

Tipo de artículo: Artículo original

Mercado de datos para el análisis de rentabilidad del cliente

Data Mart for customer profitability analysis

Maricela Pinargote-Ortega ^{1*}, <https://orcid.org/0000-0002-4018-9616>

Ivan Marcelo Quimi Ramirez ¹, <https://orcid.org/0009-0004-9078-0502>

Bryan Elías Domo Solorzano ¹, <https://orcid.org/0009-0000-6434-9664>

José Leonardo Loor Menéndez ¹, <https://orcid.org/0009-0004-4448-5414>

¹ Universidad Técnica de Manabí. Ecuador.

*Autor para la correspondencia. maricela.pinargote@utm.edu.ec

RESUMEN

El artículo analizó el mercado de datos en la empresa "La Fabril" de Ecuador, enfocándose en la rentabilidad del cliente a través del desarrollo de un almacén de datos. Se estableció como objetivo principal identificar y analizar patrones que contribuyan a optimizar la asignación de recursos. La metodología incluyó un enfoque analítico-sintético, modelado y la aplicación de la metodología Hefesto, que permitió la limpieza de datos, modelado dimensional en PostgreSQL y la creación de tableros interactivos en Power BI. Los resultados revelaron tendencias estacionales en los ingresos, destacando el "aceite de palma" como el producto más rentable y la identificación de grupos de clientes según su rentabilidad. Se observó una relación directa entre la frecuencia de compra y la lealtad del cliente, lo que subraya la necesidad de implementar estrategias de fidelización. Además, se identificaron áreas de alto costo y se propusieron



Esta obra está bajo una licencia *Creative Commons* de tipo **Atribución 4.0 Internacional**
(CC BY 4.0)

mejoras logísticas para optimizar la rentabilidad. En conclusión, la investigación demuestra que el uso efectivo de un almacén de datos puede transformar la toma de decisiones estratégicas y mejorar la rentabilidad en el contexto empresarial actual, y se planea ampliar el análisis a otros sectores y explorar técnicas avanzadas de análisis predictivo y aprendizaje automático para profundizar en la comprensión del comportamiento del cliente y optimizar aún más las estrategias empresariales.

Palabras clave: mercado de datos; rentabilidad del cliente; segmentación; análisis de patrones.

ABSTRACT

The paper analyzed the data mart in the company "La Fabril" in Ecuador, focusing on customer profitability through the development of a data warehouse. The main objective was to identify and analyze patterns that contribute to optimizing resource allocation. The methodology included an analytical-synthetic approach, modeling, and the application of the Hefesto methodology, which allowed for data cleaning, dimensional modeling in PostgreSQL, and the creation of interactive dashboards in Power BI. The results revealed seasonal trends in revenues, highlighting "palm oil" as the most profitable product and identifying customer groups based on their profitability. A direct relationship was observed between purchase frequency and customer loyalty, underscoring the need to implement loyalty strategies. Additionally, high-cost areas were identified, and logistical improvements were proposed to optimize profitability. In conclusion, the research demonstrates that the effective use of a data warehouse can transform strategic decision-making and enhance profitability in the current business context. There are plans to expand the analysis to other sectors and explore advanced techniques in predictive analysis and machine learning to deepen the understanding of customer behavior and further optimize business strategies.

Keywords: data mart; customer profitability; segmentation; pattern analysis.

Recibido: 14/09/2024

Aceptado: 08/12/2024

En línea: 01/01/2025



Esta obra está bajo una licencia *Creative Commons* de tipo **Atribución 4.0 Internacional**
(CC BY 4.0)

Introducción

En la era de los macrodatos, la capacidad de las empresas para analizar y comprender su información se ha convertido en un factor fundamental para el éxito. Los mercados de datos emergen como herramientas esenciales que permiten a las organizaciones consolidar y analizar información específica, facilitando así la toma de decisiones informadas. El avance tecnológico ha demostrado que la adopción de herramientas de inteligencia empresarial y la integración de datos en la nube pueden mejorar significativamente la toma de decisiones, proporcionando información de calidad en tiempo real (Gonçalves et al., 2023; Karabtsev et al., 2023).

La visualización de datos a través de tableros interactivos no solo facilita la evaluación de indicadores clave de rendimiento, sino que también permite una comprensión más profunda de los datos financieros y del comportamiento del cliente (Gonçalves et al., 2023; Munawar et al., 2024). En este contexto, los métodos adaptativos optimizan el aprendizaje en entornos diversos, ajustando el análisis de datos según las características específicas de cada conjunto, lo que se traduce en una mayor eficiencia y precisión en la toma de decisiones (Sal et al., 2024; Zhai et al., 2024).

El modelado dimensional del almacén de datos destaca la importancia de las tablas de hechos y la agrupación de atributos analíticos dentro de las dimensiones, estableciendo puntos clave de análisis fundamentales para la comprensión de la información. Esta estructura no solo optimiza la organización de los datos, sino que también facilita la identificación de patrones y tendencias (Daza et al., 2024).

La implementación de herramientas de inteligencia empresarial ha demostrado mejorar la capacidad analítica y optimizar la toma de decisiones estratégicas en las organizaciones (Gonçalves et al., 2023). Este enfoque se alinea con la revisión scientométrica sobre almacenes de datos, que resalta la importancia de la limpieza y transformación de datos en la creación de almacenes efectivos (Fan et al., 2022).

La investigación se llevó a cabo en la empresa "La Fabril", ubicada en Ecuador, que gestiona más de 1,152 productos en materia prima y 4,527 productos elaborados. Esta empresa abarca diversas líneas de productos, incluyendo cosméticos, chocolates, grasas, detergentes, plásticos, biodiésel y absorbentes. Durante el análisis, se identificaron limitaciones en la generación de reportes y gráficos estadísticos que presentaran información resumida, agregada y consolidada, así como un acceso deficiente a la información en tiempo



real. Por lo tanto, el objetivo de esta investigación fue desarrollar un almacén de datos para identificar y analizar patrones que contribuyan a optimizar la asignación de recursos.

En este sentido, se diseñó y construyó un almacén de datos con el fin de aumentar la eficiencia en el análisis de datos de productos y facilitar el proceso de toma de decisiones. Se procesaron datos desde el año 2022 hasta 2023 para analizar el comportamiento histórico de las ventas, lo que permitió segmentar a los clientes según su rentabilidad y los productos más adquiridos en diversas regiones y ciudades de Ecuador. Se inició con una limpieza exhaustiva de los datos, seguida de la creación de un diagrama en estrella en PostgreSQL y el desarrollo de un proceso de extracción, transformación y carga que garantiza la integridad y calidad de la información. Asimismo, se elaboraron tableros interactivos orientados a la evaluación de indicadores financieros y la segmentación de clientes, ofreciendo un análisis exhaustivo mediante visualizaciones y filtros que favorecen una comprensión más profunda de los datos

Este artículo está estructurado de la siguiente manera: en la Sección 2 se describe la metodología utilizada; la Sección 3 presenta los resultados obtenidos; y, finalmente, se exponen las conclusiones y las líneas de trabajo futuro.

Métodos o Metodología Computacional

En esta investigación, se emplearon el método analítico-sintético y la modelación. El enfoque analítico permitió establecer la metodología Hefesto para el desarrollo del almacén de datos, así como examinar detalladamente el conjunto de datos de la empresa “La Fabril”. Por otro lado, el enfoque sintético se aplicó para validar los indicadores de desempeño financiero y evaluar los resultados del sistema desarrollado.

La modelación se empleó en la estructuración y visualización del diseño del almacén de datos, lo que facilitó su implementación en PostgreSQL y su posterior integración con herramientas de visualización, favoreciendo así la toma de decisiones informadas.

Metodología de desarrollo

En el diseño del almacén de datos, se aplicó la metodología Hefesto, la cual se desarrolló en los siguientes pasos:



Paso 1: Análisis de requisitos

Se utilizó un conjunto de datos de la empresa “La Fabril”, que incluía información relacionada con alimentación, cosmética, producción de biocombustibles y datos de los clientes. Este conjunto abarcó aspectos demográficos, historial de transacciones y productos adquiridos. Para garantizar la protección de la información sensible, se anonimizó la identidad de los clientes, empleando etiquetas como "Cliente 1", "Cliente 2", entre otras. Este enfoque permitió asegurar la privacidad de los datos mientras se mantenía la integridad del análisis.

Paso 2: Definición de indicadores

Se establecieron diversos indicadores para el análisis financiero y de clientes, los cuales ofrecieron una visión detallada del desempeño financiero de la empresa y permitieron comprender su impacto en el negocio (ver Tabla 1). Además, se definieron indicadores específicos para el análisis de costos y clientes, facilitando la evaluación de las distintas ciudades, la proyección de la clientela y el establecimiento de líneas de tiempo relacionadas con los costos generados (ver Tabla 2).

Asimismo, se identificaron indicadores clave para determinar las regiones, sectores industriales, estados y productos con mayor movimiento en ganancias, lo que facilitó la toma de decisiones informadas y estratégicas en el contexto del mercado actual (ver Tabla 3).

Tabla 1. Indicadores financieros y clientes

Indicador	Descripción
Ingresos	Visualización de ingresos totales por cliente y por período.
Margen de ganancia	Gráfica de margen de ganancia por producto y cliente.
Número de transacciones	Gráfico de barras mostrando transacciones por cliente.
Segmentación de clientes	Segmentos de clientes basados en rentabilidad.
Frecuencia de compra	Indicador de frecuencia media de compra por cliente.

Tabla 2. Indicadores de costos y clientes

Indicador	Descripción
Costos asociados a productos y clientes	Visualización de costos asociados a productos y clientes.
Tasa de retención	Gráfico de retención de clientes por período.
Costos proyectados	Indicadores de la suma de costo variable y la suma de impuestos.
Ciudad costo de transporte	Árbol de mapas de las ciudades con mayor costo de transporte.
Lealtad del cliente	Indicador basado en la frecuencia y repetición de compras.



Tabla 3. Indicadores de ingresos y ventas.

Indicador	Descripción
Ingresos por región	Visualización geográfica de ingresos por región.
Productos más vendidos	Top de los productos más vendidos.
Proyección de ventas	Gráfico de línea que indica las ventas realizadas durante el año.
Margen de ganancia por industria	Indicador que muestra los mayores márgenes por industria.
Ganancia por estado	Muestras los estados que brindan más ganancias.

Paso 3: Modelado del almacén de datos en PostgreSQL

Se consideró la tabla de hechos, que se centró en la rentabilidad por cliente y contenía métricas clave como ingresos, costos y márgenes de ganancia. Asimismo, se definió la tabla de dimensiones, que incluyó variables como tiempo, producto, estado, industria y cliente, proporcionando diversas perspectivas para un análisis más completo (ver Tabla 4).

Una vez creadas las tablas en la base de datos PostgreSQL, se implementó el almacén de datos utilizando un esquema de estrella o de copo de nieve (ver Figura 1), lo que permitió organizar la información de manera que se simplificara la navegación y el análisis de los datos.

Tabla 4. Dimensiones y hechos resultantes del conjunto de datos de rentabilidad del cliente

Dimensiones/Hechos	Atributos	Descripción
dm_bu	bu, division, executive_id, id_bu	Unidad de Negocio, División dentro de la Unidad de Negocio, Identificador del Ejecutivo asignado, Identificador único de la Unidad de Negocio
dim_clientes	ciudad, ciudad_region, codigo_postal, estado, id_cliente, id_industria, nombres	Ciudad del cliente, Región de la ciudad, Código postal, Estado, Identificador único del cliente, Identificador de la industria, Nombres del cliente
dim_ejecutivo	id_ejecutivo, lmg, nombre	Identificador único del Ejecutivo, Imagen del Ejecutivo, Nombre del Ejecutivo
dim_escenario	id_escenario, escenario	Identificador único del escenario, Nombre del escenario
dim_fecha	anio, anio_periodo, fecha, mes, periodo, qtr, qtrid	Año, Año del período, Fecha específica, Mes, Número del período, Trimestre, Identificador del trimestre
dim_industria	id_industria, Image, industry	Identificador único de la industria, Imagen representativa, Nombre de la industria
dim_producto	id_producto, producto	Identificador único del producto, Nombre del producto
hrel_factura	anio_periodo, costo_material, costo_terceros, costo_viaje, gana_costo_viaje, ganancia,	Año del período, Costos de materiales, Costos a terceros, Costos de viaje, Ingresos de costos de viaje, Ganancia total, Identificador de la Unidad de Negocio, Identificador del cliente, Identificador del escenario, Identificador del



	id_bu, id_cliente, id_escenario, id_producto, impuesto, subscripcion	producto, Impuestos aplicados, Ingresos por suscripciones. Hecho que representa la relación entre los ingresos, costos y otros parámetros financieros relacionados con clientes, productos, escenarios y unidades de negocio.
--	--	---

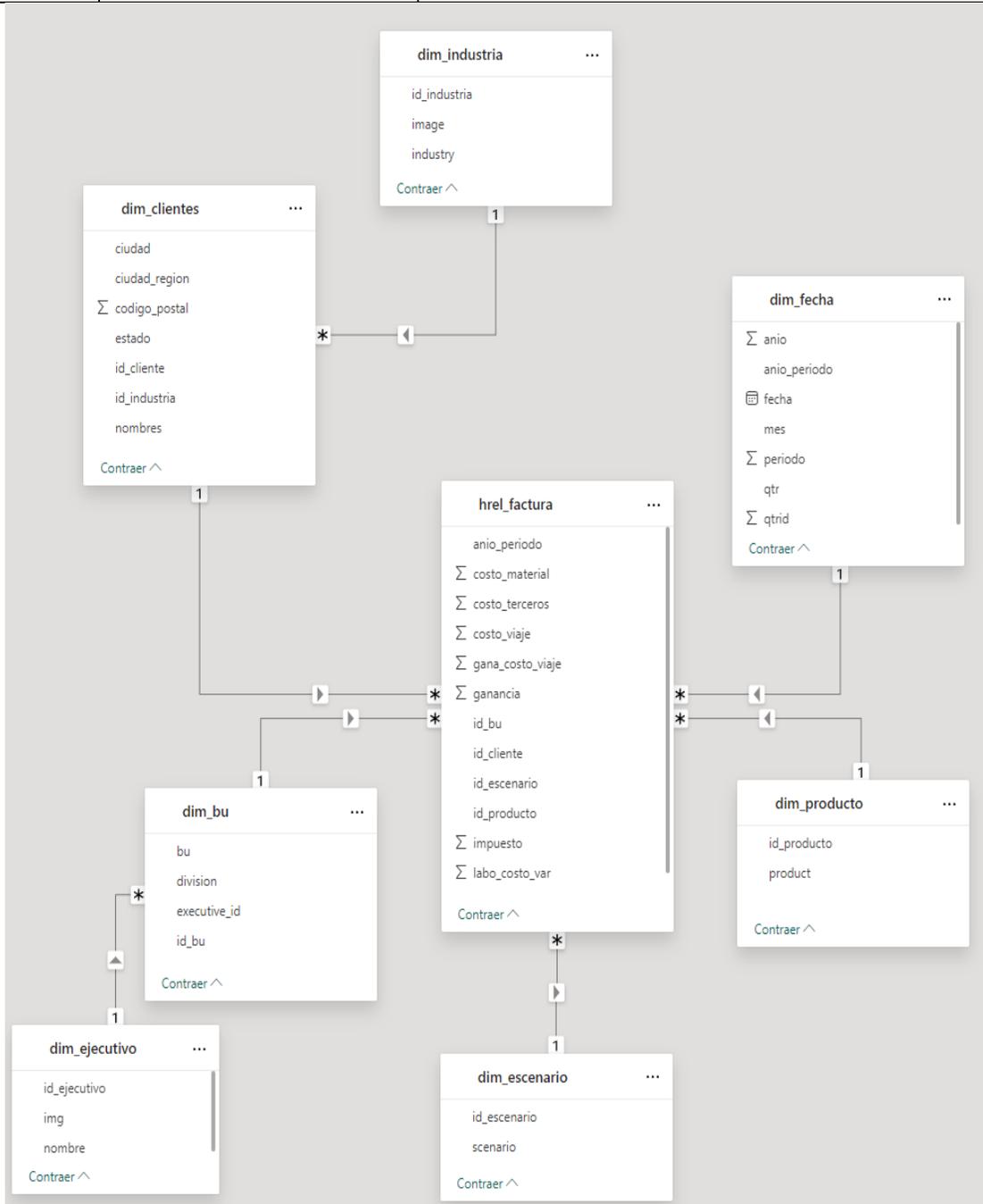


Fig 1. Almacén de datos en Power BI.



Paso 4: Proceso de extracción, transformación y carga

En la fase de extracción, se importó el archivo CSV a la base de datos SQL utilizando la herramienta DBeaver, posteriormente, en la etapa de transformación, se normalizaron los datos con el fin de eliminar inconsistencias y estandarizar formatos. Las tablas SQL se estructuraron para representar diferentes entidades del negocio, como clientes, productos y ventas, lo que facilitó una mejor organización de la información.

Una vez completada la transformación, se definieron claves primarias y foráneas en DBeaver para establecer relaciones entre las tablas, creando así un esquema relacional eficiente. Las tablas se cargaron en el almacén de datos, donde quedaron disponibles para su análisis e informes. Finalmente, se estableció una conexión entre la herramienta de visualización y el almacén de datos, lo que permitió la creación de informes y paneles interactivos, optimizando así la toma de decisiones (ver Figura 2).

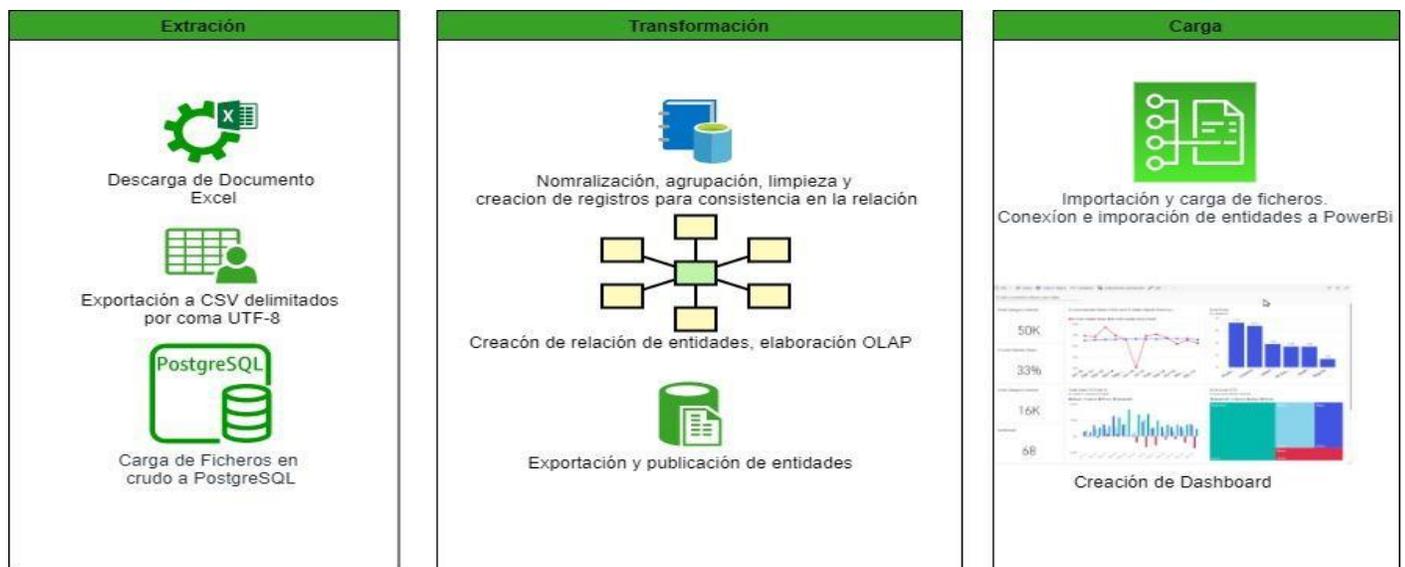


Fig 2. Proceso de extracción, transformación y carga.

Paso 5: Visualización en Power BI

Se desarrollaron paneles interactivos que mostraron los indicadores clave de rentabilidad por cliente (ver Tabla 5), lo que facilitó un análisis detallado de la segmentación de clientes y las tendencias de compra. Estos indicadores se visualizaron mediante gráficos, lo que permitió una comprensión más clara de la



información y proporcionó una herramienta eficaz para la toma de decisiones, ayudando a los analistas a identificar patrones y oportunidades de mejora en la gestión de clientes.

Tabla 5. Indicadores y perspectivas

Indicador	Perspectiva	Descripción
Ingresos	Financiera	Monto total de ventas realizadas por cliente.
Costos	Financiera	Costos asociados a los productos adquiridos por cliente.
Margen de ganancia	Financiera	Diferencia entre ingresos y costos.
Número de transacciones	Operativa	Cantidad de compras realizadas por cada cliente.
Frecuencia de compra	Relación con cliente	Promedio de tiempo entre compras consecutivas.
Segmentación de clientes	Marketing	Clasificación de clientes en diferentes segmentos según rentabilidad.
Satisfacción del cliente	Relación con cliente	Medición basada en encuestas y retroalimentación de clientes.
Tasa de retención	Relación con cliente	Porcentaje de clientes que continúan comprando en un período dado.
Nuevos clientes	Crecimiento	Número de clientes nuevos adquiridos en un período dado.
Ventas proyectadas	Futuro	Proyección de ventas basada en tendencias históricas.

Resultados y discusión

Los hallazgos de esta investigación ofrecen un análisis integral de los indicadores financieros y de clientes en “La Fabril”, revelando patrones significativos en los ingresos y márgenes de ganancia a lo largo del tiempo. Se identificaron tendencias estacionales y variaciones en la rentabilidad de productos, así como oportunidades clave para optimizar la asignación de recursos. Además, el estudio evaluó los costos y la retención de clientes, proporcionando información valiosa para la formulación de estrategias efectivas. A continuación, se presentan los resultados más relevantes.

Análisis de indicadores financieros y clientes

Los ingresos y márgenes de ganancia asociados a cada cliente se desglosaron por fecha (ver Figura 3a), permitiendo un análisis exhaustivo de su evolución a lo largo del tiempo. Este desglose facilitó la identificación de tendencias y patrones estacionales, destacando un aumento significativo en los ingresos durante el tercer trimestre del año, lo que sugiere un comportamiento estacional en las compras. Asimismo, se observó una variación notable en los márgenes de ganancia entre diferentes clientes, lo que indica que ciertos segmentos resultaron ser más rentables en períodos específicos, especialmente al inicio del tercer trimestre.



El margen de ganancia segmentado por producto reveló que el "aceite de palma" fue el producto más rentable, destacándose en varios clientes, con márgenes significativos en productos como el "aceite crudo de soya" (ver Figura 3b). Sin embargo, se notó una tendencia descendente en los márgenes conforme avanzaron los clientes, indicando una disminución en la rentabilidad de ciertos productos. Aunque el "aceite de palma" mantuvo su dominio, otros productos como el "aceite girasol" generaron márgenes más bajos, lo que resaltó la necesidad de centrar las estrategias de venta en los productos más rentables para optimizar los ingresos.

Los clientes se agruparon en función de sus ingresos y márgenes de ganancia (ver Figura 3c). El grupo "altamente rentable" predominó por su sólido desempeño en ambos aspectos, mientras que el grupo "moderadamente rentable" mostró un rendimiento inferior. Por otro lado, el grupo "menos rentable" presentó ingresos y márgenes negativos, lo que indicó la necesidad de mejorar su rentabilidad.

El análisis de las transacciones ayudó a identificar a los clientes frecuentes y a diseñar estrategias de fidelización (ver Figura 3d). A partir de 2022, algunos clientes realizaron múltiples compras, mientras que otros destacaron por su mayor lealtad. Esto subrayó la importancia de implementar estrategias como programas de recompensas, descuentos exclusivos y ofertas personalizadas para fortalecer la relación con los clientes frecuentes y atraer nuevos.

El recuento de transacciones realizadas por cada cliente proporcionó una visión cuantitativa de su actividad (ver Figura 3e). El análisis mostró que el Cliente 2 lideró con 103 transacciones, seguido por el Cliente 1 con 64, el Cliente 7 con 54 y el Cliente 19 con 44. Estos datos reflejaron la variabilidad en la frecuencia de compras entre los clientes, lo que pudo indicar diferentes niveles de compromiso o patrones de consumo.

Análisis de costos y clientes

El análisis de costos entre productos y segmentos de clientes permitió identificar áreas de mayor gasto, facilitando la optimización de precios y la reducción de costos (ver Figura 4a). En particular, se registró un alto consumo del producto "aceite crudo de soya com rat" por parte del Cliente 1, Cliente 2 y Cliente 11, lo cual destacó áreas clave para la gestión de costos.

La retención de clientes, analizada a través de un filtro de fecha, permitió evaluar tanto la lealtad del cliente como la efectividad de las estrategias de fidelización (ver Figura 4b). Se percibió que clientes como el Cliente 16 y el Cliente 15 presentaron los niveles de suscripción más altos, lo que indica una relación sólida



y continua con la empresa. Estos resultados sugieren la conveniencia de centrar los esfuerzos de fidelización en aquellos clientes con menor suscripción, como el Cliente 7, para maximizar la retención general.

La evolución proyectada de los costos variables e impuestos facilitó la identificación de tendencias y la planificación del presupuesto futuro (ver Figura 4c). Se distinguió una disminución en los costos laborales hasta febrero de 2023, mientras que los impuestos se mantuvieron estables, con picos destacados en octubre de 2022 y mayo de 2023.

Las ciudades con mayores costos de transporte fueron "Quito" y "Guayaquil", lo que permitió enfocar la optimización de rutas logísticas para reducir gastos (ver Figura 4d).

La rentabilidad, en función de la frecuencia y repetición de compras, proporcionó una visión clara del compromiso de los clientes con la marca (ver Figura 4e). Es importante señalar que la mayor rentabilidad provino del segmento "altamente rentable", que representó el 87,24% de los costos materiales, lo que subrayó la necesidad de centrarse en este grupo para maximizar los ingresos.

Análisis de ingresos y ventas

Los ingresos generados en cada área geográfica facilitaron la identificación de regiones con mayor rendimiento y potencial de crecimiento, destacando una mayor cantidad de ganancias en la región "sierra" (ver Figura 5a). El análisis de los productos más vendidos permitió ajustar las estrategias de ventas según las preferencias del mercado, destacando en este caso el "aceite de palma", el "aceite de soya", el "aceite de canola crudo" y el "aceite de girasol" (ver Figura 5b). Las ventas anuales permitieron identificar tendencias clave y prever el rendimiento futuro, con picos de demanda en meses como septiembre, octubre y noviembre de 2022 y 2023, lo que ayudó a "La Fabril" a ajustar su producción y distribución según las fluctuaciones del mercado ecuatoriano (ver Figura 5c).

Además, identificar los márgenes de ganancia por industria permitió a "La Fabril" enfocar sus esfuerzos en sectores más rentables, como los servicios, destacando productos energéticos y materias primas especializadas (ver Figura 5d). Las provincias que generaron mayores ganancias fueron "Pichincha", "Guayas" y "Manabí" (ver Figura 5e).



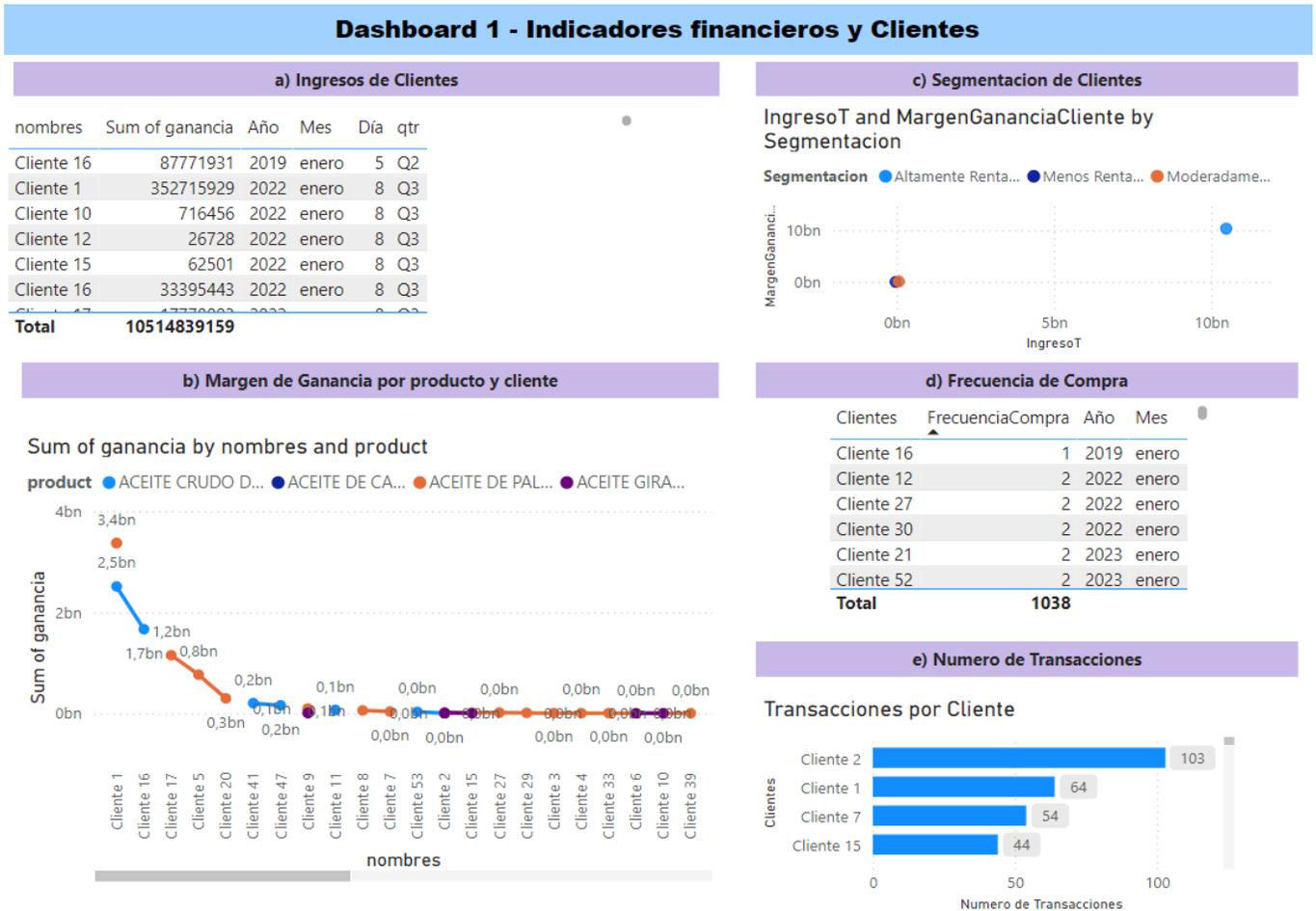


Fig 3. Panel de indicadores financieros y clientes



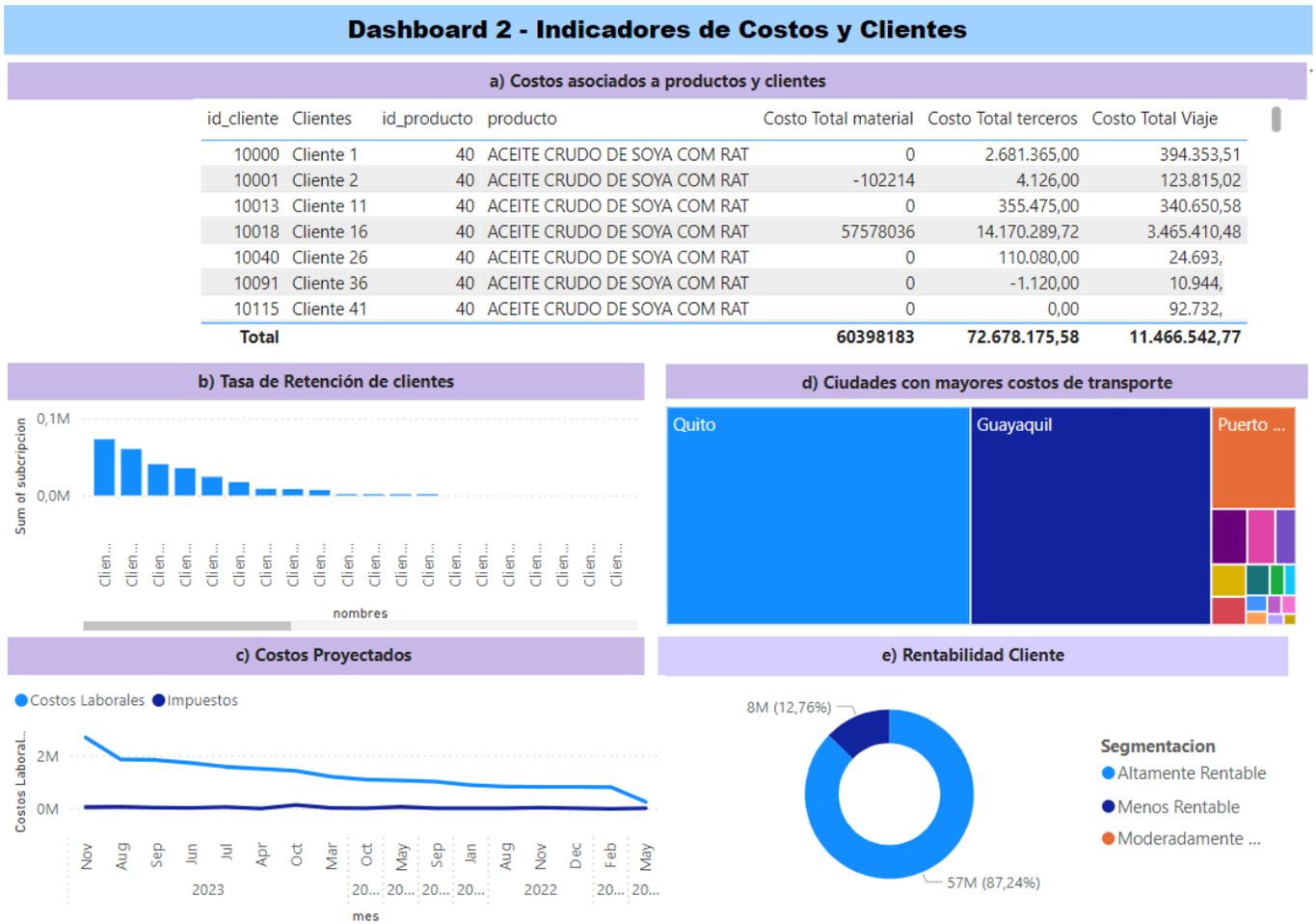


Fig 4. Panel de costos y clientes



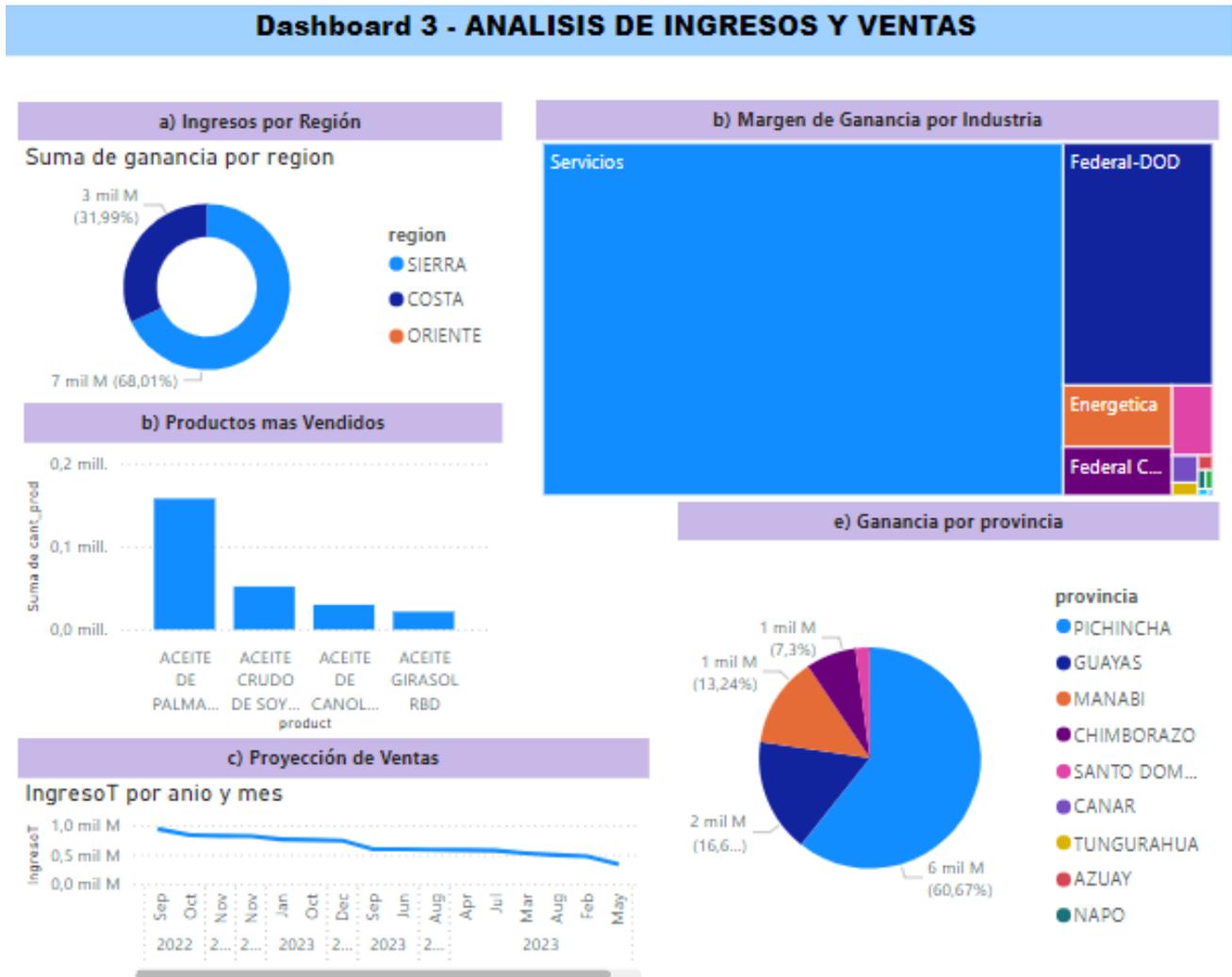


Fig 5. Panel de análisis de ingresos y ventas

La implementación del almacén de datos en "La Fabril" permitió establecer relaciones lógicas entre las diversas dimensiones de los datos, optimizando tanto el rendimiento del sistema como la representación visual de la información. Este proceso se asemejó a la construcción de un cubo multidimensional, facilitando la transformación de datos operacionales en información estratégica. Esta transformación fue fundamental para asegurar un análisis confiable y preciso, garantizando así una base de datos de alta calidad para la toma de decisiones. Este enfoque se alinea con las prácticas adoptadas por otros autores que



aplicaron procesos de extracción, transformación y carga de datos para asegurar la calidad de la información, mejorando la precisión y consistencia de los datos para el análisis (Wang et al., 2024).

Los conocimientos obtenidos de este análisis permitieron ajustar las estrategias de marketing y ventas, enfocándose en los clientes de alta rentabilidad y optimizando costos. Estos resultados se alinean con investigaciones previas que demuestran cómo la mejora en las campañas de marketing y la asignación eficiente de recursos operativos pueden impactar positivamente en la rentabilidad (Liu et al., 2024). Además, el análisis de datos de clientes para la segmentación y personalización de servicios resultó ser fundamental para la mejora del marketing y la retención de clientes, corroborando hallazgos de estudios recientes que destacan la importancia de estas prácticas en la fidelización del cliente (Antal et al., 2022).

La capacidad de aplicar filtros temporales fue fundamental para identificar tendencias, destacando la relevancia de analizar datos a lo largo del tiempo en el ámbito de la planificación financiera. Este enfoque se complementa con la inclusión de etiquetas de valor y una leyenda que distinguió los productos, lo que facilitó una interpretación más efectiva de la información y resultó esencial para detectar oportunidades de optimización en la estrategia de productos. Este hallazgo se alinea con el estudio de Purnamasari et al. (2019), que identificó patrones de comportamiento y segmentó a los clientes, lo que a su vez permitió una personalización más efectiva de las estrategias de marketing y ventas.

El análisis de la frecuencia de compra fue clave para mejorar la retención de clientes, ya que facilitó la comprensión de la relación entre la frecuencia de compras y los ingresos generados. Estos resultados se alinean con investigaciones sobre mercados de datos implementados, que buscan garantizar una integración eficiente de la información, optimizando así recursos y estrategias (Rosado et al., 2017; Oluwatoyin et al., 2023).

Se optimizó la segmentación de clientes al analizar la relación entre la cantidad de productos comprados y las ganancias, lo que facilitó una gestión más eficiente de los productos y reveló patrones de compra esenciales para incrementar la rentabilidad. Estos hallazgos son coherentes con investigaciones que identifican los factores que influyen en la deserción de clientes, lo que resulta vital para desarrollar estrategias efectivas de retención. Además, este análisis permitió visualizar y mejorar la lealtad en diversas industrias, tal como se señala en el estudio de Oluwatoyin et al. (2023), donde se enfatiza la relevancia de entender el comportamiento del cliente para fomentar relaciones más sólidas y duraderas.



El análisis del número de transacciones reveló una relación directa con el valor de vida del cliente, ya que una mayor frecuencia de compras no solo incrementó los ingresos, sino que también mejoró la retención. Los clientes que compran con mayor frecuencia tienden a desarrollar una lealtad más sólida hacia la marca, lo que se traduce en un aumento en el valor de vida del cliente. Además, el análisis en múltiples niveles permitió una comprensión más profunda del comportamiento del cliente y su impacto en las estrategias de ventas.

Por otro lado, el estudio de estos ingresos proporcionó información valiosa sobre las áreas donde se podían implementar estrategias de desarrollo. Estos resultados son coherentes con investigaciones que destacan el impacto de los mercados de datos, evidenciando que una infraestructura adecuada puede potenciar aún más estos ingresos. Esto subraya la importancia de evaluar continuamente el rendimiento económico en diferentes regiones (Kvíčala et al., 2024; Rephann, 2022). Además, este tipo de análisis, junto con las metodologías de integración y visualización de datos aplicadas en mercados de datos, facilitó una toma de decisiones más informada al procesar correctamente la información almacenada en un modelo dimensional de datos, como se menciona en el trabajo de Hernández et al. (2017).

En este contexto, Piñera et al. (2021) destacan que el uso de metodologías de almacén de datos y procesos de extracción, transformación y carga asegura una integración eficiente de la información, permitiendo un análisis preciso y estructurado de los costos. Un análisis de rentabilidad del cliente, combinado con la disponibilidad de datos consistentes, facilita una toma de decisiones más informada y eficiente (Rodríguez et al., 2017). Además, la implementación de sistemas automatizados y bases de datos, como los utilizados por Rodríguez et al. (2016) para la planificación del presupuesto, ha demostrado ser una estrategia eficaz para mejorar la precisión y eficiencia en la gestión de los costos proyectados. La adopción de métodos de automatización a través de aplicaciones integradas permite obtener análisis casi en tiempo real, mejorando la capacidad de adaptación a las fluctuaciones del mercado y facilitando una planificación más dinámica.

Este enfoque respalda las conclusiones de Gupta et al. (2021), Mahfuza et al. (2022) y Sousa et al. (2024), quienes enfatizan la importancia de la toma de decisiones estratégicas en sectores clave basadas en datos analíticos, subrayando así la relevancia de una infraestructura de datos robusta para el éxito organizacional.



Conclusiones

La segmentación de clientes y el análisis de patrones de compra han proporcionado valiosos conocimientos que pueden guiar las estrategias de marketing y ventas. Al identificar segmentos de alto valor y optimizar la asignación de recursos, se ha constatado que las ciudades de Quito y Guayaquil son las que registran la mayor cantidad de adquisiciones de aceites. Esta tendencia subraya la importancia de estos mercados urbanos, donde la demanda es notablemente alta. No obstante, es fundamental considerar que el costo de transporte hacia estas ciudades también es elevado, lo que exige la implementación de estrategias logísticas eficientes para mantener la competitividad y rentabilidad. La optimización de rutas y la negociación de tarifas de transporte se presentan como elementos clave para mitigar estos costos adicionales y maximizar los márgenes de ganancia.

Una de las limitaciones del estudio fue el tamaño reducido de la muestra. Para futuras investigaciones, se planea ampliar esta muestra y explorar el uso de técnicas avanzadas de análisis predictivo y aprendizaje automático, lo que contribuirá a obtener una comprensión más profunda de los patrones de comportamiento del cliente y a mejorar la efectividad de las estrategias comerciales, facilitando así una toma de decisiones aún más informada y precisa.

Agradecimientos

Se agradece a "La Fabril" por su valiosa colaboración en facilitar datos y apoyo para el desarrollo de esta investigación, así como a la Universidad Técnica de Manabí por el respaldo académico y logístico recibido. A todas las personas que contribuyeron directa o indirectamente con su tiempo y conocimientos, se les expresa también un sincero agradecimiento.

Referencias

Antal, R.; Marasová, D.; Hájiček, R.; Klapko, P.; Mitrik, V. Implementation of business intelligence system to analyze the data for mining mechanization – case study. *Acta Montanistica Slovaca*, 2022, 27(3): 744–753. DOI 10.46544/AMS.v27i3.14.



- Daza, A., de la paz tito, E. M., Cajas, Y. R. A. Impact of business intelligence on incident management in the control center of a security company. *Bulletin of Electrical Engineering and Informatics*, 2024, 13(1): 422–435. DOI 10.11591/eei.v13i1.5819.
- Fan, C., Zeng, L., Huang, K. Research Landscape, Emerging Trends and New Developments in Data Warehouse: A Scientometric Analysis (1985 - 2021). 2022 12th International Workshop on Computer Science and Engineering, 2022: 126–136. DOI 10.18178/wcse.2022.06.019.
- Gonçalves, C. T., Gonçalves, M. J. A., Campante, M. I. Developing Integrated Performance Dashboards Visualisations Using Power BI as a Platform. *Information (Switzerland)*, 2023, 14(11): 614. DOI 10.3390/info14110614.
- Gupta, P., Antal, R., Daza, A. Implementation of demand forecasting – A comparative approach. *Journal of Physics: Conference Series*, 2021, 1714(1): 012003. DOI 10.1088/1742-6596/1714/1/012003.
- Hernández, Y. P., Piñera, I. G. Mercado de Datos para el área de Recursos Humanos. *Serie Científica de la Universidad de las Ciencias Informáticas*, 2017, 10(5): 56–71. Recuperado de: <http://publicaciones.uci.cu>
- Karabtsev, S., Gonçalves, C. T., Piñera, I. G. Building data marts to analyze university faculty activities using Power BI. *E3S Web of Conferences*, 2023, 419: 02014. DOI 10.1051/e3sconf/202341902014.
- Kvíčala, D., Králová, M., Suchánek, P. The impact of online purchase behaviour on customer lifetime value. *Journal of Marketing Analytics*, 2024. DOI 10.1057/s41270-024-00328-9.
- Liu, Y. T., Lai, H. F. Constructing an Automated Quality Control Indicator Dashboard by Linking Cloud-Based Forms with Power BI. *Studies in health technology and informatics*, 2024, 315: 673–674. DOI 10.3233/SHTI240273.
- Munawar, G., Arsyad, Z., Hodijah, A. The development of business intelligence dashboard to measure the KPI of the study program (Case study: LAM Infokom instrument). *AIP Conference Proceedings*, 2024, 2952(1): 120004. DOI 10.1063/5.0211908.
- Oluwatoyin, A. M., Gupta, P., Daza, A. Customer Churn Prediction in Banking Industry Using Power BI. *Lecture Notes in Networks and Systems*, 2023, 421: 767–774. DOI 10.1007/978-981-19-1142-2_60.



- Piñera, I. G., Pérez, A. A., Vaillant, A. E. D. Mercado de Datos para el proceso de perfeccionamiento educacional en Cuba. *Serie Científica de la Universidad de las Ciencias Informáticas*, 2021, 14(3): 58–72. Recuperado de: <http://publicaciones.uci.cu>
- Purnamasari, A. M., Gupta, P., Liu, Y. T. Business Intelligent in an E-Commerce Industry. *IOP Conference Series: Materials Science and Engineering*, 2019, 598(1). DOI 10.1088/1757-899X/598/1/012085.
- Rephann, T. J. State Economic Impacts of Industrial Park Development: Evidence from Virginia’s Tobacco Region Megasite Program. *Review of Regional Studies*, 2022, 52(2): 249–273. DOI 10.52324/001c.37971.
- Rodríguez, F. D., Chavez, A. J. V. Mercado de Datos para la Dirección de Mercadotecnia Operativa de la División de Servicios Fijos de ETECSA. *Serie Científica de la Universidad de las Ciencias Informáticas*, 2016, 9(9): 70–83. Recuperado de: <http://publicaciones.uci.cu>
- Rodríguez, M. A. F., Morales, C. J. T., Grass, O. Y. R. Mercado de datos de energías renovables. *Serie Científica de la Universidad de las Ciencias Informáticas*, 2017, 10(7): 39–48. Recuperado de: <http://publicaciones.uci.cu>
- Rosado, M. R., Torres, J. A., Legrá, E. E. C. Mercado de Datos para el Departamento de Mercadotecnia perteneciente a la Dirección de Transferencia de Tecnología y Conocimiento. *Serie Científica de la Universidad de las Ciencias Informáticas*, 2017, 10(5): 25–42. Recuperado de: <http://publicaciones.uci.cu>
- Sal, B., Liu, Y. T., Gupta, P. Domain-specific languages for the automated generation of datasets for industry 4.0 applications. *Journal of Industrial Information Integration*, 2024, 41: 100657. DOI 10.1016/J.JII.2024.100657.
- Wang, J., Liu, B. Design of ETL Tool for Structured Data Based on Data Warehouse. *ACM International Conference Proceeding Series*, 2020. DOI 10.1145/3424978.3425101.
- Zhai, R., Wang, J., Liu, B. Adaptive client selection and model aggregation for heterogeneous federated learning. *Journal of Computer and System Sciences*, 2024.

Conflicto de interés

Los autores autorizan la distribución y uso de su artículo.



Esta obra está bajo una licencia *Creative Commons* de tipo **Atribución 4.0 Internacional** (CC BY 4.0)

Contribuciones de los autores

1. Conceptualización: José Leonardo Loor Menéndez
2. Curación de datos: Bryan Elías Domo Solorzano
3. Análisis formal: José Leonardo Loor Menéndez
4. Investigación: Ivan Marcelo Quimi Ramirez
5. Metodología: Ivan Marcelo Quimi Ramirez, José Leonardo Loor Menéndez, Maricela Pinargote Ortega, Bryan Elías Domo Solorzano
6. Administración del proyecto: Maricela Pinargote Ortega
7. Recursos: Bryan Elías Domo Solorzano
8. Software: José Leonardo Loor Menéndez
9. Supervisión: Maricela Pinargote Ortega
10. Validación: Maricela Pinargote Ortega
11. Visualización: Ivan Marcelo Quimi Ramirez
12. Redacción – borrador original: Ivan Marcelo Quimi Ramirez, José Leonardo Loor Menéndez, Maricela Pinargote Ortega, Bryan Elías Domo Solorzano
13. Redacción – revisión y edición: Ivan Marcelo Quimi Ramirez, José Leonardo Loor Menéndez, Maricela Pinargote Ortega, Bryan Elías Domo Solorzano

Financiación

La investigación no requirió fuente de financiamiento.

