de información de los usuarios y se obtienen las preguntas claves del negocio. Luego, se deben identificar los indicadores resultantes de los interrogativos y sus respectivas perspectivas de análisis, mediante las cuales se construirá el modelo conceptual de datos del Data Warehouse. Después, se analizarán los OLTP para determinar cómo se construirán los indicadores, señalar las correspondencias con los

datos fuentes y para seleccionar los campos de estudio de cada perspectiva. Una vez hecho esto, se pasará a la construcción del modelo lógico del depósito, en donde se definirá cuál será el tipo de esquema que se implementará.

Consecutivamente, se confeccionarán las tablas de dimensiones y las tablas de hechos, para luego efectuar sus respectivas uniones. Por último, se utiliza las técnicas de limpieza, calidad de datos, procesos ETL, se definirán políticas y estrategias para la Carga Inicial del DW y su respectiva actualización. (2010).

### Proceso de Herramienta en acción

Los Jobs son las tareas iniciales que definen el flujo de ejecución de los diversos ETL.

En estos ETL consideramos los procesos de extracción del modelo de .Ventas.

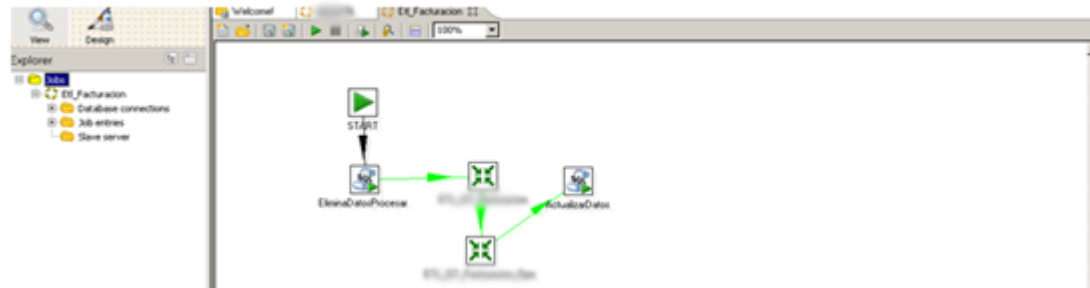


Gráfico No. 6 Jobs de Carga ETL

Fuente: Datos alcanzados en el estudio

Los ETL se dividen en dos grupos, los ETL que manejan la información de las dimensiones, las cuales generalmente pasan solo por un proceso de carga no de transformación.



Gráfico No. 7 – Carga de Dimensiones ETL

Fuente: Datos alcanzados en el estudio

Y los ETL que contienen la lógica del negocio con la cual se realiza la carga de las medidas, estas necesitan pasar por un proceso de transformación y carga.

Finalmente, se utilizan técnicas de limpieza y calidad de datos, se definirán políticas y estrategias para la Carga Inicial del DataWareHouse y su respectiva actualización.

ANÁLISIS E IMPLEMENTACIÓN DE PENTHAHO EN EL CONTROL DE DATOS

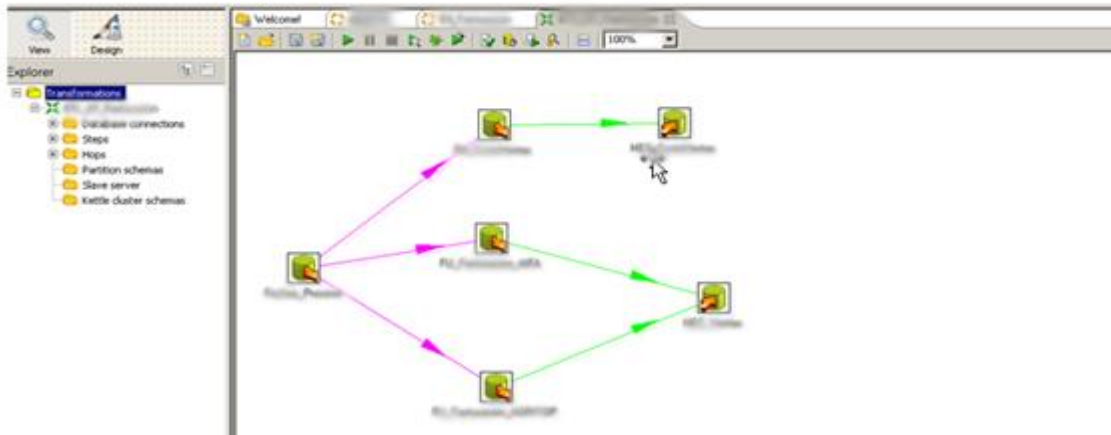
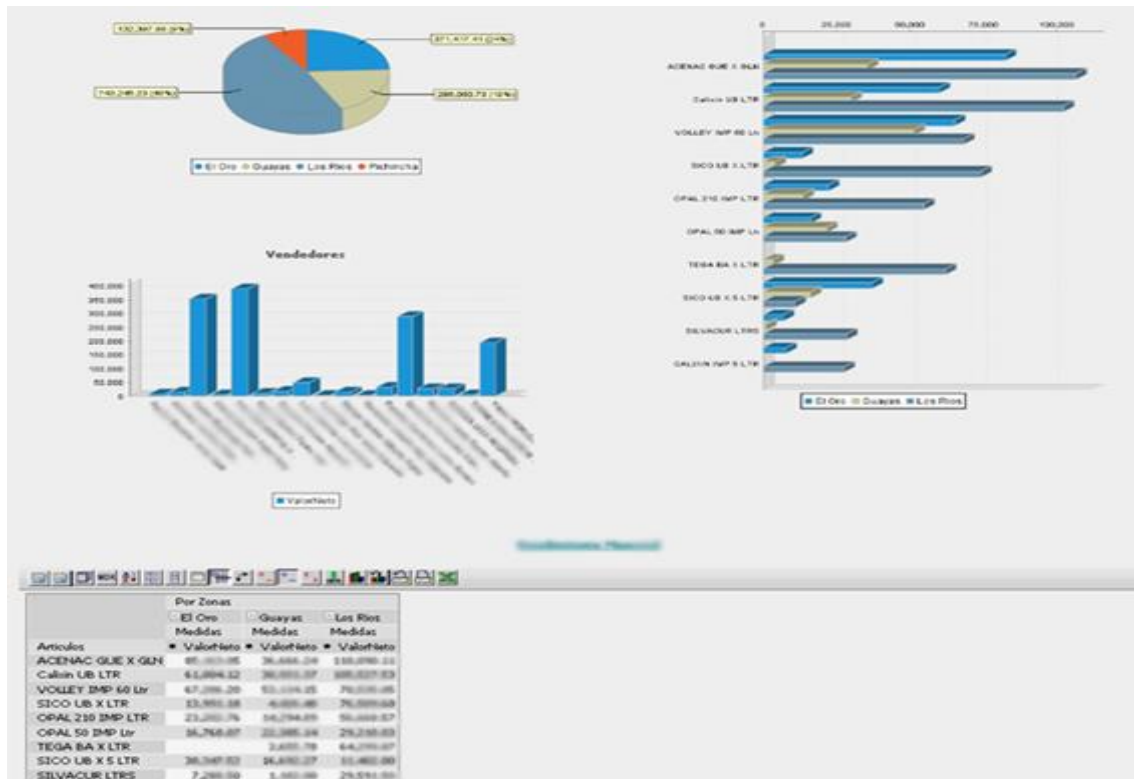


Gráfico No. 8 Proceso de Cargas de Métricas ETL

Fuente: Datos alcanzados en el estudio

Cuadro 2: Resultado de Caso



Fuente: Datos alcanzados en el estudio

Resultado de Caso:

- Se debe manejar mejor, el flujo de información para poder realizar un mejor análisis profundo con el concurso de la herramienta Pentaho, con la finalidad de tomar de forma óptica, las decisiones que ayuden al crecimiento de la misma.

- Contribuye de manera significativa a disminuir la cantidad de procesos manuales para la emisión de informes, a su vez, arroja información vital que antes no se disponía para el crecimiento del negocio y optimización de recursos, lo que permite la facilidad y rapidez al cruzar información y de esta forma poder estar en capacidad de desarrollar un plan estratégico efectivo, lo cual permitirá a la empresa potenciar, su ventaja competitiva y detectar a tiempo oportunidades de negocios que se traduzcan en un mejor posicionamiento dentro del mercado local.
- Mejora la satisfacción por parte de los empleados a nivel operativo debido a que existen la herramienta para facilitar su trabajo lo cual significan disminución de horas de trabajo, y la presentación eficaz de información solicitada.

Al no contar con información precisa sobre su cartera de clientes, desarrollo de cuentas claves y evolución de productos, se acostumbra a inferir sobre la situación global, en lugar de trabajar sobre bases sólidas y reales que permitan direccionar las inversiones de recursos hacia sectores que lo ameriten.

Al poder ya establecer las condiciones, se utilizan los datos de ponderación para satisfacer las estrategias de desarrollo, ya que permite un crecimiento en los diferentes sectores del mercado, que estudia la muestra y define un porcentaje de incremento/ganancia. Y fija las competencias en sus campos de acción.

## CONCLUSIONES

Las organizaciones líderes son las que están a la vanguardia de la tecnología, aquellas que tienen la certeza de contar con información real, precisa en el tiempo requerido para la toma de decisiones, y el uso de una herramienta eficaz como Pentaho que maximiza su productividad, en los constante cambios de entorno actual, porque logra predefinir mejores resultados en los objetivos propuestos, y a generar utilidad, a largo plazo en una organización.

Para llevar a un flujo más real en los datos se deben crear políticas que ayuden a un mejor manejo de los datos en el proceso de la obtención de información, que cumple con un seguimiento de indicadores, para establecer estrategias en la inteligencia de negocios y su mercado.

## RECOMENDACIONES

Es conveniente que los directivos y en general, se involucren en el uso de las tecnologías para dar mayor apertura en vista en que contribuyen la perspectiva sobre las amenazas, y las transforma en oportunidades para una mejor inserción en el mercado. Aplica estrategias que estén orientadas a la satisfacción del usuario-cliente, establece las condiciones que permiten, continuar con el proceso de expansión sobre una de las bases de datos sólidas.

Es vital la creación de un plan estratégico basado en el conocimiento de la situación actual de la empresa, que identifica y potencia las marcas con las cuales se puede lograr un mejor desarrollo.

A través del conocimiento de patrones de consumo y tendencias proporcionados Pentaho puede segmentar la base de datos, lo cual permite identificar nichos de mercado e implementar estrategias de desarrollo que estén alineados con la visión de una organización e impulsa el desarrollo de los datos claves para aumentar la producción.

#### REFERENCIAS BIBLIOGRÁFICAS

Cebotarean, Elena. (2010). Business intelligence (BI):

Haro Valle, V. A., & Pérez Rocano, W. R. (2014). Tesis. Recuperado a partir de <http://dspace.ucuenca.edu.ec/handle/123456789/19878>

Hiltbrand, Troy Burke, Marsha, Hiltbrand, Troy, Burke, Marsha (2011) How Gamification will change Business Intelligence .Journal: Business Intelligence Journal

Khan, RA, Quadri. (2012) .Business intelligence: an integrated .Journal: Business Intelligence Journal

Journal: Journal of Knowledge Management, Economics and Information Technology Negash, Solomon. (2010). Journal: Communications of the Association for Information Systems.

Hernández, N. B., et al. (2016). "Responsabilidad social, pobreza, derecho ambiental y naturaleza." Revista Magazine de las Ciencias. ISSN 2528-8091 1(2): 01-06.

Hernández, N. B., et al. (2017). "El desarrollo local y la formación de la competencia pedagógica de emprendimiento. Una necesidad en el contexto social de Cuba." Revista Didasc@ lia: Didáctica y Educación. ISSN 2224-2643 8(5): 213-226.

Pesantez Sarmiento, J. (2010). Tesis. Recuperado a partir de <http://dspace.ucuenca.edu.ec/handle/123456789/717>

Roldan, Maria Karina (2010). Pentaho 3.2 Data Integration Beginner's .Guide.ISBN.978-1-847199-54-6.

Understanding, Customers (2011).Business Intelligence Software