

Revista EDUCATECONCIENCIA.

Volumen 30, No.37

ISSN: 2007-6347

E-ISSN: 2683-2836

Periodo: Octubre- diciembre 2022

Tepic, Nayarit. México

Pp. 28-43

Doi: <https://doi.org/10.58299/edu.v30i37.574>

Recibido: 5 de agosto del 2022

Aprobado: 06 de octubre del 2022

Publicado: 20 de diciembre del 2022

**Modelo de *Data Mart* para mejorar la productividad de las empresas privadas,
Caso empresa inmobiliaria**

**Data Mart model to improve the productivity of private companies, Case study
in real estate company**

Víctor Darío Sosa Jauregui

*Universidad Nacional de San Antonio Abad del
Cusco, Perú.*

victor.sosa@unsaac.edu.pe

<https://orcid.org/0000-0002-0341-1206>

Vanessa Maribel Choque Soto

*Universidad Nacional de San Antonio Abad del
Cusco, Perú.*

vanessa.choque@unsaac.edu.pe

<https://orcid.org/0000-0002-4484-0074>

Modelo de *Data Mart* para mejorar la productividad de las empresas privadas, Caso empresa inmobiliaria

Data Mart model to improve the productivity of private companies, Case study in real estate company

Víctor Darío Sosa Jauregui

Universidad Nacional de San Antonio Abad del Cusco, Perú.

victor.sosa@unsaac.edu.pe

<https://orcid.org/0000-0002-0341-1206>

Vanessa Maribel Choque Soto

Universidad Nacional de San Antonio Abad del Cusco, Perú.

vanessa.choque@unsaac.edu.pe

<https://orcid.org/0000-0002-4484-0074>

Resumen

El artículo se desprende de la investigación de Sosa (2017) y utiliza un *Data Mart* para mejorar la productividad en las empresas. De acuerdo con Tavera y Ríos (2021), la industria 4.0 es un conjunto de tecnologías que las empresas requieren para impulsar respuestas rápidas en contextos dinámicos, estas tecnologías incluyen OLAP (procesamiento analítico en línea), ETL (extracción, transformación y carga), inteligencia de negocios. La investigación tiene un enfoque cuantitativo, en el que se construye un modelo de *Data Mart* con el objetivo de mejorar la toma de decisiones y la productividad en empresas privadas, tomando como caso de estudio el rubro inmobiliario, rubro con mayor proyección en el país. Las conclusiones indican que la elaboración del modelo ha servido para una reducción significativa de tiempos y como efecto mejoras en la productividad empresarial, por tanto de ser replicado se espera un aporte positivo en empresas.

Palabras clave: Análisis de datos, *Data Mart*, empresa, *ETL*, inmobiliaria, *OLAP*, procesamiento de datos.

Abstract

The article stems from Sosa (2017) research and uses a Data Mart to improve productivity in companies. According to Tavera and Ríos (2021), industry 4.0 is a set of technologies that companies require to drive rapid responses in dynamic contexts, these technologies include OLAP (online analytical processing), ETL (extraction, transformation and loading), business intelligence. The research has a quantitative approach, in which a Data Mart model is built with the aim of improving decision-making and productivity in private companies, taking the real estate sector as a case study, an area with the greatest projection in the country. The conclusions indicate that the elaboration of the model has served for a significant reduction of times and improvements in business productivity, therefore, if it is replicated, a positive contribution is expected in companies.

Keywords: company, data analysis, data mart, data processing, ETL, OLAP, real estate

Introducción

Justificación de la investigación

La investigación se fundamenta y justifica en el hecho indiscutible de que en la actualidad la empresa privadas e instituciones consideran fundamental el uso de los sistemas de información. Según Laudon (2016) las empresas utilizan sistemas de información para lograr excelencia operacional; nuevos productos, servicios y modelos de negocios, cercanía con el cliente/proveedor; toma de decisiones mejorada entre otros. El desarrollar un modelo de *Data Mart* permitirá entonces contar con una solución que produzca información oportuna y pertinente; y mostrada además de una manera gráfica para su entendimiento y aporte a la toma de decisiones y mejora de la productividad de las empresas.

Situación problemática

Arévalo *et al* (2018) menciona que el World Economic Forum reconoció que los países con mayor uso de TI experimentan una productividad empresarial siete veces mayor que los países que no invierten en TI. Cedeño & Townsend (2021) afirman que la percepción de las empresas es que las inversiones en TI generan un alto costo por lo que son vistas como un gasto y no una inversión y muchas de estas inversiones no se ajustan a las necesidades de la empresa.

En las empresas medianas se ha demostrado que el aumento de la competitividad está asociado a la adopción de tecnología (Muñoz *et al* ,2019), entonces las empresas conscientes de lo anterior están en la búsqueda de tecnologías de la información que permita la mejora de sus procesos, toma de decisiones y por ende de su productividad.

Antecedentes

Helmer *et al* (2021) proponen el proyecto “Creating and implementing a COVID-19 recruitment *Data Mart*” e indica que su innovador sistema *Data Mart* es escalable y prometedor como una solución generalizable para recomendar simultáneamente individuos de un grupo de pacientes frente a un grupo de oportunidades de ensayos urgentes.

Adicionalmente Medina & Castillo-Rojas (2018) proponen el desarrollo de un Data Mart (DM) para la obtención de indicadores de productividad académica el uso de 2 *Data Marts* satisfacen las especificaciones de requerimientos definidos en concordancia con los usuarios finales.

Así mismo Dahr *et al* (2022) proponen el Proyecto “Implementing sales decision support system using *Data Mart* based on OLAP, KPI, and Data mining approaches” la solución DSS de ventas ayuda en la toma de decisiones y ayuda a los gerentes a desarrollar nuevos conocimientos y estrategias.

En otra investigación Chávez (2021) propone “Implementación de un Datamart para el soporte de la toma de decisiones en el área comercial de una empresa privada en la ciudad de Lima – 2021” cuyo propósito fue obtener una herramienta que permita tener la información centralizada y actualizada para su posterior explotación y análisis, mejoras en la reportería que facilita un manejo óptimo de los indicadores de los negocios.

Jaleel y Abbas (2020) proponen “Design and Implementation of Efficient Decision Support System Using *Data Mart* Architecture”, investigación que propone un marco del Sistema de Soporte de Decisiones para la Empresa Pública de Información y Telecomunicaciones mediante el cual se analiza la ventaja sobre el diseño, la integración del *Data Mart* para los tomadores de decisiones.

Objetivo general

Proponer un modelo de *Data Mart* para mejorar la productividad en empresas privadas de la región Cusco, Perú, caso: rubro inmobiliario.

Objetivos específicos

- Identificar los requerimientos de análisis de información para el área de Ventas.
- Elaborar un modelo de base de datos multidimensional que permita el análisis y explotación de la información identificada.
- Construir el *Data Mart* para mostrar la información que se necesita para poder tomar decisiones estratégicas en el área de ventas.

Materiales y método

Según la clasificación de métodos de investigación de Hernández *et al* (2014) la presente investigación es de enfoque cuantitativo, porque se recolecta datos de la realidad objetiva, es secuencial y probatoria. Es de tipo correlacional, porque mide el grado de relación entre variables. Es de nivel explicativo ya que los resultados buscarán explicar el fenómeno encontrado y con diseño no experimental, ya que se realiza sin manipular variables. Las variables

consideradas en la investigación son el proceso de toma de decisiones y la aplicación del modelo de Data Mart.

Participantes

La población son las empresas privadas de la región del Cusco, Perú, según el censo 2020 del Banco Central de Reserva del Perú ascienden 1453 empresas formales clasificadas como pequeñas empresas. El caso de estudio es una empresa del rubro bienes raíces (inmobiliarias) de la región del Cusco. El rubro inmobiliario en la ciudad del Cusco siempre ha gozado de alta rentabilidad, razón por la cual se seleccionó este rubro para el caso de estudio.

Técnica e instrumentos

Se utilizó como recurso las bases de datos relacionales proporcionadas por la empresa caso de estudio.

Como tecnología se utilizará la creación de almacenes de datos que Utami *et al.* (2020) refiere es subconjunto de un almacén de datos que respalda las necesidades de información de departamentos, funcionalidad y expectativa de la presente investigación.

Procedimiento

El procedimiento consistió en las siguientes etapas:

Requerimientos

La empresa solicitó los siguientes requerimientos:

- Requerimiento 1 (RE1): Seguimiento de las visitas comparando los años y meses desde la primera visita. De esta manera se puede determinar la afluencia de clientes potenciales de la inmobiliaria desde su creación a la actualidad con la finalidad de determinar sus estrategias publicitarias por épocas del año.
- Requerimiento 2 (RE2): Observar las variaciones en las visitas para las distintas ciudades a las que pertenecen las propiedades. De esta manera se puede determinar que ciudades presentan mayor demanda en bienes inmuebles y dirigir estrategias de búsqueda de adquisición de propiedades a esas zonas.
- Requerimiento 3 (RE3): Observar visitas realizadas por empleado para así medir la productividad de sus empleados.

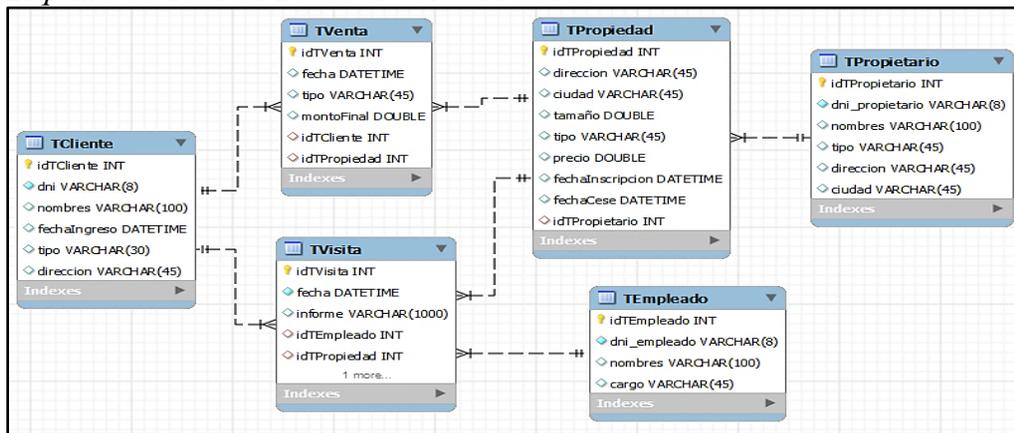
- Requerimiento 4 (RE4): Comparar las cantidades de visitas realizadas por tipo de clientes para determinar la clase de clientes potenciales de la empresa. Estos tipos son básicamente dos, Personas Naturales y Jurídicas. De acuerdo a esta información se pueden dirigir estrategias publicitarias más exclusivas y personales.

Descripción de las fuentes de datos de origen

La inmobiliaria cuenta con la siguiente estructura de base de datos relacional desarrollada en el software de gestión de bases de datos denominada MySQL, en la figura 1 se muestra el esquema.

Figura 1

Esquema de Base de Datos de la Inmobiliaria



Fuente: propia (2022)

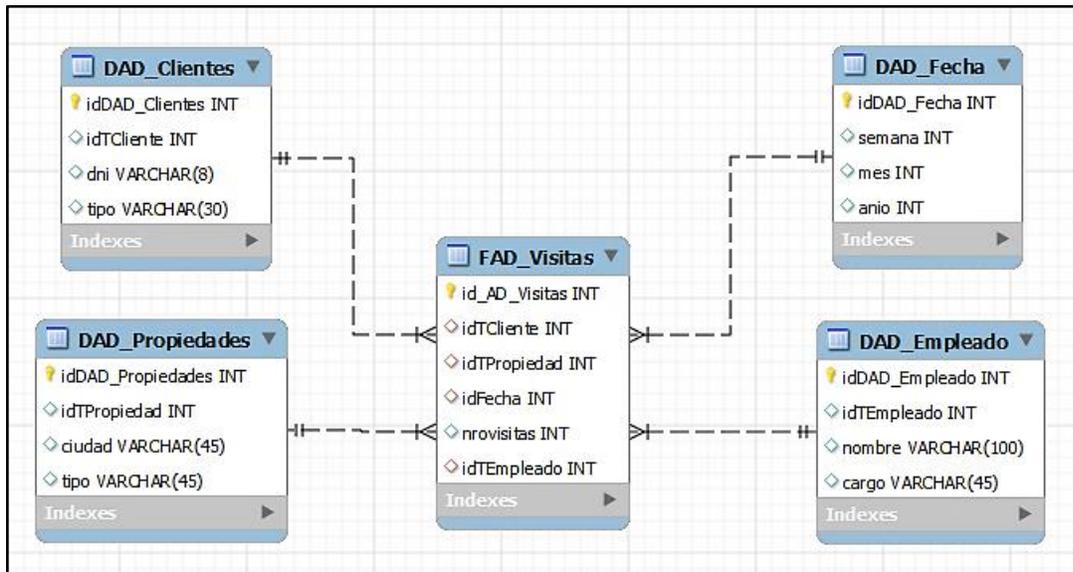
Diseño del Data Mart

El siguiente es un esquema conceptual, modelo estrella (Kimball & Ross, 2011), que fue diseñado a partir de los requerimientos indicados por la inmobiliaria en cuanto a las visitas.

Se definieron 4 dimensiones: DAD_Fecha, DAD_Empleado, DAD_Clientes de las que se extrajo el tipo y DAD_Propiedades de las que se tomara la ciudad a la que pertenece el inmueble. La dimensión filtro denominada FAD_Visitas guarda el número de visitas realizadas lo que representa el KPI (Key Performance Indicator) o indicador clave de desempeño, el esquema se visualiza en la figura 2.

Figura 2

Modelo Estrella de Base de Datos de la Inmobiliaria



Fuente: propia (2022)

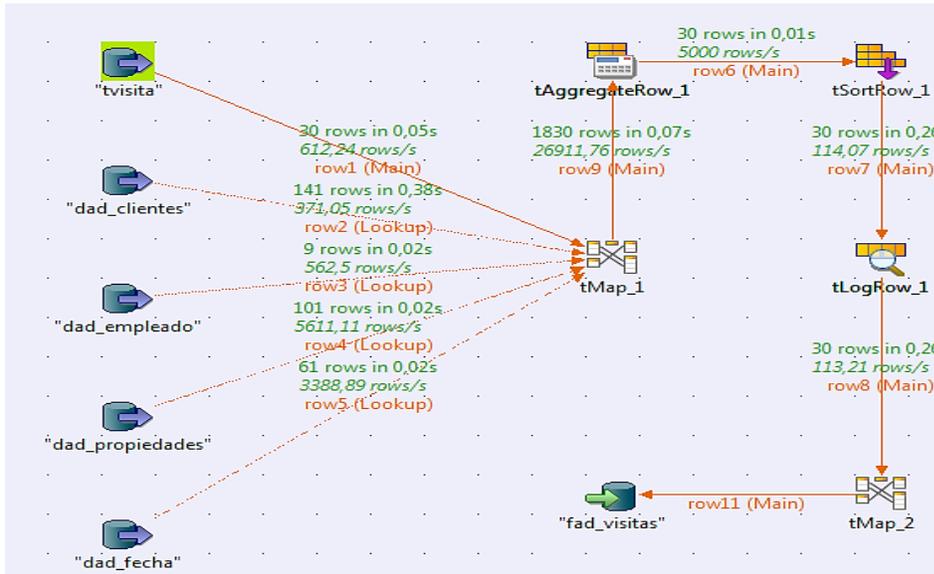
Nota: DAD_Fecha: Tabla de dimensión Fecha, DAD_Empleado: Tabla de dimensión Empleado, DAD_Clientes: Tabla de dimensión Clientes, DAD_Propiedades: Tabla de dimensión Propiedades, FAD_Visitas: Tabla de hechos Visitas.

Proceso ETL

Es el proceso que permite a las organizaciones mover datos desde múltiples fuentes, reformatearlos y limpiarlos, y cargarlos en otra base de datos, *Data Mart*, o *Data Warehouse*. En el proceso *ETL* (Extract-Transform-Load) los datos se extraen de varias fuentes y se transforman antes de cargarse en el almacén de datos (Diouf & Ndiaye, 2018, 1). Es parte fundamental de la fase de Preparación de Datos, necesaria para garantizar la calidad del conjunto de datos (Choque, 2019).

La herramienta utilizada en el proceso ETL fue Talend Open Studio (Talend Open Studio, 2021). La siguiente figura 3 muestra una captura del proceso en la tabla Vistas

Figura 3.
 Diagrama ETL de la Tabla Visitas



Fuente: propia (2022)

Nota: tAggregateRow: Tabla para agregar nueva fila, tMap_i: Tabla de mapeo, tSortRow_1: Tabla de Ordenamiento de filas, tLogRow_1: Tabla de carga de filas, tMap_2: Tabla de mapeo, fad_visitas: base de datos de visitas.

El proceso de mapeo extrae registros de cada tabla dimensión como parte del proceso ETL el resultado se visualiza en la figura 4.

Figura 4.
 Mapeo de la Tabla Dimensional Visitas



Fuente: propia (2022)

Nota: fecha: registra la fecha de la visita, idCliente: atributo que identifica al cliente, tipo_cliente: atributo que indica el tipo de cliente, idEmpleado: atributo que identifica al empleado, nombre_empleado: atributo que guarda el nombre del empleado, cargo_empleado: atributo que guarda el cargo del empleado, idPropiedad: atributo que identifica a la

propiedad, ciudad: ciudad donde se ubica la propiedad, tipo_propiedad: atributo tipo de propiedad, num_visistas: atributo cantidad de visitas.

Resultados y discusiones

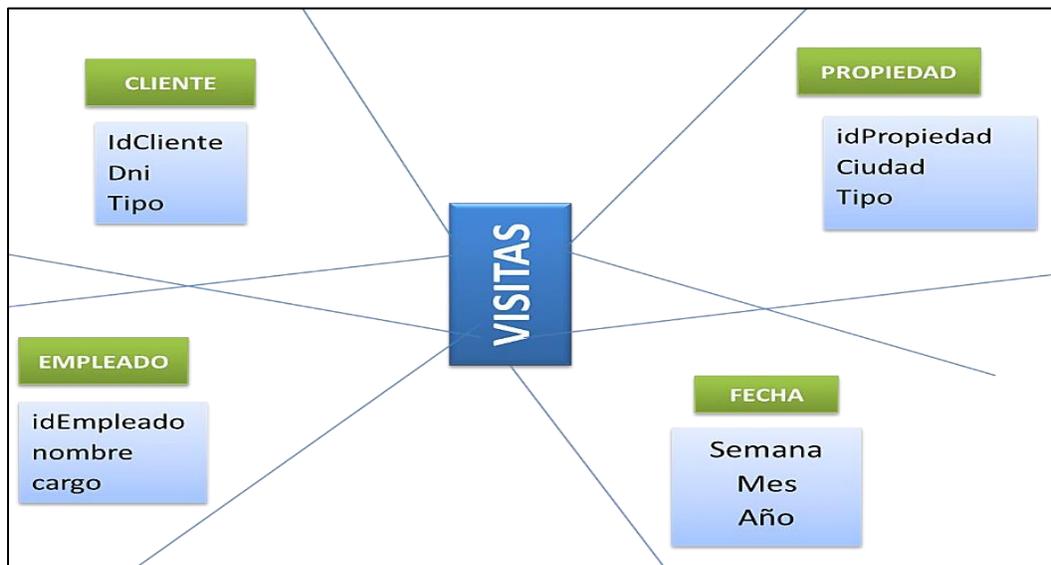
En atención a los objetivos de la investigación planteados en la introducción y a los requerimientos planteados, los resultados giran en torno a la cantidad de visitas se obtuvo un *Data Mart* con la cantidad de visitas tomando 3 dimensiones:

Cantidad de Visitas

Lo importante es analizar la cantidad de visitas realizadas por fecha (año, mes), propiedad (ciudad en la que está ubicada) y cliente (tipo de cliente que realiza la visita). Visualizar la figura 5.

Figura 5.

KPI número de visitas y sus dimensiones



Fuente: propia (2022)

En la tabla 1 se muestra la matriz de trazabilidad por requerimiento y dimensiones del *Data Mart*.

Tabla 1

Requerimientos y dimensiones del Data Mart

Requerimiento y Dimensiones	RE1	RE2	RE3	RE4
Cliente	X			X
Empleado	X		X	
Propiedad	X	X		
Fecha	X			

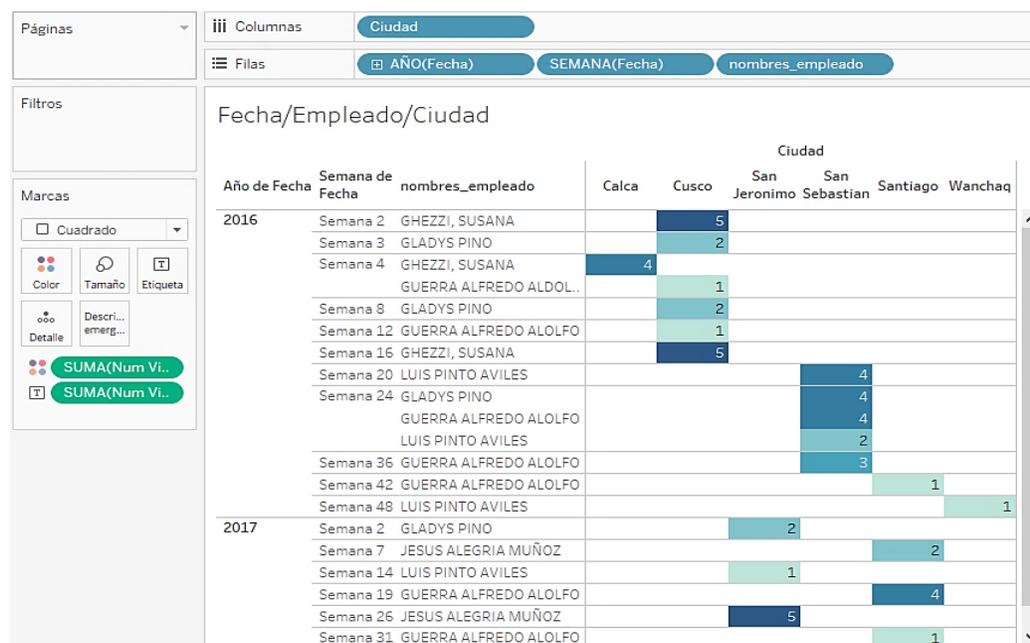
Fuente: propia (2022)

A continuación, se muestran algunos resultados gráficos de indicadores obtenidos y su respectiva validación

- La dimensión Fecha/Empleado/Ciudad muestra la combinación de 3 atributos, los datos de fecha, el nombre del empleado y la ciudad en la que se efectuó una visita, lo que se visualiza en las figuras 6 a modo de tabla y 7 como histograma.

Figura 6.

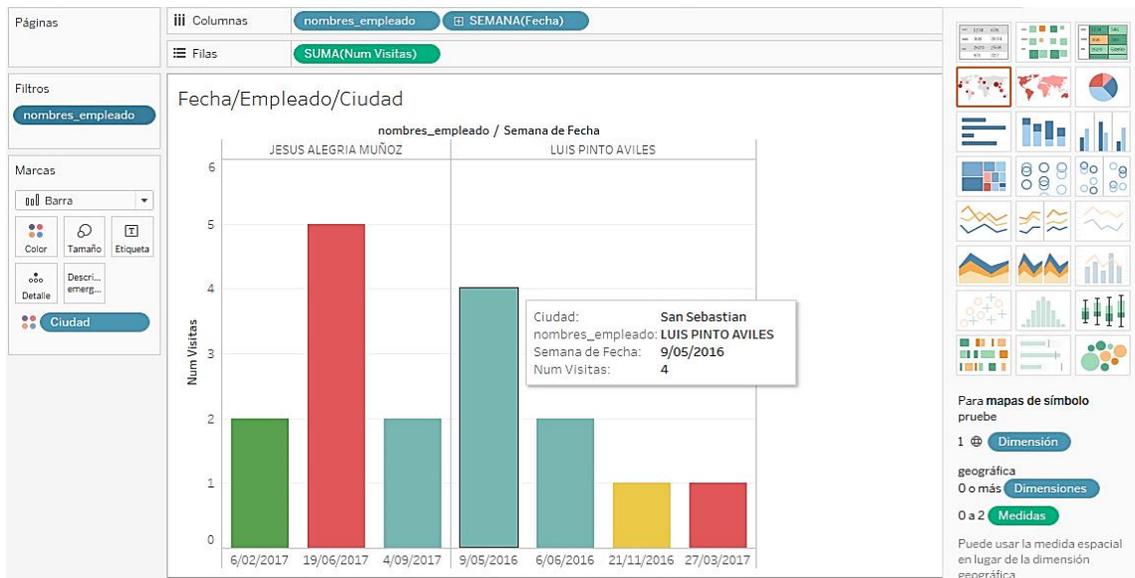
Dimensión Fecha/Empleado/Ciudad



Fuente: propia (2022)

Figura 7.

Histograma dimensión Fecha/Empleado/Ciudad



Fuente: propia (2022)

- La dimensión Fecha/TipoCliente/Ciudad en la figura 8 representada en un mapa y la figura 9 como gráfico circular.

Figura 8.

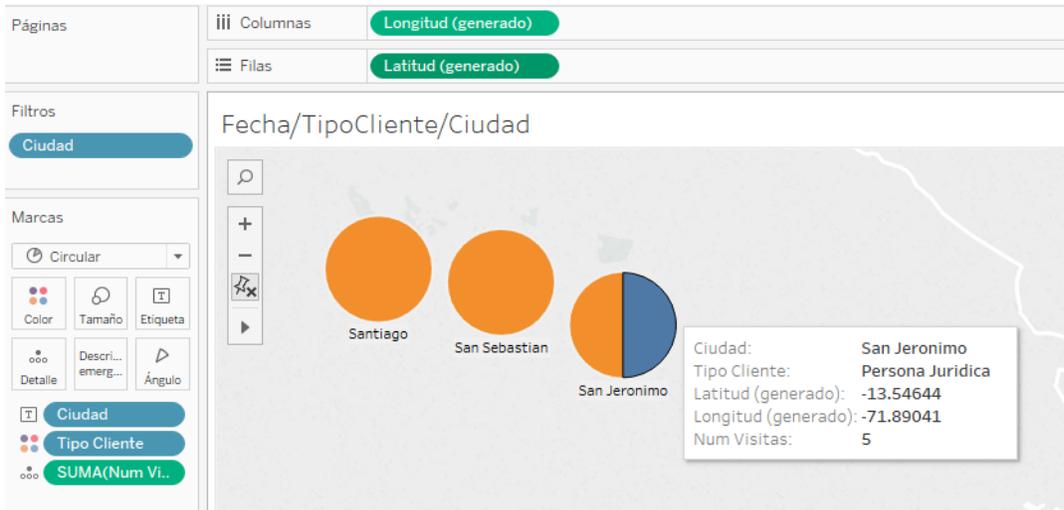
Dimensión Fecha/TipoCliente/Ciudad



Fuente: propia (2022)

Figura 9.

Dimensión Fecha/TipoCliente/Ciudad



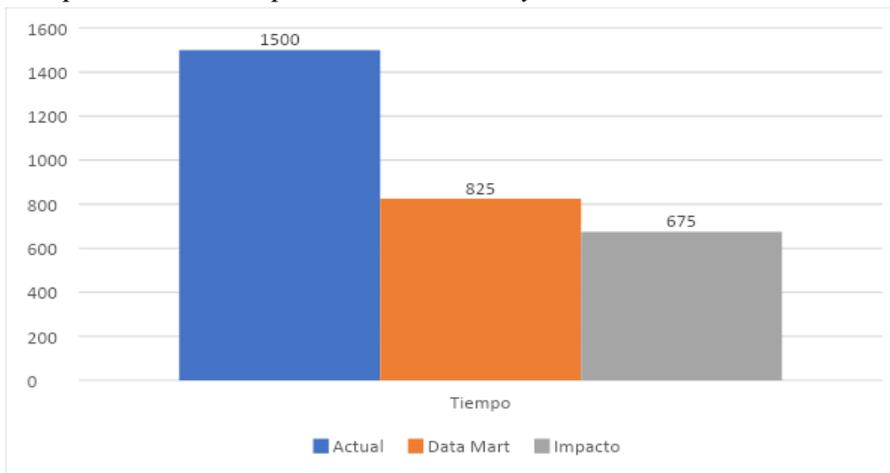
Fuente: propia (2022)

Los resultados demuestran que se puede mejorar la toma de decisiones con la implementación de un *Data Mart*, produciéndose entonces la reducción del tiempo en la elaboración de reportes; para este indicador se usó la ficha de recolección de datos y se registraron las muestras de tiempo en la búsqueda de información personalizada de la inmobiliaria antes de la implementación del *Data Mart*.

Para registrar los tiempos se tomó la investigación de referencia de Hamoud *et al.*(2021); de los resultados se obtienen el gráfico de la figura 10:

Figura 10.

Comparativa de tiempos entre lo Actual y el Data Mart



Fuente: propia (2022)

Lo que indica que existe una mejora casi el 50% de tiempo en la elaboración de reportes y búsqueda de información personalizada. Por tanto, el *Data Mart* es una propuesta de solución.

Discusión

Tara *et al.* (2021) proponen un innovador sistema *Data Mart* es escalable y prometedor como una solución generalizable en el ámbito de salud, el modelo de *Data Mart* propuesto es también una solución generalizable a empresas privadas del rubro inmobiliario.

Por otro lado Medina y Castillo-Rojas (2018) busca la obtención de indicadores de productividad académica concluyendo en el uso de 2 *Data Marts* que satisfacen las especificaciones de requerimientos definidos. La presente investigación sin embargo encuentra la satisfacción de los requerimientos de usuario en un modelo basado en un solo *Data Mart*.

Finalmente Dahr *et al.* (2022) usa un modelo de *Data Mart* basado en cuadros OLAP, KPI, y Data mining obteniendo una solución DSS de ventas que ayudan en la toma de decisiones y a desarrollar nuevos conocimientos y estrategias. El modelo de *Data Mart* construido es una herramienta que mejora la toma de decisiones y como efecto la productividad empresarial.

Conclusiones

El tiempo en la elaboración de reportes administrativos antes de la implementación del modelo fue de 1500 minutos (100%), en comparación con el *Data Mart* fue de 825 minutos (55.75%), lo cual determina una reducción notable de 675 minutos (44.25%) en la elaboración de reportes de ventas.

Monitorear la confiabilidad de las operaciones diarias en el sistema de ventas, ya que la base de datos transaccional sirve para alimentar el *Data Mart* y si la data no es confiable en términos de calidad, el modelo no mostrará información objetiva.

Realizar los *ETL* en un horario no laborable, porque si no es así, es posible que los usuarios perciban ralentización de la actividad del servidor debido a la carga de proceso. Si el *ETL* incluye una gran cantidad de datos, planifique la hora de inicio de forma que los usuarios resulten mínimamente afectados.

Se sugiere hacer un plan de replicación de base de datos para salvaguardar los datos transaccionales y por consiguiente la fuente de alimentación del *Data Mart*.

Mediante el modelo se propone una herramienta de tecnologías de la información útil, accesible y que mejora la toma de decisiones y productividad empresarial.

Como línea futura se sugiere que la generación del modelo de *Data Mart* sea una funcionalidad incorporada a las bases de datos transaccionales de las empresas.

Se sugiere también que de ser necesario se escale el modelo de acuerdo a la proyección de las empresas o se genere mas de un modelo de *Data Mart* que satisfaga los requerimientos de la empresa.

Referencias

- Arévalo, D., Nájera, S., & Piñero, E. (2018). La Influencia de la Implementación de las Tecnologías de Información en la Productividad de Empresas de Servicios. *Información tecnológica*, 29(6), 199-212. <https://ieeexplore.ieee.org/abstract/document/9172264>
- Cedeño T, F., & Townsend Valencia, J. (2021). Evaluación de la inversión en tic como factor de competitividad de las empresas Pymes del cantón Guayaquil. *Revista Universidad y Sociedad*, 13(2), 452-462. <https://rus.ucf.edu.cu/index.php/rus/article/view/1985>
- Chávez G, M. A. (2021). *Implementación de un Datamart para el soporte de la toma de decisiones en el área comercial de una empresa privada en la ciudad de Lima-2021*. [tesis de Pre grado, Universidad Tecnológica del Perú]. Repositorio institucional <https://repositorio.utp.edu.pe/handle/20.500.12867/5066>
- Choque, V. M. (2019). *Minería de datos aplicada a la identificación de factores de deserción universitaria en programas de pre grado* [tesis de maestría, Universidad Nacional de San Agustín]. Repositorio institucional <http://190.119.145.154/handle/UNSA/11015>
- Dahr, J., Hamoud, A., Najm, I., & Ahmed, M.. (2022). Implementing sales decision support system using Data mart based on OLAP, KPI, and Data mining approaches. *Journal of engineering science and technology*, 17(1), 0275-0293. https://www.researchgate.net/publication/358061092_implementing_sales_decision_support_system_using_data_mart_based_on_olap_kpi_and_data_mining_approaches
- Diouf, P., & Ndiaye, S. (2018, 12 de mayo). Variety of data in the ETL processes in the cloud: State of the art. [sesión de conferencia]. *IEEE International Conference on Innovative Research and Development (ICIRD)*, Bangkok, Thailand. <https://ieeexplore.ieee.org/abstract/document/8376308>
- Hamoud, A. K., Kamil Hussein, M., Alhilf, Z., & Hassan Sab, R. (2021). Implementing data-driven decision support system based on independent educational data mart. *International*

- Journal of Electrical and Computer Engineering (IJECE)*, 11(6), 5301.
https://www.researchgate.net/profile/Alaa-Khalaf-5/publication/353382539_Implementing_data-driven_decision_support_system_based_on_independent_educational_data_mart/links/60f97b821e95fe241a7ee1cf/Implementing-data-driven-decision-support-system-based-on-in
- Helmer, T., Lewis, A., McEver, M., Delacqua, F. (2021). Creating and implementing a COVID-19 recruitment Data Mart. *Journal of Biomedical Informatics*, 117, 103765.
<https://pubmed.ncbi.nlm.nih.gov/33798717/>
- Hernández Sampieri, R., Fernández Collado, C., & Baptista Lucio, P. (2014). *Metodología de la investigación*. McGraw-Hill
- Jaleel, R. & Abbas, T. (2020). Design and Implementation of Efficient Decision Support System Using Data Mart Architecture. [sesión de conferencia]. *In 2020 International Conference on Electrical, Communication, and Computer Engineering (ICECCE) (pp. 1-6)*. IEEE.
<https://ieeexplore.ieee.org/abstract/document/9179313>
- Kimball, R., & Ross, M. (2011). *The data warehouse toolkit: the complete guide to dimensional modeling*. John Wiley & Sons.
- Laudon, K. C. L. (2016). *Sistemas de información gerencial (14 ed.)*. Pearson.
- Medina, F., & Castillo-Rojas, W. (2018). Data Mart para obtención de indicadores de productividad académica en una universidad. *Ingeniare revista chilena de ingeniería*, 26, 88-101. https://www.scielo.cl/scielo.php?pid=S0718-33052018000500088&script=sci_arttext
- Mukherjee, R., & Kar, P. (2017, 07 de enero). A Comparative Review of Data Warehousing ETL Tools with New Trends and Industry Insight. *IEEE International Advance Computing Conference, IACC*) [sesión de conferencia]. Hyderabad, India <https://ieeexplore.ieee.org/abstract/document/7976926>
- Muñoz, G., Inda, A. D., González, M., & Álvarez, C. V. (2019). Las micro, pequeñas y medianas empresas, una estrategia de aplicación de tecnología para aumentar su competitividad. *Revista Espacios*, 40(20), p.2.
<http://www.revistaespacios.com/a19v40n20/19402002.html>
- Queiroz-Sousa, P. O., & Salgado, A. C. (2020, 02). A Review on OLAP Technologies Applied to Information Networks. *ACM Transactions on Knowledge Discovery from Data*, 14(1), 1-25. <https://dl.acm.org/doi/abs/10.1145/3370912>
- Sosa, V. D. (2017). *Modelo de Data Mart para Empresas Inmobiliarias. CASO: Inmobiliaria Alegría S.R.L.* grado [tesis de maestría, Instituto Científico y Tecnológico del Ejército ICTE]. Repositorio institucional

<http://repositorio.icte.ejercito.mil.pe/handle/123456789/156>

Talend Open Studio. (2021). Talend Open Studio: Open-source ETL and Free Data Integration. Talend (v 6.4.1) [software]. <https://www.talend.com/products/talend-open-studio/>

Tavera R., C. A., & Ríos Prado, A. (2021, 09). Business intelligence: business evolution after industry 4.0. *Sustainability 2021*, 13(18). <https://doi.org/10.3390/su131810026>

Utami, A., Pratama, B., & Widiyanto, S. (2020). Data Mart design in BKPP Bandung using from enterprise models to dimensional models method. *Jurnal Ilmu Pengetahuan dan Teknologi Komputer*, 5(1), 279-284.
<http://ejournal.nusamandiri.ac.id/index.php/jitk/article/view/1219>