

Análisis de datos para mejorar la toma de decisiones en la distribución de agua potable del cantón Ibarra, utilizando Business Intelligence

Autor-Andrés Fernando CÁRDENAS PEPINÓS¹

Universidad Técnica del Norte, Av. 17 de Julio, Ibarra, Imbabura

afcardenasp@utn.edu.ec

Resumen. En las instituciones públicas del país, el análisis de información es parte de las estrategias básicas para mejorar sus procesos con el objetivo de mejorar la entrega de sus productos y servicios. En este aspecto la data generada día a día por los sistemas transaccionales se ha convertido en un activo importante para cada una de las instituciones, por lo tanto, el desarrollo e implementación de sistemas para la toma de decisiones son más frecuentes. La presente investigación, pretende adaptar el concepto de integridad de datos de varias instituciones y realizar una base de datos única que ayude a la toma de decisiones en varios ejes, con el objetivo de mejorar la distribución de agua potable tomando en cuenta factores de crecimiento poblacional, área de catastros, niveles de educación etc. Para llevar a cabo esta investigación se la realizó con un enfoque cualitativo descriptivo, en la cual se basa en revisiones bibliográficas de aplicación de análisis de datos, para la integración de datos se realizó entrevistas con los usuarios finales de la herramienta (stakeholder), con la finalidad de presentar una herramienta que se adapte a los requerimientos de cada uno de ellos y mejorar la toma de decisiones. La implementación de este tipo de herramientas permitió conocer los patrones de crecimiento poblacional, así como las necesidades en cuanto a la distribución de agua potable.

Palabras Claves

Análisis de Datos, Sistema Integrado, Inteligencia de Negocios, Almacenes de Datos, Toma de Decisiones.

Abstract. In the public institutions of the country, the analysis of information is part of the basic strategies to improve processes with the aim of improving the delivery of their products and services. In this aspect, transactional data have become an important element for each of the institutions, for the development and implementation of systems for making more frequent decisions. The present investigation intends to adapt the concept of the integrity of the data and to realize a unique database that helps to make

decisions in several axes, with the objective of improving the distribution of drinking water taking into account factors of population growth, cadastral area, education levels, etc. In order to carry out this research, the analysis is carried out with a qualitative descriptive approach, based on the bibliographical references of the data analysis application. For the integration of the data, the interviews with the final users of the tool were carried out. (stakeholder), in order to present a tool that adapts to the requirements of each of them and improves decision making. The implementation of this type of tools will allow to know the patterns of population growth, as well as the needs regarding the distribution of drinking water.

Keywords

Data Analysis, Integrated System, Business Intelligence, Data Warehouses, Decision Making.

1. Introducción

Esta investigación es parte del macro proyecto “Ibarra Verde”, proyecto que tiene el objetivo principal de crear una herramienta de análisis, basado en un mapeo integral de la zona urbana de la ciudad de Ibarra, en los Andes Norte del Ecuador, que sirva como modelo para una planificación urbana técnicamente alta, y que, a la vez, respete el sentido cultural, las normas sociales y el estado ambiental de la región. (Frolich, y otros, 2017)

En colaboración con empresa de agua potable (EMAPA-I—Empresa Municipal de Agua Potable y Alcantarillado—Ibarra), se obtuvo datos del uso de agua potable, de cada usuario doméstico igual que industrial, sobre los últimos tres años, debidamente digitalizados. Como parte del análisis de datos, estamos incorporando la ubicación geográfica, con coordenadas precisas de cada usuario para poder analizar la distribución de uso por sectores geográficos y tendencias históricas del uso de agua

y energía eléctrica en un mapeo integrado e interactivo. (Frolich, y otros, 2017)

Se analizó las tendencias temporales y la distribución geográfica del uso de este recurso, los datos se unificaron en una base de Oracle para aplicar las herramientas de Oracle BI (Business Intelligence) una poderosa herramienta basada en tecnología Web con interfaz gráfica amigable para los usuarios y que permitió reunir todos los datos y transformarlos en información que sirve para toma de decisiones. (Frolich, y otros, 2017)

EMAPA-I, es una empresa que a través de una administración eficiente, suministra servicios de agua potable y saneamiento con calidad, para mejorar las condiciones de vida de los habitantes del cantón, enmarcados en valores, principios y normativas vigentes. (emapa, 2017)

Se ha realizado un previo análisis general de todas las instituciones que intervienen en esta investigación, descubriendo que cada una maneja su información aisladamente y por lo tanto no pueden realizar análisis de información conjunta para toma de decisiones.

Al tener varios sistemas y bases de datos transaccionales en cada una de las instituciones, que generan datos importantes con el objetivo de cumplir y ejecutar procesos en cada una de estas.

El no tener una infraestructura lógica y física donde se puedan consolidar y analizar la información generada por las diversas entidades a través de una herramienta de Business Intelligence (BI), se dificulta la presentación de información depurada, consolidada y sobre todo de fácil visualización, lo que provoca realizar trabajos extensos para realizar análisis propios de la empresa y más aún cuando los análisis son dependientes de la información proveniente de otras instituciones.

Al ser el agua potable un recurso natural no renovable, necesita ser tratado de una manera especial en cuanto a su distribución, consumo y desaprovechamiento. Por lo cual si no se implementan acciones que ayuden a optimizar el consumo y distribución, sobre todo a disminuir el desaprovechamiento, tendremos problemas de abastecimiento en un futuro, afectando el estilo de vida de la población del cantón Ibarra que va en aumento día a día.

Al tener gran y valiosa información generada por las diversas instituciones se podrían generar análisis los cuales permitan unir diversos índices que permitan mejorar los niveles de consumo y sobre todo de desperdicio de agua potable.

2. Materiales y Métodos

Para el desarrollo de la investigación se realizó utilizando la inteligencia de negocios, en la cual detallamos ciertos enunciados a continuación:

Inteligencia de Negocios: Desde un punto de vista más pragmático, y asociándolo directamente con las tecnologías de la información, podemos definir Business Intelligence como el conjunto de metodologías, aplicaciones y tecnologías que permiten reunir, depurar y transformar datos de los sistemas transaccionales e información desestructurada (interna y externa a la compañía) en información estructurada, para su explotación directa (reporting, análisis OLTP / OLAP, alertas...) o para su análisis y conversión en conocimiento, dando así soporte a la toma de decisiones sobre el negocio.

La inteligencia de negocio actúa como un factor estratégico para una empresa u organización, generando una potencial ventaja competitiva, que no es otra que proporcionar información privilegiada para responder a los problemas de negocio: entrada a nuevos mercados, promociones u ofertas de productos, eliminación de islas de información, control financiero, optimización de costes, planificación de producción, análisis de perfiles de clientes, rentabilidad de un producto concreto, etc.

Los sistemas y componentes del BI se diferencian de los sistemas operacionales en que están optimizados para preguntar y divulgar sobre datos. Esto significa típicamente que, en un data warehouse, los datos están des-normalizados para apoyar consultas de alto rendimiento, mientras que en los sistemas operacionales suelen encontrarse normalizados para apoyar operaciones continuas de inserción, modificación y borrado de datos. En este sentido, los procesos ETL (extracción, transformación y carga), que nutren los sistemas BI, tienen que traducir de uno o varios sistemas operacionales normalizados e independientes a un único sistema des normalizado, cuyos datos estén completamente integrados.

Aplicación de Inteligencia de Negocios para análisis de consumo de agua potable.

La aplicación de políticas en cuanto al consumo y distribución del agua potable es imprescindible para aprovechar al máximo el recurso y evitar el desabastecimiento. La importancia del agua potable en la actualidad se da debido al desabastecimiento en algunas épocas del año por factores ambientales y por el crecimiento de las urbes lo que significa un aumento de necesidad de agua en diversos sectores.

Con la utilización de la inteligencia de negocios se analiza la información para identificar consumos anómalos que no cumplan con los patrones establecidos como litros diarios por persona, consumos máximos mensuales por tarifas, márgenes de pérdidas, etc.

Para el análisis de información se realizó utilizando técnicas de recopilación de datos de diversas instituciones que nos faciliten comprender las tendencias de consumo.

Con la información recopilada se procedió aplicar la técnica OLAP, propia de la inteligencia de negocio, que permite realizar la construcción de los cubos multidimensionales, de los cuales obtenemos los elementos

de análisis o facts como son máximos consumo, porcentajes de perdidas, crecimiento por zonas, y sectores.

Para complementar la investigación se aplicó la técnica de Data Mining, propia de la inteligencia de negocios, la cual nos permitió identificar patrones y correlaciones en cada uno de los análisis realizados.

Metodología Kimbal

Para el desarrollo de este proyecto se hace referencia a la metodología propuesta por Kimbal, la cual es empleada para la construcción de almacenes de datos DW orientados a un determinado ámbito de análisis. (WEB 2.0 - MediaWiki, 2014).

Como se aprecia en la Ilustración propuesta por Brito M, se representa el ciclo de vida dimensional del negocio.

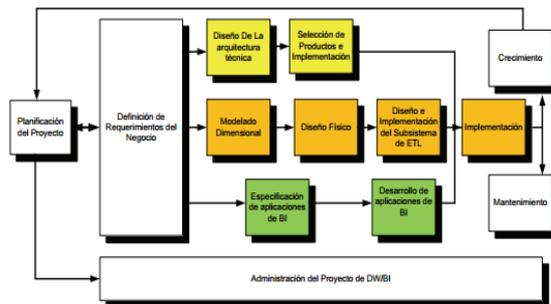


Ilustración 1: Ciclo de Vida Kimball.

Fuente:

<http://ucv2014inteligenciadenegocios.blogspot.com/2014/09/post-3-metodologia-de-kimball.html>

La metodología se basa en el Ciclo de Vida Dimensional del negocio, que se encuentra basado en cuatro principios básicos.

Al escribir el texto, el estilo "Básico" tiene que ser utilizado.

Concentrarse en el negocio.- Se basa en identificar los requerimientos del negocio y su valor asociado, para desarrolla las relaciones sólidas con el negocio y enfocar los análisis de mejor manera. (Brito, 2014).

Construir una infraestructura de información adecuada.- Se enfoca en el diseño de la base de información única, integrada y de fácil uso, con un rendimiento alto para reflejar los requerimientos de implementación de una forma adecuada. (Azuaje, 2014)

Realizar entregas en incrementos significativos.- Se lo aplica en la creación de los almacenes de datos DW, en incrementos entregables cada cierto tiempo, por recomendación podría ser en plazos de 6 a 12. Los

incrementos deben ser identificados para determinar el orden de aplicación, se recomienda la utilización de metodologías ágiles de construcción de software. (Azuaje, 2014)

Ofrecer la solución completa.- Se lo efectúa al proporcionar los elementos necesarios para realizar las entregas de valor necesario a los usuarios finales. Se debe tener en este punto los almacenes de datos sólidos, bien diseñados y con calidad, se entregará las herramientas de consulta ad hoc, aplicaciones de informes, análisis avanzado, capacitación, soporte y documentación. (Azuaje, 2014)

3. Resultados

Al realizar análisis de los datos en cuanto al consumo de agua potable, se pudo identificar factores por los cuales el consumo no era el adecuado, se pudo evidenciar suministros que tenían tarifas con beneficios de uso residencial y que en realidad el consumo eran para otras actividades muy alejadas a la realidad. Al realizar los cruces con variables como número de población se identificó que los consumos tomando en cuenta el número de habitantes era demasiado alto al promedio sugerido por la OMS, que es de 100 litros diarios por persona.

Se pudo identificar que los índices de pérdida de agua potable superaban el 30% es decir que de cada 100 litros de agua potable distribuidos por cada planta solo 70 son facturados, los 30 restantes son perdidas ya sean por fallas en el sistema de distribución o por utilización de agua de forma fraudulenta.

La plataforma BI está desarrollada en base a los requerimientos de cada una de las entidades requerentes, con lo cual los resultados han sido evaluados por parte del personal de las entidades involucradas, se ha enfocado el análisis de resultados a cada área de negocio, se han realizado análisis tanto de resultados obtenidos en la herramienta BI, así como también el análisis estadístico a partir de la información generada por el BI mediante la utilización de software especializado con es el PSP.

Para facilitar el análisis por parte de los usuarios finales se desarrolló un Dashboard el cual cuenta con informes interactivos de cada una de las necesidades de cada empresa, en el cual se puede acceder a información necesaria para la toma de decisiones.

Al generar conocimiento a partir de información de varias instituciones nos ayuda a comprender que el desarrollo de las zonas urbanas y por ende de la utilización de recursos, depende de varios factores socioeconómicos y sobre todo de la organización que se lleve al momento de planificar la expansión de las ciudades.

Los resultados deben expresar los resultados de los experimentos descritos en el Material y Métodos y presentar las pruebas que apoyan tales resultados, sea en forma de figuras, tablas o en el mismo texto.

Los resultados deben poder ser vistos y entendidos de forma rápida y clara. Es por ello por lo que la construcción de esta sección debe comenzar por la elaboración de las tablas y figuras, y sólo posteriormente redactar el texto pertinente en función de ellas. El primer párrafo de este texto debe ser utilizado para resumir en una frase concisa, clara y directa, el hallazgo principal del estudio. Esta sección debe ser escrita utilizando los verbos en pasado.

Aunque se utiliza mucho la voz pasiva o el impersonal (“se ha encontrado que...”) como prefieren algunos editores. En el texto se deben citar todas las tablas y figuras, en caso de ser tomadas de otros estudios deberán constar las referencias bibliográficas. Todas las tablas y figuras deben tener su respectiva leyenda.

En cuanto al formato, esta sección se puede organizar en subtítulos, y cada uno de estos no puede subdividirse nuevamente. Para un ejemplo referirse a los anexos.

4. Conclusiones

Se ha desarrollado la herramienta BI bajo la utilización de la metodología de Kimball, la cual nos ayudó a realizar una implementación sistematizada de la herramienta, con la cual se pudo aplicar para realizar análisis de consumos en diversos aspectos, al obtener resultados se pudo identificar varios factores por los cuales los consumos de agua potable son altos, entre los cuales están la utilización de agua en actividades que son diferentes a los de consumos residenciales así como el desperdicio ya sean estos por consumo no registrado hurto o fallas en los sistemas de distribución.

Se utilizó técnicas propias de la inteligencia de negocios como son OLAP y Data Mining, lo que permitió realizar análisis detallados y sobre todo de fácil comprensión por los usuarios finales, se realizó análisis basados en tendencias de consumo, máximos, mínimos, así como de correlaciones entre datos de diferentes instituciones.

Al realizar la integración de datos de EMAPA con las diferentes instituciones, se consolidó información valiosa para los análisis, en los cuales se obtuvo conocimiento sobre las tendencias de consumo de agua potable, así como los índices de pérdidas o desperdicio de agua potable que nos permiten tomar medidas para mejorar la distribución y disminuir las pérdidas.

La implementación de esta herramienta mejoró notablemente el tiempo en el cual se puede acceder a información importante para la toma de decisiones, aunque en la actualidad existe una resistencia en la implementación de herramientas gerenciales de toma de decisiones en

instituciones públicas, esto generado por la falta de visión y sobre todo por los costos de inversión que representa, tanto para la adquisición como tal y falta de personal capacitado.

La aplicación de controles de mandos permitió a los usuarios finales acceder a la información mediante informes interactivos fácilmente con el objetivo de obtener resultados de forma rápida y de fácil comprensión.

En la actualidad las empresas privadas son las que están a la vanguardia de la implementación de herramientas BI, sobre todo las empresas que están el mercado de la banca y venta de productos, en la parte pública las empresas pioneras que tiene implementado esta infraestructura son las de generación y comercialización de energía eléctrica.

Agradecimientos

A la gloriosa Universidad Técnica del Norte, quienes me brindaron la oportunidad de obtener mi Maestría en Ingeniería de Software, así ayudándome a crecer tanto personal como profesionalmente con todos los conocimientos impartidos.

A mi director de tesis Ing. Jorge Caraguay Msc, por todos los consejos impartidos, para cada día avanzar y culminar con éxito el presente trabajo de grado.

A los compañeros del proyecto Ibarra verde, por la dedicación y entrega presta para que el trabajo sea el mejor y poder alcanzar nuestros objetivos.

A la Empresa de Agua Potable y Alcantarillado de la ciudad de Ibarra, por la ayuda incondicional en cuanto a la obtención de la información.

Referencias Bibliográficas

- [1] Abreu, J. L. (2014). El Método de la Investigación.
- [2] Abril Frade, D. O., & Pérez Castillo, J. N. (2017). Estado actual de las tecnologías de bodega de datos y OLAP aplicadas a bases de datos espaciales. REVISTA INGENIERÍA E INVESTIGACIÓN.
- [3] Azuaje, A. (2014). Metodología de Kimball. CARACAS.
- [4] Brito, D. (26 de 01 de 2014). Inteligencia de Negocios. Obtenido de <http://inteligenciadenegociosdiegobrito.blogspot.com/2014/01/la-metodologia-de-kimball.html>
- [5] Castro Roza, F. E. (2013). Indicadores de gestión para la toma de decisiones basada en Inteligencia de Negocios. Caldas: TIA.
- [6] Cota, M. (12 de Julio de 2011). BI-SPAIN. Obtenido de BI-LATINO: <http://www.bi-spain.com/articulo/71397/oracle/energia-utilities/mexico-df-ahorra-11-millones-de-litros-de-agua-por-hora-gracias-al-business-intelligence-de-oracle-video-entrevista-exclusiva>
- [7] Cueva Andrade, C., Jerez Cevallos, S., Díaz Zuñiga, P., & Ron Egas, M. (2014). DETERMINACIÓN DE LA LÍNEA BASE PARA LA APLICACIÓN DE SISTEMAS DE BI (BUSINESS INTELLIGENCE) EN EL ECUADOR, DENTRO DE UN PROGRAMA PARA EL USO DE TECNOLOGÍA DE AVANZADA EN LA ADMINISTRACIÓN PÚBLICA Y PRIVADA. Quito: ESPE.

- [8] DATAPRIX. (s.f.). Dataprix, el portal profesional sobre Software y Servicios IT para la empresa. Obtenido de <http://www.dataprix.com/24-or-genes-antecedentes-las-bases-datos>
- [9] emapa. (3 de 10 de 2017). Emapa Ibarra - Filosofía. Obtenido de Emapa Ibarra: <http://www.emapaibarra.gob.ec/index.php/filosofia>
- [10] Esri. (s.f.). ArcGis Resources. Obtenido de <http://resources.arcgis.com/es/help/getting-started/articles/026n00000014000000.htm>
- [11] Frolich, L. M., Caraguay Procel, J. A., Rosales, O., Cárdenas Pepinós, A. F., Castro, P., Llanga Ochoa, P. I., . . . Sarmiento, F. (2017). Ecología humana y la urbe inteligente: Utilizando mapeo interactivo para el análisis socio uso de agua y de energía eléctrica en Ibarra. *CIENCIA*, 227-240.
- [12] GAMBOA CRUZADO, J. (2015). DESARROLLO DE BUSINESS INTELLIGENCE, BASADO EN LA METODOLOGÍA DE RALPH KIMBALL, PARA MEJORAR EL PROCESO DE TOMA DE DECISIONES EN EL ÁREA DE ADMISIÓN DE LA UNIVERSIDAD AUTÓNOMA DEL PERÚ. Lima.
- [13] Gómez Morales, A. J. (2013). Inteligencia de negocios, una ventaja competitiva para las organizaciones. Trujillo.: Revista CIENCIA Y TECNOLOGÍA.
- [14] GUEVARA VEGA, C. P. (2015). DESARROLLO DE UNA PLATAFORMA DE BUSINESS INTELLIGENCE PARA FACILITAR EL ANÁLISIS DE DATOS DE LAS COMPETENCIAS GENERALES DE FORMACIÓN APLICADAS EN EL DESEMPEÑO LABORAL DE LOS EGRESADOS DE LA UNIVERSIDAD TÉCNICA DEL NORTE. LATACUNGA.
- [15] Ibarra, G. A. (12 de 11 de 2017). *Gobierno Autónomo Descentralizado de San Miguel de Ibarra*. Obtenido de <https://www.ibarra.gob.ec/web/>
- [16] Instituto Nacional de Estadística y Censos. (15 de 11 de 2017). *Instituto Nacional de Estadística y Censos*. Obtenido de Instituto Nacional de Estadística y Censos: <http://www.ecuadorencifras.gob.ec/institucional/home/>
- [17] Lascano Sumbana, K. V. (2016). DISEÑO DE UN MODELO PARA DETERMINAR LA DEMANDA DE AGUA POTABLE PARA JUNTAS DE AGUA POTABLE DE LA PROVINCIA DE TUNGURAHUA: CASO DE APLICACIÓN PARROQUIA HUAMBALO. Ambato.
- [18] Oracle. (1 de 04 de 2014). *Technetwork*. Obtenido de <http://www.oracle.com/technetwork/es/articles/database-performance/pdbs-backup-recovery-dbpitr-2192211-esa.html>
- [19] Oracle. (2016). Guía del Usuario de Oracle Business Intelligence Enterprise Ediion. Oracle.
- [20] PowerData. (19 de 09 de 2015). PowerData Solutions S.L. Obtenido de <https://blog.powerdata.es/el-valor-de-la-gestion-de-datos/bid/406542/qu-son-las-bases-de-datos-multidimensionales>
- [21] Quest. (2014). Quest. Obtenido de <https://www.quest.com/toad/>
- [22] Rodríguez Almeida, A. M., & Da Silva Camargo, S. (2015). ACADEMIC ANALYTICS: APLICANDO TÉCNICAS DE BUSINESS INTELLIGENCE SOBRE DATOS DE PERFORMANCE ACADÉMICA EN ENSEÑANZA SUPERIOR. Aracaju: EXATAS E TECNOLÓGICAS.
- [23] Sinnexus. (2016). Sinnexus - Business Intelligence Informática estratégica. Obtenido de http://www.sinnexus.com/business_intelligence/arquitectura.aspx
- [24] Universidad Poliecnica de Valencia. (s.f.). Museo Informatica. Obtenido de <http://histinf.blogs.upv.es/author/>
- [25] WEB 2.0 - MediaWiki. (30 de 01 de 2014). Inteligencia de Negocios. Obtenido de <http://inteligenciadenegociosval.blogspot.com/2014/01/metodologia-de-kimball.html>