



Escuela de Posgrados

Caso K: Analítica de Datos y Estrategias de Transformación Digital

Máster en Gerencia de Datos y Negocios

Camila Fernanda Rojas Avila

Ana Carolina Armijos

Cuenca, Ecuador 2026

DEDICATORIA

Dedico este trabajo, a mi papá, Juan Pablo, quien ha sido mi mayor ejemplo de esfuerzo, integridad y perseverancia. Gracias por creer siempre en mí, por impulsarme a alcanzar mis metas y por brindarme tu apoyo incondicional en cada etapa de mi vida. Eres el hombre que más admiro y una de mis mayores inspiraciones.

AGRADECIMIENTO

A mi mamá, Fernanda, por su amor infinito, sus palabras de aliento y por enseñarme, con su ejemplo, el valor de la dedicación y la fortaleza. Gracias por acompañarme siempre y celebrar cada uno de mis logros como si fueran tuyos.

Y a mi segundo papá, Carlos, por su cariño, confianza y apoyo constante. Gracias por estar presente en este camino, por motivarme a seguir adelante y por formar parte fundamental de mi vida.

Este logro también les pertenece a ustedes, porque sin su amor, su guía y su apoyo incondicional, este sueño no habría sido posible.

Resumen ejecutivo

La presente tesis integra cinco áreas de la Maestría en Gerencia de Datos y Negocios para demostrar cómo el uso estratégico de los datos fortalece la toma de decisiones, la eficiencia operativa y la competitividad empresarial. El trabajo aborda soluciones de Big Data, procesamiento y visualización de datos, gestión de proyectos, análisis financiero y transformación digital, aplicados a casos empresariales y sociales relevantes. En primer lugar, se propone una solución Big Data para Mercado Libre, orientada a optimizar la operación comercial y evaluar su expansión hacia Ecuador mediante una arquitectura lakehouse cloud-first, modelos analíticos, gobierno de datos y tableros gerenciales. En segundo lugar, se analiza la relación entre turismo y desigualdad de ingresos en Ecuador, evidenciando que la visualización de datos permite identificar patrones útiles para decisiones públicas y empresariales. Además, se desarrolla un proyecto de analítica predictiva para campañas bancarias, con el fin de mejorar la segmentación de clientes y aumentar la efectividad comercial. Asimismo, la tesis incorpora un análisis financiero comparativo de Indurama, Mabe y Fibroacero, destacando diferencias en ventas, rentabilidad, liquidez y sostenibilidad durante el periodo 2021-2024. Finalmente, se plantea una estrategia de transformación digital basada en la actualización de SAP S/4HANA en una empresa manufacturera, con el propósito de reducir riesgos operativos, asegurar continuidad del negocio y fortalecer la competitividad. En conjunto, el documento concluye que la gestión de datos no debe entenderse sólo como un recurso tecnológico, sino como una capacidad gerencial clave para generar valor, anticipar riesgos, optimizar recursos y construir ventajas competitivas sostenibles.

Palabras clave: Big Data, inteligencia de negocios, transformación digital, visualización de datos, toma de decisiones

Abstract

This thesis integrates five areas of the Master's Program in Data and Business Management to demonstrate how the strategic use of data strengthens decision-making, operational efficiency, and business competitiveness. The study addresses Big Data solutions, data processing and visualization, project management, financial analysis, and digital transformation through applied business and social cases. First, the thesis proposes a Big Data solution for Mercado Libre, focused on optimizing commercial operations and evaluating its expansion into Ecuador through a cloud-first lakehouse architecture, analytical models, data governance, and executive dashboards. Second, it analyzes the relationship between tourism and income inequality in Ecuador, showing how data visualization helps identify useful patterns for public and business decision-making. In addition, the study develops a predictive analytics project for banking marketing campaigns to improve customer segmentation and increase commercial effectiveness. The thesis also includes a comparative financial analysis of Indurama, Mabe, and Fibroacero, highlighting differences in sales, profitability, liquidity, and financial sustainability during the 2021-2024 period. Finally, it presents a digital transformation strategy based on the upgrade of SAP S/4HANA in a manufacturing company, aimed at reducing operational risks, ensuring business continuity, and strengthening competitiveness. Overall, the document concludes that data management should not be understood only as a technological resource, but as a key managerial capability to generate value, anticipate risks, optimize resources, and build sustainable competitive advantages.

Keywords: Big Data, business intelligence, digital transformation, data visualization, decision making



Índice

Índice de Tablas	xi
Índice de Figuras	xii
Capítulo 1: Big Data Solutions	1
1.1 Introducción del capítulo	1
1.2 Definición del problema desde datos	1
1.3 Objetivo general del capítulo	3
1.4 Objetivos específicos	3
1.5 Alcance de la solución Big Data	3
1.6 Soporte teórico: Big Data como capacidad estratégica	4
1.7 Arquitectura propuesta: lakehouse cloud-first con streaming	5
1.7.1 Capa de ingesta	5
1.7.2 Capa de almacenamiento	6
1.7.3 Capa de procesamiento	6
1.7.4 Capa analítica y machine learning	6
1.7.5 Capa de visualización y consumo	7
1.8 Fuentes de datos requeridas	7
1.8.1 Fuentes internas	7
1.8.2 Fuentes externas	8
1.9 Estrategia de integración de datos	9
1.9.1 Identificadores maestros	9
1.9.2 Contratos de datos	9
1.9.3 Calidad de datos	10
1.9.4 Seguridad y privacidad	10
1.9.5 Observabilidad	10
1.10 Casos de uso analíticos prioritarios	10
1.10.1 Recomendaciones y personalización	10
1.10.2 Pricing dinámico y promociones inteligentes	10
1.10.3 Predicción de demanda y abastecimiento	11
1.10.4 Optimización logística y ETA	11
1.10.5 Detección de fraude y riesgo	12
1.10.6 Expansión a Ecuador	12
1.11 Indicadores de éxito	12
1.12 Riesgos principales de la solución	14

Conclusión del capítulo	14
Perspectiva gerencial	15
Capítulo 2: Procesamiento y Visualización de Datos	16
Análisis de la relación entre turismo y desigualdad de ingresos en Ecuador	16
2.1 Introducción del capítulo	16
2.2 Planteamiento del problema	16
2.3 Objetivo general del capítulo	17
2.4 Objetivos específicos	17
2.5 Fuentes de datos utilizadas	18
2.5.1 Fuente de datos turísticos	18
2.5.2 Fuente de datos sobre desigualdad	18
2.5.3 Selección de Ecuador como caso de estudio	19
2.6 Procesos ETL/ELT aplicados al análisis	19
2.6.1 Extracción de datos	19
2.6.2 Transformación de datos	19
2.6.3 Carga de datos	20
2.7 Calidad de datos	21
2.7.1 Dimensiones de calidad consideradas	21
2.7.2 Riesgos de calidad de datos	22
2.8 Gobernanza de datos	22
2.8.1 Definición de variables	22
2.8.2 Responsabilidad sobre los datos	23
2.8.3 Control de versiones	23
2.8.4 Uso ético de los datos	23
2.9 Modelado de datos	23
2.9.1 Modelo conceptual	23
2.9.2 Modelo lógico	24
2.9.3 Modelo analítico	24
2.10 Visualización de datos	24
2.10.1 Tipo de visualización recomendada	25
2.10.2 Diseño visual sugerido	25
2.10.3 Interpretación de la visualización	26
2.11 Storytelling con datos	27
2.12 Aplicación práctica empresarial y pública	28
2.12.1 Aplicación para el sector público	28

2.12.2 Aplicación para empresas turísticas	28
2.12.3 Aplicación para inversionistas	28
2.13 Limitaciones del análisis	29
2.14 Recomendaciones analíticas	29
2.15 Indicadores propuestos para seguimiento	30
Conclusión del capítulo	30
Perspectiva gerencial	31
Capítulo 3: Gestión y Diseño de Proyectos Big Data	32
3.1 Introducción del capítulo	32
3.2 Planteamiento del problema	32
3.3 Objetivo general del capítulo	33
3.4 Objetivos específicos	33
3.5 Metodología del proyecto: CRISP-DM	34
3.5.1 Fase 1: Comprensión del negocio	34
3.5.2 Fase 2: Comprensión de los datos	34
3.5.3 Fase 3: Preparación de datos	35
3.5.4 Fase 4: Modelado	35
3.5.5 Fase 5: Evaluación	36
3.5.6 Fase 6: Despliegue propuesto	36
3.6 Alcance del proyecto	37
3.6.1 Dentro del alcance	37
3.6.2 Fuera del alcance	37
3.7 Planificación del proyecto	37
3.8 Recursos necesarios	38
3.8.1 Recursos humanos	38
3.8.2 Recursos tecnológicos	39
3.8.3 Recursos de gestión	39
3.9 Matriz RACI del proyecto	39
3.10 Gestión de riesgos	40
3.10.1 Riesgos técnicos	40
3.10.2 Riesgos operativos	41
3.10.3 Riesgos éticos y regulatorios	41
3.10.4 Matriz de riesgos	41
3.11 KPIs del proyecto	42
3.11.1 KPIs de gestión del proyecto	42

3.11.2 KPIs técnicos del modelo	42
3.11.3 KPIs de impacto comercial	43
3.12 Resultados del proyecto analítico	43
3.12.1 Preparación y selección de atributos	43
3.12.2 Comparación de modelos supervisados	44
3.12.3 Clasificación no supervisada	45
3.13 Interpretación gerencial de resultados	46
3.14 Aplicación práctica empresarial	46
3.14.1 Etapa 1: Piloto analítico	46
3.14.2 Etapa 2: Campaña segmentada	47
3.14.3 Etapa 3: Monitoreo y mejora	47
3.15 Consideraciones éticas	47
3.16 Recomendaciones para la gestión del proyecto	47
Conclusión del capítulo	48
Perspectiva gerencial	49
Capítulo 4: Economía y Finanzas para la Gestión de Proyectos	50
4.1 Introducción del capítulo	50
4.2 Planteamiento del problema financiero	50
4.3 Objetivo general del capítulo	51
4.4 Objetivos específicos	51
4.5 Soporte teórico financiero	52
4.6 Descripción de las empresas analizadas	52
4.6.1 Indurama Ecuador S.A.	52
4.6.2 Mabe Ecuador S.A.	53
4.6.3 Fibroacero S.A.	54
4.7 Análisis de ventas	55
4.7.1 Indurama: liderazgo en escala	55
4.7.2 Mabe: variabilidad comercial	55
4.7.3 Fibroacero: menor escala y presión competitiva	56
4.8 Análisis de utilidad bruta	56
4.8.1 Indurama: recuperación progresiva	56
4.8.2 Mabe: utilidad bruta dependiente del volumen	56
4.8.3 Fibroacero: mejora bruta con debilidad neta	57
4.9 Análisis de márgenes financieros	57
4.9.1 Margen bruto	57

4.9.2 Margen neto	57
4.10 Análisis de liquidez	58
4.10.1 Mabe: mayor holgura financiera	58
4.10.2 Indurama: posición intermedia	58
4.10.3 Fibroacero: riesgo de liquidez	59
4.11 Evaluación financiera del proyecto empresarial	59
4.11.1 ROI aplicado al caso	59
4.11.2 VAN aplicado al caso	60
4.11.3 TIR aplicada al caso	60
4.12 Costos del proyecto de análisis financiero	61
4.12.1 Costos directos	61
4.12.2 Costos indirectos	61
4.12.3 Costos de oportunidad	61
4.13 Impacto económico esperado	61
4.13.1 Mejora en control de costos	62
4.13.2 Fortalecimiento de rentabilidad	62
4.13.3 Recuperación financiera	62
4.13.4 Mejor toma de decisiones	62
4.14 Análisis de riesgos financieros	62
4.14.1 Riesgo de rentabilidad	62
4.14.2 Riesgo de liquidez	62
4.14.3 Riesgo operativo	62
4.14.4 Riesgo financiero	63
4.14.5 Riesgo estratégico	63
4.15 Matriz de riesgos financieros	63
4.16 KPIs financieros propuestos	64
4.17 Recomendaciones financieras por empresa	64
4.17.1 Recomendaciones para Indurama	64
4.17.2 Recomendaciones para Mabe	65
4.17.3 Recomendaciones para Fibroacero	65
4.18 Aplicación gerencial del análisis financiero	66
Conclusión del capítulo	66
Perspectiva gerencial	67
Capítulo 5: Competitividad y Transformación Digital	69
5.1 Introducción del capítulo	69

5.2 Planteamiento del problema	69
5.3 Objetivo general del capítulo	70
5.4 Objetivos específicos	70
5.5 Transformación digital como eje de competitividad	71
5.6 Riesgos estratégicos del proyecto	79
Conclusión del capítulo	81
Perspectiva gerencial	82
Bibliografía	83

Índice de Tablas

Tabla 1. Modelos supervisados	45
Tabla 2. Modelo no supervisado	45
Tabla 3. Estado financiero Indurama	53
Tabla 4. Estado financiero Mabe	54
Tabla 5. Estado financiero Fibroacero	55
Tabla 6. Márgenes e ingresos 2021-2024	67

Índice de Figuras

Figura 1. Relación entre turismo e Índice de Gini en Ecuador

26

Capítulo 1: Big Data Solutions

Proyecto de Big Data para la optimización comercial y expansión de Mercado Libre en Ecuador

1.1 Introducción del capítulo

El presente capítulo desarrolla la propuesta de Big Data Solutions para Mercado Libre, orientada a fortalecer la toma de decisiones basada en datos, optimizar la experiencia del usuario, mejorar la eficiencia comercial y logística, reducir riesgos operativos y evaluar la expansión hacia el mercado ecuatoriano. La propuesta parte del trabajo base del módulo, donde se plantea que Mercado Libre requiere una plataforma de datos capaz de integrar información transaccional, comportamental, logística, financiera y externa para responder preguntas estratégicas sobre demanda, oferta, fraude, costos de entrega y potencial de crecimiento en Ecuador.

Desde una perspectiva gerencial, el problema no se limita a almacenar grandes volúmenes de información. El reto principal consiste en convertir datos dispersos en capacidades analíticas accionables para decidir qué productos priorizar, en qué ciudades iniciar operaciones, qué vendedores incorporar, cómo reducir el costo logístico, cómo personalizar la experiencia de compra y cómo anticipar riesgos de fraude. En este sentido, Big Data se convierte en una herramienta estratégica porque permite integrar datos de alta variedad, velocidad y volumen dentro de una arquitectura escalable, gobernada y orientada a resultados de negocio.

Mercado Libre opera en un entorno altamente competitivo, donde la experiencia digital, la rapidez logística, la confianza en los pagos y la personalización son factores críticos de diferenciación. Su propio informe de inversionistas describe a la empresa como un ecosistema regional de comercio electrónico y fintech, con una posición relevante en América Latina y con oportunidades de crecimiento derivadas del aumento de la penetración del comercio electrónico en la región. En este contexto, una solución Big Data no debe entenderse como un proyecto tecnológico aislado, sino como una plataforma habilitadora de crecimiento, eficiencia operativa y ventaja competitiva sostenible.

1.2 Definición del problema desde datos

El problema central identificado es que Mercado Libre necesita tomar decisiones de expansión, personalización, pricing, logística y control de riesgo en un contexto caracterizado por alta complejidad de datos. La organización genera información en múltiples puntos de contacto: búsquedas, clics, sesiones, transacciones, pagos, devoluciones, inventario, campañas de marketing, interacciones con vendedores, tracking logístico y señales de fraude. Además, para evaluar Ecuador como mercado potencial, requiere incorporar información externa sobre ciudades, poder adquisitivo, comportamiento digital, infraestructura logística, estadísticas públicas, clima, mapas, datos macroeconómicos y registros de potenciales vendedores.

Según el trabajo, se identifican preguntas de negocio clave para la expansión a Ecuador: qué segmentos, categorías y ciudades muestran mayor tracción; cuál es la elasticidad de precio por categoría; qué zonas presentan brechas de oferta o tiempos de entrega altos; qué mix de marketing reduce el costo de adquisición de cliente; dónde se concentran los focos de fraude; y cuál es el tamaño de mercado, ticket promedio y capacidad logística esperada en ciudades piloto como Quito, Guayaquil y Cuenca. Estas preguntas evidencian que el problema no puede resolverse con reportes tradicionales o análisis aislados, porque requiere integrar información histórica, datos en tiempo real, modelos predictivos y tableros gerenciales.

Desde el enfoque de Big Data, el problema puede formularse de la siguiente manera:

Mercado Libre requiere una arquitectura de datos escalable, gobernada y orientada a analítica avanzada que permita optimizar sus operaciones actuales y reducir la incertidumbre de entrada al mercado ecuatoriano mediante integración multifuente, procesamiento batch y streaming, modelos de machine learning, tableros ejecutivos y mecanismos de gobierno, seguridad y calidad de datos.

La problemática se puede dividir en cinco dimensiones:

- **Dimensión comercial.** La empresa necesita incrementar conversión, el valor bruto de mercancía y el valor promedio de pedido y retención mediante personalización, recomendaciones, pricing dinámico y campañas segmentadas. Sin una plataforma integrada, las decisiones comerciales pueden depender de análisis parciales, lo que limita la capacidad de identificar patrones de compra, afinidades entre productos y oportunidades por categoría.
- **Dimensión logística.** La expansión a Ecuador exige comprender tiempos de entrega, costos por pedido, desempeño de carriers, rutas, zonas de cobertura y cumplimiento de compromiso mínimo de tiempo.. Una mala estimación logística puede deteriorar la experiencia del cliente y afectar la rentabilidad del modelo.
- **Dimensión de riesgo y fraude.** Las plataformas de comercio electrónico enfrentan riesgos asociados a contracargos, cuentas falsas, comportamiento anómalo, abuso de promociones y operaciones sospechosas. El documento base plantea la reducción de pérdidas por fraude como uno de los objetivos esperados del proyecto.
- **Dimensión de datos.** La variedad de fuentes internas y externas exige estándares de integración, calidad, metadatos, linaje, contratos de datos y seguridad. NIST (National Institute of Standards and Technology) plantea que una arquitectura de referencia de Big Data debe considerar componentes funcionales, roles, interfaces e interoperabilidad para sostener soluciones escalables y confiables.

- **Dimensión gerencial.** La alta dirección necesita indicadores confiables para decidir si avanzar, ajustar o detener la expansión. Por ello, el proyecto debe generar información para comités ejecutivos, líderes comerciales, equipos de logística, marketing, riesgo, tecnología y finanzas.

1.3 Objetivo general del capítulo

Diseñar una solución Big Data para Mercado Libre que permita integrar, procesar, gobernar y analizar datos internos y externos con el fin de optimizar la operación comercial actual y evaluar la expansión a Ecuador mediante casos de uso analíticos de alto impacto.

1.4 Objetivos específicos

1. Definir el problema empresarial desde una perspectiva de datos, vinculando las necesidades de negocio con capacidades analíticas concretas.
2. Proponer una arquitectura Big Data tipo lakehouse cloud-first con procesamiento batch, streaming, gobierno de datos y MLOps.
3. Identificar fuentes de datos internas y externas necesarias para la solución.
4. Diseñar una estrategia de integración de datos que permita trazabilidad, calidad, seguridad y disponibilidad.
5. Plantear casos de uso analíticos prioritarios para personalización, pricing, demanda, logística, fraude y expansión a Ecuador.
6. Definir indicadores de éxito que permitan evaluar el impacto técnico y gerencial del proyecto.

1.5 Alcance de la solución Big Data

El alcance funcional de la solución incluye la construcción de una plataforma de datos que permita ingesta multifuente, almacenamiento escalable, procesamiento distribuido, modelado analítico, despliegue de modelos predictivos, visualización ejecutiva y gobierno de datos. El documento base establece como alcance preliminar la ingesta multifuente, una arquitectura lakehouse, streaming, modelos de recomendación, demanda y fraude, dashboards ejecutivos y un laboratorio de mercado para Ecuador .

La solución incluye los siguientes componentes:

- Ingesta de datos internos y externos.
- Procesamiento batch para información histórica y agregada.
- Procesamiento streaming para eventos de comportamiento, fraude y tracking logístico.
- Almacenamiento en capas raw, clean y curated.

- Modelos de machine learning para recomendación, demanda, fraude y ETA.
- Tableros ejecutivos y operativos.
- Gobierno de datos, calidad, seguridad, privacidad y linaje.
- MLOps para despliegue, monitoreo y actualización de modelos.

Quedan fuera del alcance inicial el reemplazo completo del data warehouse legado, una infraestructura on-premise de gran escala y modelos generativos de atención al cliente a gran escala, tal como se definió en el trabajo base del módulo . Esta delimitación es importante porque permite controlar costos, reducir complejidad y priorizar capacidades con impacto directo en el negocio.

1.6 Soporte teórico: Big Data como capacidad estratégica

Big Data puede entenderse como un conjunto de capacidades tecnológicas, organizacionales y analíticas que permiten capturar, almacenar, procesar y analizar grandes volúmenes de datos estructurados, semiestructurados y no estructurados. Su valor estratégico no reside únicamente en la acumulación de información, sino en la capacidad de transformar los datos en conocimiento útil, oportuno y confiable para la toma de decisiones empresariales. En este sentido, Big Data permite a las organizaciones identificar patrones, anticipar riesgos, optimizar procesos y generar ventajas competitivas sostenibles mediante el uso de analítica avanzada y métricas orientadas al desempeño (Marr, 2015). Asimismo, una arquitectura Big Data debe considerar componentes funcionales, roles, interfaces e interoperabilidad para garantizar soluciones escalables, seguras y alineadas con las necesidades del negocio (Chang et al., 2019)

Desde la arquitectura, NIST propone una visión de Big Data basada en componentes funcionales, roles, interfaces, interoperabilidad y flujos de datos, lo que permite diseñar soluciones independientes de proveedor y alineadas con necesidades empresariales. Esta perspectiva es relevante para Mercado Libre porque la solución debe soportar múltiples dominios: comercio, pagos, logística, marketing, riesgo y expansión internacional.

Por otra parte, la gestión de datos requiere gobierno, calidad, seguridad y alineación estratégica. DAMA International (2017) plantea que la gestión de datos debe integrarse a los objetivos del negocio, considerando gobierno, cumplimiento, seguridad, estandarización y eficiencia operativa . En este proyecto, dichos principios se traducen en la necesidad de definir propietarios de datos, reglas de calidad, controles de acceso, políticas de privacidad y métricas de confiabilidad.

Finalmente, los modelos de machine learning requieren prácticas de machine learning operations para ser útiles en producción. Google Cloud define machine learning operations como un conjunto de prácticas para gestionar el ciclo de vida de modelos de machine learning, incluyendo experimentación, despliegue, monitoreo y reentrenamiento . Esto es clave para los casos de uso de

Mercado Libre, ya que modelos como recomendaciones, fraude, tiempo estimado de llegada o demanda pierden valor si no se monitorean frente a cambios en comportamiento de clientes, vendedores, productos o condiciones logísticas.

1.7 Arquitectura propuesta: lakehouse cloud-first con streaming

La arquitectura propuesta es un modelo lakehouse cloud-first con procesamiento batch y streaming, utilizando servicios de Google Cloud Platform como referencia tecnológica. El documento base del módulo propone GCP como proveedor, con Pub/Sub, Dataflow, Datastream o Debezium, Cloud Storage, Dataproc, Cloud Composer, BigQuery, Looker y Vertex AI como componentes centrales.

El enfoque lakehouse permite combinar la flexibilidad de un data lake con capacidades analíticas propias de un data warehouse. Esto resulta adecuado para Mercado Libre porque la empresa requiere almacenar datos crudos de gran variedad, pero también necesita datos curados y confiables para tableros, modelos y decisiones ejecutivas. La arquitectura debe responder a cuatro principios: escalabilidad, interoperabilidad, gobierno y orientación a casos de uso.

1.7.1 Capa de ingesta

La capa de ingesta se encarga de capturar datos desde fuentes internas y externas. Para eventos web y app, se recomienda usar procesamiento streaming mediante Pub/Sub y Dataflow. Para sistemas transaccionales como ERP, CRM, OMS o WMS, se puede utilizar CDC mediante Datastream o Debezium. Para fuentes SaaS y APIs externas, se propone usar Cloud Run Jobs y Workflows, con controles de rate limiting, reintentos y monitoreo.

Esta capa permite recibir datos de:

- Clics.
- Búsquedas.
- Impresiones.
- Sesiones.
- Pedidos.
- Pagos.
- Devoluciones.
- Tracking logístico.
- Inventario.
- Campañas.
- Señales de fraude.
- Información externa de mercado.

La ingesta debe estar gobernada por contratos de datos que definan esquema, frecuencia, propietario, reglas de calidad, campos sensibles y SLA. Esto evita que cambios no controlados en una fuente afecten modelos, dashboards o decisiones.

1.7.2 Capa de almacenamiento

La solución propone una estructura en tres zonas:

Zona raw. Conserva los datos tal como llegan desde las fuentes. Su función es preservar trazabilidad, permitir reprocesamiento y mantener evidencia histórica.

Zona clean. Contiene datos depurados, estandarizados y validados. Aquí se aplican reglas de calidad, normalización de identificadores, conversión de formatos, eliminación de duplicados y validación de consistencia.

Zona curated. Contiene datos listos para consumo analítico, modelos predictivos y dashboards. Incluye tablas de hechos, dimensiones, métricas certificadas y datasets por dominio.

El documento base sugiere formatos Parquet, Delta o Iceberg para analítica, Avro o JSON para eventos y CSV solo para intercambios puntuales controlados. Esta decisión es coherente con una arquitectura moderna porque facilita eficiencia de consulta, compresión, control de esquemas y escalabilidad.

1.7.3 Capa de procesamiento

La arquitectura contempla dos tipos de procesamiento:

Procesamiento batch. Se utiliza para consolidar ventas diarias, inventarios, campañas, costos, indicadores financieros, desempeño por categoría y entrenamiento de modelos históricos. Puede ejecutarse mediante Dataproc y orquestarse con Cloud Composer. (Google Cloud, s. f.-a; Google Cloud, s. f.-b)

Procesamiento streaming. Se utiliza para eventos que requieren baja latencia, como recomendaciones en tiempo real, scoring de fraude, tracking logístico, alertas operativas y actualización de señales de comportamiento. El documento base propone latencias inferiores a cinco minutos para flujos streaming y freshness menor a 60 minutos en tablas críticas. (Google Cloud, s. f.-a; Google Cloud, s. f.-b)

1.7.4 Capa analítica y machine learning

La capa analítica incluye modelos predictivos, segmentación, experimentación A/B y feature store. Para Mercado Libre, esta capa permite responder preguntas como:

- ¿Qué producto recomendar a cada usuario?
- ¿Qué precio o promoción maximiza margen y conversión?
- ¿Qué categoría tendrá mayor demanda en una ciudad?
- ¿Qué pedido tiene mayor riesgo de fraude?
- ¿Qué ruta logística reduce tiempo y costo?
- ¿Qué ciudad ecuatoriana ofrece mejor balance entre demanda, oferta y capacidad operativa?

Google Cloud señala que MLOps permite automatizar la integración, entrega y entrenamiento continuo de modelos de machine learning, lo cual es relevante para mantener modelos confiables en producción .

1.7.5 Capa de visualización y consumo

La capa de visualización debe traducir los datos en decisiones. Se propone Looker como herramienta de BI y BigQuery como motor analítico. Los tableros deben estar diferenciados por rol:

- **Alta dirección:** GMV, conversión, rentabilidad, expansión, riesgo, ROI.
- **Comercial:** AOV, CAC, LTV, conversión, categorías foco.
- **Marketing:** ROAS, performance de campañas, cohortes, retención.
- **Logística:** OTD, costo por pedido, SLA, tiempos de entrega, fill rate.
- **Riesgo:** tasa de fraude, chargebacks, alertas, precisión y recall.
- **Data Office:** calidad, freshness, disponibilidad, incidentes, adopción BI.

El documento base ya plantea dashboards ejecutivos, operativos, de marketing y riesgo con indicadores como GMV, AOV, conversión, CAC, LTV, NPS, OTD, SLA, ROAS, tasa de fraude y chargebacks .

1.8 Fuentes de datos requeridas

La solución necesita integrar fuentes internas y externas. Las fuentes internas permiten entender la operación actual, mientras que las fuentes externas reducen la incertidumbre de entrada al mercado ecuatoriano.

1.8.1 Fuentes internas

Las principales fuentes internas son:

- CRM.
- ERP.
- OMS.
- WMS.
- Catálogo de productos.
- Inventario.
- Pricing.
- Transacciones.
- Clickstream.
- Búsquedas.
- Campañas.
- Fulfillment.
- Devoluciones.
- Soporte al cliente.
- Pagos.
- Fraude.
- Tracking logístico.

Estas fuentes permiten construir una visión integral del cliente, vendedor, producto, pedido y operación logística. El documento base establece la necesidad de unificar identificadores de usuario, vendedor, producto, pedido y envío; normalizar catálogos, monedas, impuestos y duplicación transaccional; y definir SLA de calidad para freshness, completeness, consistency y validity .

1.8.2 Fuentes externas

Las fuentes externas recomendadas son:

- Estadísticas públicas de Ecuador.
- Datos macroeconómicos.
- Tendencias de búsqueda.
- Redes sociales.
- Marketplaces comparables.
- Mapas y datos viales.
- Clima.
- Información de carriers y SLA logísticos.
- Bureaus de riesgo.

- Registros tributarios y comerciales.
- Datos demográficos por ciudad.

El documento base menciona fuentes locales como instituto nacional de estadísticas y censos (INEC), Banco Central del Ecuador, servicio de rentas internas (SRI), Ministerio de Telecomunicaciones, municipios de Quito, Guayaquil y Cuenca, INAMHI y OpenStreetMap para enriquecer la evaluación de mercado, logística y riesgo .

1.9 Estrategia de integración de datos

La integración de datos debe diseñarse como una capacidad empresarial, no como una tarea técnica aislada. Para ello se propone una estrategia basada en cinco componentes.

1.9.1 Identificadores maestros

La solución debe unificar identificadores de clientes, vendedores, productos, pedidos, pagos, envíos y campañas. Esto permite conectar eventos de navegación con transacciones, devoluciones, pagos, reclamos y comportamiento logístico.

1.9.2 Contratos de datos

Cada fuente debe contar con un contrato que especifique:

- Propietario del dato.
- Descripción funcional.
- Esquema.
- Frecuencia.
- Campos obligatorios.
- Reglas de calidad.
- Clasificación de sensibilidad.
- SLA de actualización.
- Política de retención.
- Impacto aguas abajo.

El documento base incluye un catálogo mínimo de datos y contratos con hechos de pedidos, líneas de pedido, envíos, eventos de tracking, pagos, devoluciones, clics, búsquedas, impresiones, sesiones, campañas y redenciones de cupones; además de dimensiones de usuario, vendedor, producto, categoría, ciudad, canal y dispositivo.

1.9.3 Calidad de datos

Las dimensiones mínimas de calidad son completitud, validez, consistencia, unicidad, oportunidad y trazabilidad. DAMA-DMBOK considera la calidad de datos como una disciplina clave dentro de la gestión de datos, vinculada con definiciones, estándares, controles y mejora continua.

En el caso de Mercado Libre, la calidad de datos impacta directamente en decisiones de pricing, recomendaciones, detección de fraude y estimación de demanda. Un producto mal categorizado, una ciudad mal asignada o un pedido duplicado pueden alterar modelos y decisiones gerenciales.

1.9.4 Seguridad y privacidad

La arquitectura debe incluir cifrado en reposo y tránsito, control de acceso basado en roles y atributos, enmascaramiento de datos sensibles, tokenización de PII, auditoría, gestión de secretos y políticas de retención. El documento base plantea prácticas como minimización, masking, tokenization, acceso mínimo necesario, auditoría, linaje y logs inmutables.

1.9.5 Observabilidad

La plataforma debe monitorear disponibilidad, latencia, costos, calidad, drift de datos, drift de modelos y errores de pipelines. Esta observabilidad permite anticipar fallas y sostener la confianza de usuarios internos y gerentes.

1.10 Casos de uso analíticos prioritarios

1.10.1 Recomendaciones y personalización

El primer caso de uso consiste en recomendar productos relevantes según comportamiento de navegación, historial de compra, afinidad por categoría, ubicación, dispositivo, sensibilidad al precio y comportamiento de usuarios similares. El documento base propone técnicas como candidate generation, ranking, matrix factorization y modelos tipo XGBoost/GBDT, evaluados mediante uplift en CTR y conversión.

Su impacto gerencial se refleja en mayor conversión, incremento del AOV, mejora de experiencia y mayor eficiencia en campañas.

1.10.2 Pricing dinámico y promociones inteligentes

Este caso de uso busca modelar elasticidad de precio por categoría, ciudad, segmento y canal. Permite definir descuentos, cupones y promociones con mayor precisión, evitando sacrificar margen innecesariamente.

Indicadores asociados:

- Margen por categoría.
- Conversión incremental.
- Tasa de uso de cupones.
- Uplift de ventas.
- Rentabilidad por campaña.

1.10.3 Predicción de demanda y abastecimiento

La predicción de demanda permite anticipar qué productos tendrán mayor rotación por ciudad, zona y temporada. En Ecuador, este caso es clave para definir categorías semilla, profundidad de surtido y prioridades de seller onboarding.

Indicadores asociados:

- MAPE del forecast.
- Stockout rate.
- Fill rate.
- GMV por categoría.
- Cobertura de inventario.

1.10.4 Optimización logística y ETA

La optimización logística busca mejorar rutas, tiempos estimados de entrega, asignación de carriers y costos por pedido. Este caso es crítico porque la experiencia de entrega influye directamente en recompra, NPS y confianza del usuario.

Indicadores asociados:

- OTD.
- Costo logístico por pedido.
- Tiempo de entrega promedio.
- SLA por carrier.
- First mile y last mile performance.

1.10.5 Detección de fraude y riesgo

El caso de fraude combina reglas, modelos supervisados, graph analytics, behavioral signals y monitoreo en tiempo real. Su objetivo es reducir contracargos, abuso de promociones, transacciones sospechosas y pérdidas operativas.

Indicadores asociados:

- Tasa de fraude.
- Chargeback rate.
- Precisión.
- Recall.
- Pérdidas evitadas.
- Tiempo de respuesta ante alerta.

1.10.6 Expansión a Ecuador

Este caso de uso integra market sizing, análisis de ciudades, demanda potencial, capacidad logística, oferta de vendedores, categorías prioritarias, riesgo y rentabilidad. El documento base plantea monitorear la expansión mediante usuarios activos manuales, compradores activos, vendedores activos, valor bruto de mercadería, entregas a tiempo y puntaje neto de recomendación.

La expansión debe iniciar con ciudades piloto como Quito, Guayaquil y Cuenca, porque concentran actividad económica, conectividad y capacidad logística. La decisión de escalar debe depender de evidencia, no de supuestos comerciales.

1.11 Indicadores de éxito

Los indicadores deben evaluar tanto la plataforma como el negocio. El documento base propone resultados esperados en 12 meses: incremento de conversión entre 8 % y 12 %, aumento de GMV entre 10 % y 15 % en categorías prioritarias, reducción de costo logístico por pedido entre 10 % y 15 %, reducción de fraude entre 30 % y 50 %, time to insight menor a 24 horas para batch y menor a cinco minutos para streaming, además de metas de adopción inicial en Ecuador .

Los principales KPIs son:

Comerciales

- GMV
- AOV

- Conversión
- Margen por categoría

Marketing y clientes

- CAC
- LTV
- Retención M3/M6
- ROAS
- NPS

Logística

- OTD.
- Costo por pedido.
- Fill rate.
- Stockouts.
- Tiempo promedio de entrega.

Riesgo

- Tasa de fraude
- Chargeback rate
- Pérdidas evitadas
- Precisión y recall del modelo

Plataforma de datos

- Disponibilidad
- Freshness
- Completitud
- Validez
- Unicidad
- Latencia de inferencia
- Pipelines monitoreados

Adopción BI

- Usuarios activos BI

- Sesiones BI
- Tiempo hasta decisión
- Uso de dashboards por rol

1.12 Riesgos principales de la solución

Los riesgos más relevantes son:

Baja calidad de datos. Puede afectar recomendaciones, pricing, demanda y fraude. Se mitiga con reglas automáticas, contratos de datos, owners y monitoreo.

Retrasos en integración. La variedad de fuentes puede generar demoras. Se mitiga priorizando fuentes de alto impacto y usando arquitectura incremental.

Sobrecostos cloud. El procesamiento intensivo puede elevar costos. Se mitiga con FinOps, autoscaling, control de jobs, almacenamiento por capas y monitoreo.

Riesgo de privacidad. El uso de datos personales requiere controles estrictos. Se mitiga con minimización, anonimización, cifrado y acceso mínimo.

Resistencia organizacional. Los equipos pueden no adoptar dashboards o modelos. Se mitiga con capacitación, data champions, tableros por rol y decisiones visibles basadas en evidencia.

Vendor lock-in. La dependencia de un proveedor cloud puede limitar flexibilidad. Se mitiga con formatos abiertos, infraestructura como código y diseños portables.

Conclusión del capítulo

El capítulo demuestra que una solución Big Data para Mercado Libre debe diseñarse como una capacidad estratégica integral, no únicamente como infraestructura tecnológica. El problema de negocio exige integrar datos comerciales, logísticos, financieros, transaccionales, comportamentales y externos para responder preguntas críticas sobre crecimiento, eficiencia, riesgo y expansión geográfica.

La arquitectura lakehouse cloud-first con procesamiento batch y streaming permite sostener casos de uso de alto impacto como recomendaciones, pricing dinámico, predicción de demanda, optimización logística, detección de fraude y evaluación de entrada a Ecuador. Además, la incorporación de gobierno, calidad, seguridad, observabilidad y MLOps asegura que los datos y modelos puedan utilizarse de manera confiable en decisiones reales.

En el contexto del proyecto de tesis, este capítulo cumple una función fundacional: define el problema desde datos, establece la arquitectura objetivo, identifica las fuentes necesarias y conecta la tecnología con resultados gerenciales medibles. Por ello, sirve como base para los capítulos posteriores, especialmente procesamiento y visualización de datos, gestión de proyectos Big Data, evaluación financiera y transformación digital.

Perspectiva gerencial

Desde la perspectiva gerencial, la propuesta de Big Data Solutions aporta valor en tres niveles. Primero, mejora la toma de decisiones estratégicas, porque permite evaluar la expansión a Ecuador con evidencia sobre demanda, ciudades, categorías, vendedores, logística y riesgo. Segundo, incrementa la eficiencia operativa, al optimizar rutas, inventarios, campañas, precios y detección de fraude. Tercero, contribuye a la generación de valor, porque impacta indicadores como conversión, GMV, AOV, CAC, LTV, OTD, costo por pedido y pérdidas por fraude.

La alta dirección puede utilizar esta solución para decidir cuándo lanzar pilotos, qué ciudades priorizar, qué categorías impulsar, cuánto invertir en marketing, qué carriers contratar y qué umbrales exigir antes de escalar. En consecuencia, Big Data deja de ser un recurso técnico y se convierte en un sistema de gestión estratégica orientado a crecimiento rentable, control de riesgos y ventaja competitiva.

Capítulo 2: Procesamiento y Visualización de Datos

Análisis de la relación entre turismo y desigualdad de ingresos en Ecuador

2.1 Introducción del capítulo

El presente capítulo desarrolla el módulo de Procesamiento y Visualización de Datos, aplicado al análisis de la relación entre el turismo y la desigualdad de ingresos en Ecuador. El estudio parte de una problemática relevante para la gestión pública, empresarial y social: determinar si el crecimiento del turismo puede contribuir a mejorar la distribución del ingreso o, por el contrario, concentrar los beneficios económicos en determinados sectores, territorios o grupos sociales.

La actividad base del módulo tuvo como objetivo relacionar datos de turismo con la desigualdad de ingresos de un país específico mediante la elaboración de visualizaciones claras y comprensibles. Para ello, se seleccionó Ecuador como caso de estudio y se utilizaron dos fuentes principales: una base de datos turística denominada “Tourist and Economic Impact – Data Set”, obtenida de Kaggle, y una base sobre el Índice de Gini, obtenida de Datosmacro. A partir de estas fuentes, se construyó una base integrada con las variables país, año, índice de Gini, turismo y tipo de desigualdad de ingresos.

Desde el enfoque de la Maestría en Gerencia de Datos y Negocios, este capítulo no se limita a describir una visualización, sino que transforma el ejercicio académico en una propuesta analítica con valor gerencial. El procesamiento y la visualización de datos permiten identificar patrones, examinar relaciones entre variables y generar evidencia para la toma de decisiones. En este caso, el análisis busca establecer si existe una relación observable entre el turismo y la desigualdad de ingresos en Ecuador, así como valorar si dicha relación puede interpretarse como una oportunidad para diseñar políticas, inversiones y estrategias empresariales más inclusivas.

El Índice de Gini es una medida ampliamente utilizada para analizar desigualdad. El Banco Mundial define este indicador como una medida de distribución del ingreso, donde un valor de 0 representa igualdad perfecta y un valor de 100 representa desigualdad perfecta (Banco Mundial, s. f.). Esta definición es importante porque permite interpretar que una disminución del Índice de Gini refleja una mejora relativa en la distribución del ingreso.

2.2 Planteamiento del problema

El turismo es una actividad económica con capacidad para generar empleo, dinamizar servicios, atraer divisas, impulsar emprendimientos locales y fortalecer cadenas productivas. De acuerdo con el Banco Mundial (2025), el turismo constituye uno de los motores más inclusivos y escalables para la creación de empleo y el desarrollo económico. Sin embargo, su impacto sobre la desigualdad no es automático, ya que depende de la forma en que se distribuyen sus beneficios entre los distintos actores del territorio. En algunos contextos, estos beneficios pueden llegar a hoteles, restaurantes,

transporte, guías, artesanos, comunidades locales y pequeños negocios; mientras que, en otros casos, pueden concentrarse en grandes operadores, zonas urbanas específicas o grupos con mayor acceso al capital. Por ello, el turismo puede contribuir al crecimiento inclusivo siempre que esté acompañado de políticas, encadenamientos productivos y mecanismos que favorezcan la participación de las comunidades locales (UNCTAD, 2017; World Bank, 2025).

En el trabajo se plantea precisamente esta inquietud: analizar si existe relación entre los datos de turismo y la desigualdad de ingresos, y determinar si el turismo puede generar empleo y oportunidades económicas que mejoren la distribución del ingreso o si, por el contrario, puede concentrar la riqueza y aumentar la desigualdad.

En Ecuador, esta discusión resulta especialmente relevante, ya que el país posee una oferta turística diversa, sustentada en su biodiversidad, patrimonio cultural, gastronomía, naturaleza y turismo comunitario. Al mismo tiempo, enfrenta retos estructurales de desigualdad de ingresos, lo que convierte al turismo en una actividad que debe analizarse no solo por su capacidad de generar ingresos, sino también por su potencial para distribuir oportunidades económicas.

El problema de investigación del capítulo puede formularse de la siguiente manera:

¿Cómo se relaciona la evolución del turismo con la desigualdad de ingresos en Ecuador, medida a través del Índice de Gini, y qué información gerencial puede obtenerse mediante el procesamiento y visualización de datos?

Esta pregunta implica tres desafíos analíticos. Primero, integrar fuentes de datos heterogéneas. Segundo, depurar y estructurar los datos para que sean comparables por país y año. Tercero, diseñar visualizaciones claras que permitan identificar patrones sin sobrecargar al usuario.

2.3 Objetivo general del capítulo

Analizar la relación entre los datos de turismo y la desigualdad de ingresos en Ecuador mediante procesos de integración, depuración, modelado y visualización de datos, con el fin de generar información útil para la toma de decisiones gerenciales, empresariales y públicas.

2.4 Objetivos específicos

1. Integrar datos de turismo y desigualdad de ingresos en una base estructurada para Ecuador.
2. Aplicar procesos de limpieza, transformación y clasificación de datos para facilitar el análisis.
3. Definir criterios de calidad y gobernanza aplicables al conjunto de datos trabajado.

4. Diseñar visualizaciones que relacionen indicadores de turismo con la evolución del Índice de Gini.
5. Interpretar los resultados desde una perspectiva analítica, empresarial y gerencial.
6. Proponer recomendaciones para el uso de visualización de datos en decisiones vinculadas al turismo y la inclusión económica.

2.5 Fuentes de datos utilizadas

Para desarrollar el análisis se recurrió a información secundaria proveniente de dos bases de datos distintas, seleccionadas por su relación directa con las variables centrales del estudio: turismo y desigualdad de ingresos. La primera fuente corresponde al conjunto de datos **“Tourist and Economic Impact – Data Set”**, disponible en Kaggle, del cual se tomó la información relacionada con la actividad turística. La segunda fuente corresponde a **Datosmacro**, de donde se obtuvo el **Índice de Gini**, indicador utilizado para medir la desigualdad en la distribución del ingreso.

La integración de ambas fuentes permitió construir una base consolidada para Ecuador, organizada por año y orientada a comparar el comportamiento del turismo con la evolución de la desigualdad. Esta combinación de datos hizo posible identificar patrones visuales y generar una interpretación inicial sobre la posible relación entre la actividad turística y la distribución del ingreso en el país.

2.5.1 Fuente de datos turísticos

La base turística utilizada permite observar el comportamiento del turismo como variable económica. Aunque el documento del módulo no detalla todas las columnas originales de Kaggle, sí indica que se tomó el dato de turismo correspondiente a Ecuador y se integró con el Índice de Gini en una base común.

Desde el punto de vista analítico, esta fuente permite aproximarse al comportamiento del turismo a través del tiempo. Dependiendo de la estructura exacta del conjunto de datos original, la variable turística puede estar asociada con llegadas de turistas, ingresos por turismo, impacto económico o indicadores relacionados con la actividad turística. Para efectos de este capítulo, se mantiene el enfoque del documento base: utilizar la variable turismo como indicador principal para contrastarla con la desigualdad de ingresos.

2.5.2 Fuente de datos sobre desigualdad

La segunda fuente utilizada fue Datosmacro, específicamente el indicador Índice de Gini. Este índice permite medir la desigualdad en la distribución del ingreso. De acuerdo con el Banco Mundial, el Índice de Gini representa la distribución del ingreso dentro de una economía, donde valores más bajos indican mayor igualdad y valores más altos reflejan mayor desigualdad (Banco Mundial, s. f.).

La elección del Índice de Gini es adecuada porque permite comparar la evolución de la desigualdad en el tiempo y relacionarla con otras variables económicas. En este proyecto, se utiliza para identificar si los periodos de mayor turismo coinciden con disminuciones o aumentos en la desigualdad.

2.5.3 Selección de Ecuador como caso de estudio

El documento base justifica la selección de Ecuador porque es un país latinoamericano con alta biodiversidad, potencial turístico y dependencia del turismo como fuente de divisas, pero también con retos importantes en desigualdad de ingresos.

Esta elección es pertinente desde una perspectiva gerencial por tres razones. Primero, Ecuador cuenta con una oferta turística diversa que incluye naturaleza, cultura, patrimonio, gastronomía y turismo comunitario. Segundo, el turismo puede generar encadenamientos productivos en sectores como transporte, alojamiento, alimentación, artesanías, comercio y servicios locales. Tercero, la desigualdad de ingresos exige evaluar si el crecimiento turístico realmente beneficia a distintos grupos sociales o si se concentra en actores específicos.

2.6 Procesos ETL/ELT aplicados al análisis

Para desarrollar una visualización confiable, fue necesario aplicar un proceso de integración y preparación de datos. Aunque en el trabajo resume este proceso de manera general, se puede estructurar metodológicamente bajo las fases ETL: extracción, transformación y carga.

2.6.1 Extracción de datos

La extracción consistió en obtener información desde dos fuentes distintas: datos de turismo desde Kaggle y datos del Índice de Gini desde Datosmacro. Esta fase permitió identificar los registros correspondientes a Ecuador, seleccionar los años disponibles y preparar ambas bases para su integración.

Desde una perspectiva gerencial, esta etapa permite asegurar que el análisis se construya sobre información relevante para el país seleccionado y no sobre datos generales que podrían distorsionar la interpretación.

2.6.2 Transformación de datos

La transformación fue una etapa central del proceso, porque permitió convertir dos bases independientes en una sola estructura analítica. Según el documento original, la base final incluyó las columnas país, año, índice de Gini, turismo y tipo de desigualdad de ingresos.

Las principales transformaciones aplicadas fueron las siguientes:

- Filtrado por país.
Se seleccionaron únicamente los registros correspondientes a Ecuador, ya que el objetivo del análisis era estudiar la relación turismo-desigualdad en un contexto nacional específico.
- Homologación de años.
Se alinearon los años disponibles en ambas bases para que cada registro anual tuviera un dato de turismo y un dato de Índice de Gini.
- Normalización de nombres de variables.
Se definieron nombres comprensibles para las columnas: país, año, índice de Gini, turismo y tipo de desigualdad.
- Clasificación de desigualdad.
El documento indica que se creó una variable llamada tipo de desigualdad de ingresos, obtenida al comparar el valor del Índice de Gini con una tabla de referencia para clasificar la desigualdad como alta o baja.
- Preparación para visualización.
Los datos fueron organizados en una estructura que permitiera comparar visualmente la evolución del turismo con la evolución del Índice de Gini.

2.6.3 Carga de datos

La fase de carga corresponde a la incorporación de la base final depurada en una herramienta de visualización de datos, con el propósito de transformar los registros procesados en gráficos interpretables. En este caso, la base consolidada —compuesta por las variables país, año, turismo, Índice de Gini y tipo de desigualdad de ingresos— debió ser cargada en una plataforma de análisis visual, como Microsoft Excel, Power BI, Tableau o una herramienta equivalente, capaz de construir gráficos comparativos por año.

Esta etapa es importante porque permite pasar del dato tabular al análisis visual. Una vez cargada la información, las variables de turismo e Índice de Gini pueden representarse en un mismo gráfico o en visualizaciones relacionadas, facilitando la identificación de tendencias, variaciones y posibles relaciones entre ambas dimensiones. De esta manera, la carga no constituye únicamente el cierre técnico del proceso de preparación de datos, sino el punto de partida para comunicar hallazgos de forma clara, sencilla y útil para la toma de decisiones.

En el contexto del capítulo, esta fase permitió generar una visualización orientada a responder si existe una relación entre la evolución del turismo y la desigualdad de ingresos en Ecuador. Por ello,

la herramienta utilizada debía permitir comparar variables en el tiempo, diferenciar niveles de desigualdad y presentar los resultados de manera comprensible para usuarios no técnicos.

La estructura final de datos puede representarse de la siguiente manera:

País	Año	Índice de Gini	Turismo	Tipo de desigualdad
Ecuador	Año 1	Valor Gini	Valor turismo	Alta/Baja
Ecuador	Año 2	Valor Gini	Valor turismo	Alta/Baja
Ecuador	Año 3	Valor Gini	Valor turismo	Alta/Baja

Esta estructura facilita el análisis temporal y permite construir visualizaciones de tendencia, relación y comparación.

2.7 Calidad de datos

La calidad de datos es fundamental para garantizar que las conclusiones del análisis sean confiables. En este caso, el principal riesgo no está en el volumen de datos, sino en la correcta integración de fuentes, la correspondencia temporal y la interpretación adecuada de las variables.

La gestión de datos requiere prácticas de calidad, gobernanza, seguridad y trazabilidad para asegurar que la información utilizada en la toma de decisiones sea confiable (DAMA International, 2017). Bajo esta perspectiva, el análisis de turismo y desigualdad no debe limitarse a unir dos bases, sino que debe asegurar que los datos sean consistentes, comparables y correctamente interpretados.

2.7.1 Dimensiones de calidad consideradas

Compleitud.

Se debe verificar que cada año incluido en el análisis tenga información tanto de turismo como de Índice de Gini. Si un año tiene dato de turismo, pero no dato de desigualdad, la comparación puede quedar incompleta.

Consistencia.

Los valores deben tener coherencia entre fuentes. Por ejemplo, el año debe estar en el mismo formato y el país debe estar escrito de forma uniforme como Ecuador. Comparabilidad. Los datos deben corresponder al mismo periodo de análisis. No sería correcto comparar turismo de un año con desigualdad de otro año si no se especifica un rezago temporal.

Trazabilidad.

Cada dato debe poder relacionarse con su fuente original: Kaggle para turismo y Datosmacro para Índice de Gini.

2.7.2 Riesgos de calidad de datos

Los principales riesgos identificados son:

- Diferencias en años disponibles entre ambas bases.
- Ausencia de datos para determinados periodos.
- Falta de claridad sobre la unidad exacta de medición de la variable turismo.
- Posible necesidad de normalizar escalas para comparar visualmente variables de distinta magnitud.
- Riesgo de interpretar una correlación visual como causalidad directa.

Este último punto es especialmente importante. El documento base concluye que existe una relación inversa: cuando aumenta el turismo, la desigualdad disminuye, y viceversa. Sin embargo, desde una perspectiva académica rigurosa, esta relación debe interpretarse como un patrón observado en la visualización y no como una prueba concluyente de causalidad. Para sostener una relación causal, sería necesario aplicar modelos estadísticos adicionales y controlar otras variables, como empleo, gasto público, inflación, educación, inversión, informalidad y crecimiento económico.

2.8 Gobernanza de datos

La gobernanza de datos en este capítulo se entiende como el conjunto de reglas, responsabilidades y criterios que permiten asegurar que la información utilizada sea confiable, trazable y útil para la toma de decisión. De acuerdo con DAMA International (2017) plantea que los datos deben gestionarse como un activo organizacional, mediante políticas, roles, estándares y controles que permitan asegurar su calidad y uso adecuado. En este proyecto, la gobernanza es relevante porque el análisis aborda una problemática económica y social sensible: la desigualdad de ingresos.

2.8.1 Definición de variables

Cada variable debe tener una definición clara:

- **País.**
Unidad territorial del análisis. En este caso, Ecuador.
- **Año.**
Periodo anual utilizado para relacionar turismo e Índice de Gini.

- **Turismo.**
Indicador proveniente de la base turística seleccionada en Kaggle.
- **Índice de Gini.**
Indicador de desigualdad de ingresos obtenido de Datosmacro.
- **Tipo de desigualdad.**
Clasificación categórica creada a partir del valor del Índice de Gini.

2.8.2 Responsabilidad sobre los datos

En una aplicación real, los responsables serían:

Rol	Responsabilidad
Analista de datos	Depurar, transformar e integrar la información
Especialista en turismo	Validar la interpretación del indicador turístico
Especialista económico/social	Validar la lectura del Índice de Gini
Tomador de decisión	Interpretar resultados para diseñar acciones o políticas

2.8.3 Control de versiones

Cada versión de la base integrada debe conservarse con fecha de actualización, fuente utilizada, filtros aplicados y transformaciones realizadas. Esto permite reproducir el análisis y evitar cambios no documentados.

2.8.4 Uso ético de los datos

Aunque el análisis no trabaja con datos personales, sí aborda una problemática social sensible. Por ello, las conclusiones deben redactarse con prudencia, evitando afirmar que el turismo “soluciona” la desigualdad sin evidencia suficiente. El turismo puede ser una herramienta de inclusión económica, pero sus efectos dependen de cómo se distribuyen las oportunidades, quiénes participan en la cadena de valor y qué políticas acompañan el crecimiento del sector.

2.9 Modelado de datos

El modelado de datos propuesto para este análisis es sencillo, pero suficiente para responder la pregunta de investigación. Se trata de un modelo tabular de serie temporal, donde cada fila representa un año para Ecuador y cada columna representa una variable relevante.

2.9.1 Modelo conceptual

El modelo conceptual se basa en la relación entre dos dimensiones principales:

- Actividad turística.
- Desigualdad de ingresos.

La hipótesis visual es que, a medida que aumenta la actividad turística, el Índice de Gini tiende a reducirse. Esta hipótesis se desprende de la interpretación realizada en el documento base, donde se observa una relación inversa entre ambas variables.

2.9.2 Modelo lógico

El modelo lógico puede organizarse así:

Campo	Tipo de dato	Descripción
País	Texto	País analizado
Año	Numérico/fecha	Año de observación
Índice_Gini	Numérico	Medida de desigualdad
Turismo	Numérico	Indicador turístico seleccionado
Tipo_Desigualdad	Categorico	Clasificación según nivel de Gini

2.9.3 Modelo analítico

El modelo analítico permite realizar:

- Análisis de tendencia temporal.
- Comparación entre turismo e Índice de Gini.
- Identificación de periodos de aumento o disminución.
- Clasificación de años según tipo de desigualdad.
- Interpretación visual de relación inversa o directa.

Este modelo es útil porque permite comunicar el fenómeno de manera clara a usuarios no técnicos.

2.10 Visualización de datos

La visualización fue el eje central del trabajo del módulo. La rúbrica de la actividad valoraba que los datos estuvieran correctamente estructurados, que las soluciones gráficas permitieran identificar claramente el objetivo de la visualización y que los gráficos fueran sencillos, elegantes y comprensibles.

Una visualización efectiva debe facilitar la comprensión de los datos, resaltar patrones relevantes y apoyar la toma de decisiones mediante diseños claros y no sobrecargados (Tableau, s. f.). Por tanto, el diseño visual no debe entenderse como un elemento decorativo, sino como un recurso analítico para comunicar evidencia.

En este caso, la visualización debía responder dos preguntas:

1. ¿Existe relación entre los datos de turismo y desigualdad de ingresos?
2. Si existe relación, ¿el turismo puede generar empleo y oportunidades económicas que mejoren la distribución de ingresos, o puede concentrar la riqueza y aumentar la desigualdad?

2.10.1 Tipo de visualización recomendada

Para este análisis, las visualizaciones más adecuadas son:

- Gráfico de líneas con doble eje.
Permite observar la evolución del turismo y del Índice de Gini en el tiempo. Es útil para identificar si ambas variables se mueven en sentido contrario o en la misma dirección.
- Gráfico de dispersión.
Permite observar la relación entre turismo e Índice de Gini. Si los puntos muestran una pendiente descendente, se evidencia una relación inversa.
- Gráfico con color por tipo de desigualdad.
Permite diferenciar visualmente los años con desigualdad alta o baja.
- Tarjetas resumen.
Pueden mostrar el año de mayor turismo, el año de menor Gini y la tendencia general observada.

2.10.2 Diseño visual sugerido

Para cumplir con criterios de claridad y elegancia, la visualización debería:

- Usar pocos colores.
- Diferenciar claramente turismo e Índice de Gini.
- Incluir etiquetas comprensibles.
- Evitar exceso de texto.
- Mantener escalas visibles.
- Incluir título explicativo.
- Incorporar una nota metodológica breve.

- Evitar concluir causalidad directa desde el gráfico.

Un título adecuado podría ser:

“Relación entre turismo e Índice de Gini en Ecuador”

Y un subtítulo:

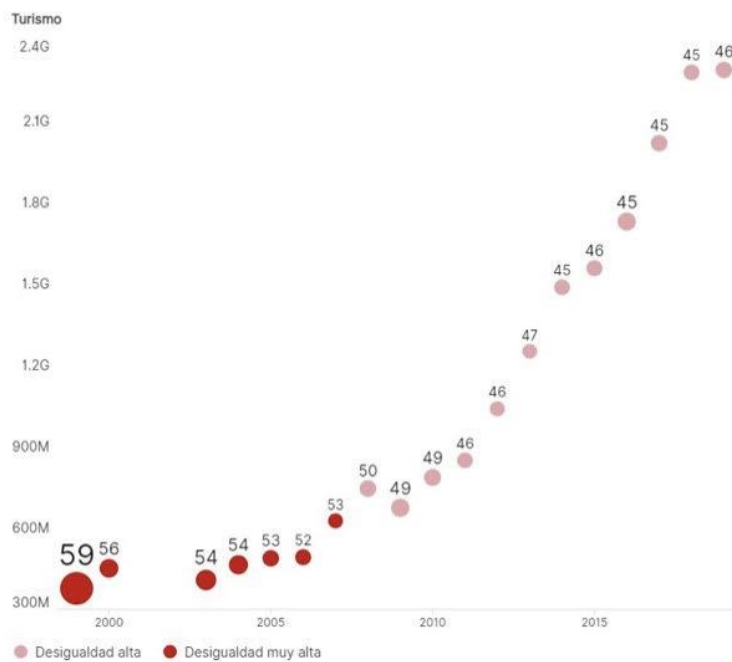
“Evolución anual de la actividad turística y la desigualdad de ingresos”

2.10.3 Interpretación de la visualización

El trabajo concluye que sí existe una relación entre los datos, ya que cuando el turismo aumenta, la desigualdad disminuye, y cuando el turismo baja, la desigualdad tiende a aumentar.

Figura 1.

Relación entre turismo e Índice de Gini en Ecuador



Esta lectura sugiere una relación inversa entre turismo y desigualdad. Desde una perspectiva económica, ello podría interpretarse como un indicio de que el turismo contribuye a generar empleo, ingresos y oportunidades económicas. UN Tourism (2025), destaca que medir el empleo turístico

permite comprender con mayor amplitud la contribución del turismo al progreso social y al cumplimiento de los Objetivos de Desarrollo Sostenible (ODS).

Sin embargo, la interpretación debe realizarse con prudencia. Una visualización permite identificar patrones, pero no demuestra por sí sola que el turismo sea la causa directa de la reducción de desigualdad. La desigualdad depende de múltiples factores, como política fiscal, empleo formal, educación, distribución territorial de inversión, acceso a crédito, informalidad laboral, productividad, gasto público y estructura productiva.

2.11 Storytelling con datos

El storytelling con datos permite convertir una visualización en una narrativa útil para la toma de decisiones. En un contexto gerencial, los datos no deben presentarse como una acumulación de gráficos, sino como una historia orientada a la acción.

Para este caso, la historia que comunican los datos puede estructurarse de la siguiente manera:

Situación inicial.

Ecuador es un país con alto potencial turístico, pero también con retos de desigualdad de ingresos.

Pregunta analítica.

Se busca conocer si el turismo se relaciona con la evolución de la desigualdad.

Proceso de datos.

Se integraron datos turísticos de Kaggle y datos del Índice de Gini de Datosmacro, construyendo una base con país, año, turismo, Índice de Gini y tipo de desigualdad.

Hallazgo visual.

La visualización muestra una relación inversa: en los periodos donde el turismo aumenta, el Índice de Gini tiende a disminuir.

Interpretación gerencial.

El turismo puede representar una oportunidad para generar empleo y dinamizar ingresos, siempre que exista una estrategia de inclusión territorial y productiva.

Advertencia analítica.

El patrón observado no debe interpretarse como causalidad absoluta. Para tomar decisiones de política pública o inversión se recomienda complementar el análisis con modelos estadísticos y variables adicionales.

Decisión sugerida.

Fortalecer estrategias de turismo inclusivo, turismo comunitario, capacitación, encadenamientos productivos y medición territorial del impacto económico.

2.12 Aplicación práctica empresarial y pública

El análisis tiene aplicación tanto para el sector público como para empresas privadas vinculadas al turismo.

2.12.1 Aplicación para el sector público

Para gobiernos nacionales y locales, la visualización puede apoyar decisiones como:

- Priorizar inversión turística en zonas con mayor pobreza o desigualdad.
- Diseñar programas de capacitación para comunidades receptoras.
- Promover turismo comunitario y rural.
- Medir si el crecimiento turístico genera beneficios distribuidos.
- Evaluar políticas de empleo turístico.
- Identificar territorios donde el turismo crece, pero la desigualdad no mejora.

Desde esta perspectiva, la visualización se convierte en una herramienta de monitoreo de política pública.

2.12.2 Aplicación para empresas turísticas

Para hoteles, agencias, restaurantes, operadores turísticos y plataformas digitales, el análisis permite comprender que la sostenibilidad social del turismo también puede ser una ventaja competitiva. Las empresas pueden usar estos resultados para:

- Diseñar cadenas de suministro con proveedores locales.
- Contratar talento de comunidades cercanas.
- Crear productos turísticos inclusivos.
- Desarrollar experiencias comunitarias.
- Medir impacto social de sus operaciones.
- Alinear su estrategia con criterios ESG.

2.12.3 Aplicación para inversionistas

Para inversionistas, el análisis permite identificar que el turismo no solo debe evaluarse por rentabilidad, sino también por impacto territorial y social. Una inversión turística puede ser más sostenible si genera empleo local, diversifica ingresos y reduce tensiones sociales.

2.13 Limitaciones del análisis

El análisis presenta varias limitaciones que deben reconocerse:

- **Primera limitación: número de variables.**
El análisis se centra en turismo e Índice de Gini. No incorpora variables como empleo, educación, informalidad, inflación, inversión pública, pobreza, salario promedio o gasto social.
- **Segunda limitación: causalidad.**
La relación visual no demuestra causalidad. Solo permite observar asociación.
- **Tercera limitación: nivel de agregación.**
El análisis se realiza a nivel nacional. Esto puede ocultar diferencias territoriales entre provincias, ciudades o zonas turísticas.
- **Cuarta limitación: definición de turismo.**
El documento no detalla si la variable turismo representa llegadas de turistas, ingresos turísticos, impacto económico u otro indicador. Esto debe aclararse en una versión final del estudio.
- **Quinta limitación: fuente secundaria.**
Los datos provienen de bases secundarias. Por tanto, dependen de la metodología y actualización de las fuentes originales.

Reconocer estas limitaciones fortalece la calidad académica del capítulo, porque evita conclusiones excesivas y plantea oportunidades de mejora.

2.14 Recomendaciones analíticas

A partir del análisis, se proponen las siguientes recomendaciones:

1. Complementar la visualización con un análisis de correlación entre turismo e Índice de Gini.
2. Incorporar variables de empleo turístico, ingreso laboral, pobreza y actividad económica regional.
3. Realizar análisis por provincias o ciudades turísticas para identificar diferencias territoriales.
4. Separar turismo nacional e internacional si los datos lo permiten.
5. Comparar Ecuador con otros países latinoamericanos para evaluar si el patrón se repite.

6. Diseñar dashboards interactivos con filtros por año, región, tipo de turismo y nivel de desigualdad.
7. Evitar afirmar causalidad sin modelos econométricos o análisis longitudinal más profundo.
8. Usar la visualización como herramienta de decisión para políticas de turismo inclusivo.

2.15 Indicadores propuestos para seguimiento

Para convertir el análisis en una herramienta gerencial, se proponen los siguientes indicadores:

Dimensión	Indicador	Propósito
Turismo	Variación anual del turismo	Medir crecimiento o reducción de la actividad turística
Desigualdad	Índice de Gini	Medir distribución del ingreso
Relación analítica	Correlación turismo-Gini	Evaluar dirección e intensidad de la relación
Inclusión	Empleo turístico	Medir generación de oportunidades laborales
Territorio	Turismo por provincia	Identificar zonas beneficiadas
Impacto social	Variación de pobreza en zonas turísticas	Evaluar efecto social territorial
Gestión pública	Inversión turística local	Medir apoyo institucional
Empresas	Proveedores locales contratados	Evaluar encadenamiento productivo

Estos indicadores permitirían pasar de una visualización académica a un sistema de seguimiento útil para toma de decisiones.

Conclusión del capítulo

El capítulo de **Procesamiento y Visualización de Datos** permitió transformar dos fuentes de información independientes —datos turísticos e Índice de Gini— en una base integrada para analizar la relación entre turismo y desigualdad de ingresos en Ecuador. A partir del procesamiento de datos, se estructuraron variables clave como país, año, turismo, Índice de Gini y tipo de desigualdad, lo que permitió generar una visualización orientada a responder una pregunta social y económica relevante.

El hallazgo principal sugiere una relación inversa entre turismo y desigualdad: cuando el turismo aumenta, la desigualdad de ingresos tiende a disminuir; cuando el turismo se reduce, la desigualdad tiende a incrementarse. Esta conclusión plantea que el turismo podría contribuir a generar empleo y oportunidades económicas favorables para una mejor distribución del ingreso. No obstante, desde una perspectiva académica rigurosa, este resultado debe interpretarse como una asociación visual y no como una demostración definitiva de causalidad.

La importancia de este capítulo dentro de la tesis radica en que demuestra cómo el procesamiento y la visualización de datos pueden convertir información dispersa en conocimiento útil para decisiones estratégicas. En este caso, la visualización no solo permite observar una relación entre variables, sino que abre una discusión gerencial sobre turismo inclusivo, desarrollo territorial, encadenamientos productivos y reducción de desigualdad.

Perspectiva gerencial

Desde una perspectiva gerencial, este capítulo aporta valor porque muestra cómo la visualización de datos puede apoyar decisiones públicas y empresariales relacionadas con el turismo. Para los responsables de política pública, el análisis permite identificar si el turismo está asociado con mejoras en la distribución del ingreso y si conviene fortalecer programas de turismo comunitario, capacitación laboral, infraestructura local y promoción de destinos emergentes.

Para las empresas turísticas, el capítulo evidencia que el crecimiento del sector puede tener un impacto social positivo si se gestiona con enfoque inclusivo. Esto implica contratar talento local, trabajar con proveedores de la zona, impulsar emprendimientos comunitarios y diseñar experiencias turísticas que distribuyan beneficios más allá de los grandes operadores.

En términos estratégicos, la visualización de datos permite reducir la dependencia de la intuición y fortalecer la toma de decisiones basada en evidencia. Un gerente, ministro, alcalde, empresario turístico o inversionista puede utilizar este tipo de análisis para identificar oportunidades, evaluar riesgos y diseñar intervenciones con mayor impacto económico y social.

En síntesis, el capítulo demuestra que los datos no solo sirven para describir una realidad, sino para orientar decisiones que generen valor. En el caso de Ecuador, el turismo puede convertirse en una palanca de desarrollo más inclusiva si se gestiona con información confiable, visualizaciones claras y una lectura gerencial orientada a resultados.

Capítulo 3: Gestión y Diseño de Proyectos Big Data

Proyecto de analítica predictiva para campañas de marketing bancario

3.1 Introducción del capítulo

El presente capítulo desarrolla el módulo de Gestión y Diseño de Proyectos Big Data, aplicado al análisis de campañas de marketing bancario orientadas a predecir si un cliente aceptará o no un depósito a plazo. A diferencia de los capítulos anteriores, que se enfocan en el diseño de soluciones de datos y en la visualización de información, este capítulo aborda la gestión integral del proyecto analítico: metodología, planificación, riesgos, recursos, entregables, métricas de control y valor gerencial.

El punto de partida del proyecto es una problemática común en el sector financiero: las entidades bancarias realizan campañas de marketing directo para ofrecer productos como depósitos a plazo, pero dichas campañas pueden generar altos costos operativos y bajas tasas de conversión cuando no existe una segmentación adecuada. En el documento base del módulo se plantea que el objetivo es analizar la efectividad de campañas bancarias y desarrollar enfoques analíticos que permitan segmentar clientes y predecir la aceptación del producto financiero mediante modelos supervisados y no supervisados.

El proyecto utiliza el Bank Marketing Dataset, un conjunto de datos asociado a campañas de marketing directo de una institución bancaria portuguesa, cuyo objetivo de clasificación consiste en predecir si un cliente suscribirá o no un depósito a plazo. La UCI Machine Learning Repository describe este dataset como información relacionada con campañas telefónicas de marketing directo, donde el objetivo es predecir la suscripción de un depósito a plazo mediante una variable binaria de respuesta (“yes” o “no”) (Moro et al., 2014).

Desde una perspectiva de gestión de proyectos Big Data, este caso es relevante porque combina datos de clientes, variables sociodemográficas, historial financiero, condiciones de contacto, técnicas de preparación de datos, algoritmos de clasificación, segmentación no supervisada y métricas de evaluación. Además, exige traducir resultados técnicos en decisiones de negocio: a quién contactar, qué clientes priorizar, cómo reducir contactos improductivos, cómo mejorar la tasa de conversión y cómo optimizar recursos comerciales.

3.2 Planteamiento del problema

Las campañas bancarias tradicionales suelen operar con criterios generales de segmentación, como edad, nivel de ingresos, historial de contacto o productos contratados. Sin embargo, cuando estas variables se utilizan de forma aislada, la entidad puede contactar a un gran número de clientes con

baja probabilidad de aceptación, elevando costos, saturando canales comerciales y reduciendo la eficiencia de los equipos de ventas.

El documento base del módulo identifica que las campañas de marketing bancario implican costos operativos altos y bajas tasas de conversión cuando no están correctamente segmentadas. Por ello, propone usar analítica de datos para comprender el comportamiento de los clientes y priorizar a quienes tienen mayor probabilidad de aceptar el depósito a plazo.

El problema puede formularse de la siguiente manera:

¿Cómo gestionar y diseñar un proyecto Big Data que permita predecir la aceptación de depósitos a plazo en campañas de marketing bancario, optimizando recursos comerciales y mejorando la toma de decisiones mediante modelos analíticos?

Este problema combina una dimensión técnica y una dimensión gerencial. La dimensión técnica se relaciona con la selección del dataset, preparación de datos, tratamiento de valores faltantes, transformación de variables, selección de atributos, entrenamiento de modelos y evaluación de métricas. La dimensión gerencial se relaciona con la definición de objetivos de negocio, planificación del proyecto, asignación de responsabilidades, gestión de riesgos, monitoreo de KPIs y adopción de resultados por parte de las áreas comerciales.

3.3 Objetivo general del capítulo

Diseñar la gestión integral de un proyecto Big Data para campañas de marketing bancario, utilizando la metodología CRISP-DM y modelos de machine learning para predecir la aceptación de depósitos a plazo, optimizar la priorización de clientes y mejorar la eficiencia comercial.

3.4 Objetivos específicos

1. Definir la metodología de gestión del proyecto Big Data aplicada al caso de campañas bancarias.
2. Establecer los objetivos de negocio y de minería de datos del proyecto.
3. Diseñar la planificación general del proyecto, incluyendo fases, actividades, responsables y entregables.
4. Identificar riesgos técnicos, operativos, éticos y gerenciales asociados al uso de modelos predictivos en banca.
5. Definir KPIs para medir el desempeño del proyecto, los modelos y el impacto comercial.
6. Interpretar los resultados de los modelos desde una perspectiva gerencial y de toma de decisiones.

3.5 Metodología del proyecto: CRISP-DM

La metodología seleccionada para este proyecto es CRISP-DM (*Cross-Industry Standard Process for Data Mining*), debido a que permite estructurar proyectos de analítica desde la comprensión del negocio hasta la evaluación y posible despliegue de los resultados. El documento base del módulo indica que el proyecto aplicó CRISP-DM como marco estructurado para gestionar el análisis desde la comprensión del problema hasta la evaluación de modelos.

CRISP-DM organiza un proyecto de minería de datos en fases como comprensión del negocio, comprensión de los datos, preparación de datos, modelado, evaluación y despliegue. La guía original de CRISP-DM describe este proceso como un modelo jerárquico compuesto por fases, tareas genéricas, tareas especializadas e instancias de proceso, lo que permite adaptar la metodología a diferentes contextos empresariales (Chapman et al., 2000).

La elección de CRISP-DM es adecuada para este caso porque el objetivo no es únicamente construir modelos predictivos, sino resolver un problema empresarial concreto: mejorar la efectividad de campañas bancarias. En este sentido, la metodología ayuda a evitar que el análisis se convierta en un ejercicio técnico desconectado de las necesidades del negocio.

3.5.1 Fase 1: Comprensión del negocio

En esta fase se define el problema desde la perspectiva de la entidad financiera. El banco necesita mejorar la efectividad de sus campañas de marketing directo, reducir contactos innecesarios y aumentar la probabilidad de conversión en depósitos a plazo.

El documento base establece tres objetivos principales de negocio: incrementar la tasa de aceptación de depósitos entre un 5 % y un 8 %, reducir entre un 10 % y un 15 % los contactos innecesarios a clientes con baja probabilidad de aceptación, y lograr que al menos el 70 % de las campañas se apoyen en criterios analíticos derivados del modelo.

Desde el enfoque gerencial, esta fase permite alinear el proyecto con resultados medibles. No se trata de “hacer machine learning”, sino de mejorar la rentabilidad y eficiencia del proceso comercial.

3.5.2 Fase 2: Comprensión de los datos

La segunda fase consiste en conocer el dataset disponible, sus variables, estructura, calidad y limitaciones. El documento base indica que se seleccionó el Bank Marketing Dataset, el cual contiene características de clientes como edad, ocupación, estado civil, nivel educativo, historial financiero y resultado de campañas telefónicas.

La UCI Machine Learning Repository señala que el dataset contiene información de campañas telefónicas de una institución bancaria portuguesa y que el objetivo de clasificación es predecir si el cliente suscribirá un depósito a plazo. También identifica el conjunto como un problema de clasificación dentro del área de negocios, con variables categóricas e integer (Moro et al., 2014).

Esta fase es clave porque permite identificar si los datos son suficientes para responder el problema empresarial. Además, permite reconocer retos como desbalance de clases, valores faltantes, variables categóricas, necesidad de normalización y selección de atributos.

3.5.3 Fase 3: Preparación de datos

La preparación de datos fue una fase crítica del proyecto. Según el documento base, se cargó el dataset mediante el operador Read CSV, se verificó la correcta lectura del archivo, se confirmaron las variables categóricas y numéricas, y se definió claramente la variable objetivo deposit para el proceso de clasificación.

Posteriormente, se aplicaron transformaciones como:

- Reemplazo de valores faltantes mediante promedio en variables numéricas.
- Conversión de variables nominales a numéricas mediante *dummy coding*.
- Normalización de variables numéricas.
- Selección de variables mediante *Information Gain*.
- Reducción de dimensionalidad mediante *Select by Weights*.

Desde la gestión de proyectos, esta fase debe controlarse cuidadosamente porque la calidad del modelo depende directamente de la calidad de los datos preparados. Un modelo entrenado con variables mal codificadas, datos inconsistentes o desbalance no tratado puede generar decisiones comerciales equivocadas.

3.5.4 Fase 4: Modelado

La fase de modelado consistió en comparar algoritmos supervisados y no supervisados. En la clasificación supervisada, el objetivo fue predecir si un cliente aceptaría el depósito a plazo. Para ello se evaluaron modelos como W-ZeroR, W-OneR, Naive Bayes, Random Forest, Árbol de Decisión y W-MultilayerPerceptron.

En la clasificación no supervisada, se aplicó K-Medoids con $K = 5$ para segmentar registros a partir de atributos seleccionados como *contact*, *month*, *pdays*, *poutcome* y variables numéricas como

balance. Esta segmentación permitió identificar grupos de clientes con patrones distintos de recencia y comportamiento financiero.

Desde una perspectiva gerencial, la combinación de modelos supervisados y no supervisados es valiosa porque permite responder dos preguntas distintas: quién tiene mayor probabilidad de aceptar el producto y qué grupos de clientes comparten características similares.

3.5.5 Fase 5: Evaluación

La evaluación del proyecto comparó modelos mediante métricas como accuracy, precision, recall y AUC. La guía CRISP-DM plantea que la evaluación no debe limitarse a revisar la precisión del modelo, sino que debe valorar si los resultados cumplen los objetivos de negocio definidos inicialmente (Chapman et al., 2000).

En el proyecto, el Árbol de Decisión se destacó como el modelo con mejor desempeño general, alcanzando una accuracy de 78,79 % y un AUC de 0,829, superando a Naive Bayes, Random Forest y la red neuronal en el contexto específico analizado.

La evaluación también mostró que no basta con observar el accuracy. Para campañas comerciales, el comportamiento sobre la clase “yes” es especialmente relevante, porque representa a los clientes que sí aceptarían el depósito y, por tanto, justifican el contacto comercial. El documento base destaca que precision y recall deben analizarse especialmente en la clase de interés para interpretar los resultados desde una perspectiva empresarial.

3.5.6 Fase 6: Despliegue propuesto

Aunque el documento base se concentra en el análisis y evaluación, desde una visión de gestión de proyectos Big Data se propone una fase de despliegue gradual. Esta fase debería integrar el modelo seleccionado, preferentemente el Árbol de Decisión por su equilibrio entre desempeño e interpretabilidad, en el proceso real de gestión de campañas.

El despliegue no debe automatizar decisiones de manera absoluta. En banca, el uso de modelos predictivos debe acompañarse de validación humana, revisión de sesgos, criterios de transparencia y controles de seguridad. El propio documento base advierte que una aplicación real debe garantizar confidencialidad, seguridad y evitar conclusiones automáticas que puedan derivar en decisiones injustas o erróneas para determinados perfiles de clientes.

3.6 Alcance del proyecto

El alcance del proyecto incluye el diseño, entrenamiento, comparación y evaluación de modelos predictivos y de segmentación para campañas de marketing bancario. El foco principal es transformar datos históricos de campañas en información útil para mejorar la priorización de clientes.

3.6.1 Dentro del alcance

- Selección y carga del Bank Marketing Dataset.
- Validación inicial de variables y estructura del dataset.
- Tratamiento de valores faltantes.
- Conversión de variables categóricas.
- Normalización de variables numéricas.
- Selección de atributos mediante Information Gain.
- Entrenamiento de modelos supervisados.
- Aplicación de clustering con K-Medoids.
- Comparación de métricas de desempeño.
- Interpretación gerencial de resultados.
- Propuesta de KPIs y riesgos del proyecto.

3.6.2 Fuera del alcance

- Implementación real en un CRM bancario productivo.
- Automatización completa de campañas.
- Uso de datos personales reales identificables.
- Evaluación con datos recientes fuera del dataset.
- Cálculo financiero detallado de ROI, que corresponde al capítulo financiero.
- Auditoría regulatoria completa del modelo.

Esta delimitación permite que el proyecto sea viable dentro del contexto académico, pero con potencial de aplicación empresarial.

3.7 Planificación del proyecto

La planificación del proyecto se organiza en fases alineadas con CRISP-DM. A continuación, se presenta una propuesta estructurada para convertir el trabajo académico en un proyecto gestionable.

Fase	Actividades principales	Entregables	Responsable principal
1. Comprensión del negocio	Definir problema, objetivos y criterios de éxito y alcance	Acta del proyecto y objetivos de negocio	Líder del proyecto / área comercial
2. Comprensión de datos	Revisar dataset, variables y calidad y distribución de clases	Informe exploratorio de datos	Analista de datos
3. Preparación de datos	Limpiar, transformar, codificar, normalizar y seleccionar atributos	Dataset preparado	Data analyst / data engineer
4. Modelado	Entrenar modelos supervisados y no supervisados	Modelos entrenados	Científico de datos
5. Evaluación	Comparar accuracy, precision, recall y AUC	Matriz comparativa de modelos	Científico de datos / negocio
6. Interpretación gerencial	Traducir resultados técnicos en decisiones comerciales	Informe gerencial	Líder de analítica
7. Despliegue piloto	Integrar modelo en proceso controlado de campaña	Piloto de priorización de clientes	Comercial / TI / analítica
8. Monitoreo	Medir conversión, contactos evitados y desempeño del modelo	Dashboard de seguimiento	de Data Office

3.8 Recursos necesarios

3.8.1 Recursos humanos

Rol	Función
Sponsor ejecutivo	Asegura apoyo institucional y priorización del proyecto
Líder de proyecto	Coordina alcance, cronograma, riesgos y entregables
Representante comercial	Define criterios de campaña y necesidades del negocio
Analista de datos	Prepara, explora y valida la información
Científico de datos	Entrena, compara e interpreta modelos
Especialista TI	Evalúa integración técnica con sistemas internos
Oficial de riesgos/compliance	Revisa privacidad, ética y uso responsable de datos
Usuario de negocio	Valida utilidad práctica de los resultados

3.8.2 Recursos tecnológicos

Los recursos tecnológicos mínimos son:

- Herramienta de analítica o minería de datos, como RapidMiner.
- Repositorio de datos en formato CSV.
- Herramienta de visualización de métricas.
- Ambiente de pruebas para modelos.
- Control de versiones de datasets y experimentos.
- Documentación de transformaciones aplicadas.

3.8.3 Recursos de gestión

- Cronograma del proyecto.
- Matriz de riesgos.
- Matriz RACI.
- Criterios de éxito.
- Registro de decisiones.
- Informe de resultados.
- Plan de despliegue piloto.

3.9 Matriz RACI del proyecto

Actividad	Sponsor	Líder proyecto	Comercial	Analista datos	Científico datos	TI	Riesgos/Compliance
Definir objetivos de negocio	A	R	C	I	I	I	C
Seleccionar dataset	I	A	C	R	C	I	C
Preparar datos	I	C	I	R	C	I	C
Entrenar modelos	I	C	I	C	R	I	C
Evaluar métricas	I	A	C	C	R	I	C

Actividad	Sponsor	Líder proyecto	Comercial	Analista datos	Científico datos	TI	Riesgos/Compliance
Interpretar impacto comercial	C	A	R	C	C	I	C
Validar riesgos éticos	I	A	C	C	C	I	R
Diseñar piloto	A	R	R	C	C	C	C
Monitorear resultados	I	A	R	C	R	C	C

R = Responsable de ejecutar.

A = Responsable final.

C = Consultado.

I = Informado.

3.10 Gestión de riesgos

La gestión de riesgos es fundamental en proyectos Big Data, especialmente cuando se trabaja con datos bancarios y modelos predictivos que pueden influir en decisiones comerciales. A continuación, se identifican los principales riesgos del proyecto.

3.10.1 Riesgos técnicos

- **Riesgo de baja calidad de datos.**

Datos incompletos, mal codificados o inconsistentes pueden afectar el desempeño del modelo. En el proyecto se mitigó parcialmente mediante reemplazo de valores faltantes, conversión de variables nominales, normalización y selección de atributos.

- **Riesgo de desbalance de clases.**

El documento base señala que el dataset presenta desbalance entre clientes que aceptan y no aceptan el producto, lo que representa un reto analítico. Este riesgo puede llevar a modelos con buen accuracy general, pero baja capacidad para detectar la clase de interés.

- **Riesgo de sobreajuste.**

Modelos más complejos pueden ajustarse demasiado al dataset de entrenamiento y perder capacidad de generalización. Por ello se recomienda validar con datos no vistos y revisar métricas por clase.

- **Riesgo de interpretación incompleta de métricas.**

Usar solo accuracy puede llevar a decisiones incorrectas. Para campañas bancarias,

precisión y recall de la clase “yes” son críticos porque representan clientes con mayor probabilidad de conversión.

3.10.2 Riesgos operativos

- **Riesgo de baja adopción comercial.**
El equipo comercial puede no confiar en el modelo si no comprende sus reglas o si el resultado no se integra al flujo real de campaña.
- **Riesgo de contacto excesivo a clientes.**
Si el modelo prioriza mal, se puede saturar a ciertos clientes, afectando la experiencia y la percepción de la marca.
- **Riesgo de integración tecnológica.**
Pasar de un modelo académico a un sistema real requiere integración con CRM, canales de contacto, bases internas y sistemas de seguimiento.

3.10.3 Riesgos éticos y regulatorios

- **Riesgo de sesgo algorítmico.**
El modelo puede favorecer o excluir ciertos perfiles si las variables reflejan patrones históricos sesgados.
- **Riesgo de privacidad.**
En un contexto real, los datos bancarios son sensibles y deben protegerse con controles de acceso, anonimización y políticas de uso.
- **Riesgo de automatización injusta.**
Las decisiones no deben depender exclusivamente del modelo. El documento base recomienda control humano, revisión de sesgos y criterios de transparencia en aplicaciones reales.

3.10.4 Matriz de riesgos

Riesgo	Probabilidad	Impacto	Nivel	Mitigación
Datos incompletos o inconsistentes	Media	Alto	Alto	Validaciones, limpieza y reglas de calidad
Desbalance de clases	Alta	Alto	Alto	Métricas por clase, balanceo y ajuste de umbrales
Baja adopción del modelo	Media	Alto	Alto	Capacitación, interpretabilidad y pilotos controlados

Riesgo	Probabilidad	Impacto	Nivel	Mitigación
Sobreajuste del modelo	Media	Medio	Medio	Validación cruzada y prueba con datos nuevos
Sesgo algorítmico	Media	Alto	Alto	Revisión ética, monitoreo y variables justificadas
Privacidad de datos	Media	Alto	Alto	Anonimización, accesos restringidos y cumplimiento normativo
Mala interpretación de métricas	Alta	Medio	Alto	Capacitación gerencial y explicación de precision/recall
Falta de integración al CRM	Media	Medio	Medio	Piloto técnico gradual

3.11 KPIs del proyecto

Los KPIs deben medir tres niveles: desempeño del proyecto, desempeño del modelo e impacto comercial.

3.11.1 KPIs de gestión del proyecto

KPI	Fórmula	Meta sugerida
Cumplimiento del cronograma	$\frac{\text{Actividades completadas}}{\text{actividades planificadas}} \times 100$	$\geq 90 \%$
Cumplimiento de entregables	$\frac{\text{Entregables aprobados}}{\text{entregables planificados}} \times 100$	$\geq 95 \%$
Riesgos mitigados	$\frac{\text{Riesgos mitigados}}{\text{riesgos identificados}} \times 100$	$\geq 80 \%$
Participación de usuarios clave	$\frac{\text{Usuarios participantes}}{\text{usuarios convocados}} \times 100$	$\geq 80 \%$
Documentación completa	$\frac{\text{Documentos aprobados}}{\text{documentos requeridos}} \times 100$	$\geq 95 \%$

3.11.2 KPIs técnicos del modelo

KPI	Propósito	Meta sugerida
Accuracy	Medir clasificación global correcta	$\geq 75 \%$
Precision clase "yes"	Medir acierto al predecir clientes que aceptan	$\geq 75 \%$

KPI	Propósito	Meta sugerida
Recall clase “yes”	Medir capacidad de detectar clientes que aceptarían	$\geq 60 \%$
AUC	Medir capacidad de discriminación entre clases	$\geq 0,80$
Falsos negativos	Clientes que aceptarían, pero el modelo no prioriza	Reducir $\geq 10 \%$
Estabilidad del modelo	Comparar desempeño entre muestras	Variación controlada

Estas metas se alinean con los objetivos de Data Mining del documento base, donde se plantea alcanzar una precisión global mínima del 75 %, un recall para la clase “acepta” de al menos 60 % y reducir falsos negativos en al menos 10 %.

3.11.3 KPIs de impacto comercial

KPI	Fórmula	Meta sugerida
Tasa de aceptación	$\frac{\text{Clientes que aceptan}}{\text{clientes contactados}} \times 100$	+5 % a +8 %
Reducción de contactos innecesarios	$\frac{\text{Contactos evitados}}{\text{contactos base}} \times 100$	10 % a 15 %
Conversión por segmento	$\frac{\text{Aceptaciones por segmento}}{\text{contactos por segmento}} \times 100$	Incremento progresivo
Costo por conversión	$\frac{\text{Costo total campaña}}{\text{clientes convertidos}}$	Reducción progresiva
Uso analítico en campañas	$\frac{\text{Campañas basadas en modelo}}{\text{campañas totales}} \times 100$	$\geq 70 \%$

Estos objetivos están directamente relacionados con los resultados esperados del proyecto base: aumentar la aceptación, reducir contactos improductivos y mejorar la toma de decisiones comerciales mediante criterios analíticos.

3.12 Resultados del proyecto analítico

3.12.1 Preparación y selección de atributos

El proyecto aplicó técnicas de preparación de datos orientadas a mejorar la calidad y utilidad del dataset. Se reemplazaron valores faltantes, se transformaron variables nominales en numéricas, se normalizaron variables y se aplicó Weight by Information Gain para evaluar la relevancia de los atributos respecto a la variable objetivo. Los resultados indicaron que duration fue la variable con mayor peso informativo, seguida de outcome, contact, pdays y previous, lo que muestra que las

características relacionadas con la interacción y el historial de contacto son influyentes en la aceptación del producto.

Desde la perspectiva del negocio, este hallazgo es relevante porque demuestra que no todas las variables del cliente tienen el mismo peso. La forma, duración y resultado de contactos previos pueden ser más determinantes que variables generales cuando se busca predecir la aceptación de una oferta bancaria.

3.12.2 Comparación de modelos supervisados

La comparación de modelos permitió evaluar distintas alternativas de clasificación. El modelo W-ZeroR se utilizó como línea base y obtuvo un accuracy de 52,89 %, pero no resultó útil para la toma de decisiones porque no segmentaba adecuadamente. W-OneR alcanzó 74,10 % de accuracy y mostró buen equilibrio entre precision y recall, demostrando que reglas simples pueden aportar valor cuando son interpretables.

Naive Bayes obtuvo 70,14 % de accuracy y AUC de 0,804, mostrando capacidad aceptable de discriminación, aunque con menor recall para la clase “yes”. Random Forest alcanzó 69,42 % de accuracy, con precision para “yes” de 74,83 % y recall de 63,57 %, pero no superó al Árbol de Decisión ni a W-OneR en esta corrida.

El Árbol de Decisión fue el modelo más sólido, con accuracy de 78,79 % y AUC de 0,829. Además, presentó mejores valores de precision y recall, especialmente en la identificación de clientes que aceptarían el depósito.

La red neuronal W-MultilayerPerceptron alcanzó 70,50 % de accuracy, con precision para “yes” de 81,19 %, pero recall de 57,55 %, lo que implica que cuando predice “yes” suele acertar, aunque deja fuera una parte relevante de clientes que sí aceptarían.

Tabla 1.*Modelos supervisados*

accuracy: 63.08% +/- 1.48% (micro average: 63.08%)			
	true yes	true no	class precision
pred yes	2140	546	79.67%
pred no	3149	4165	56.95%
class recall	40.40%	88.41%	

**PerformanceVector**

```

PerformanceVector:
accuracy: 70.14%
ConfusionMatrix:
True: yes no
yes: 3122 819
no: 2167 3892
precision: 64.24% (positive class: no)
ConfusionMatrix:
True: yes no
yes: 3122 819
no: 2167 3892
recall: 82.62% (positive class: no)
ConfusionMatrix:
True: yes no
yes: 3122 819
no: 2167 3892
AUC (optimistic): 0.804 (positive class: no)
AUC: 0.804 (positive class: no)
AUC (pessimistic): 0.804 (positive class: no)

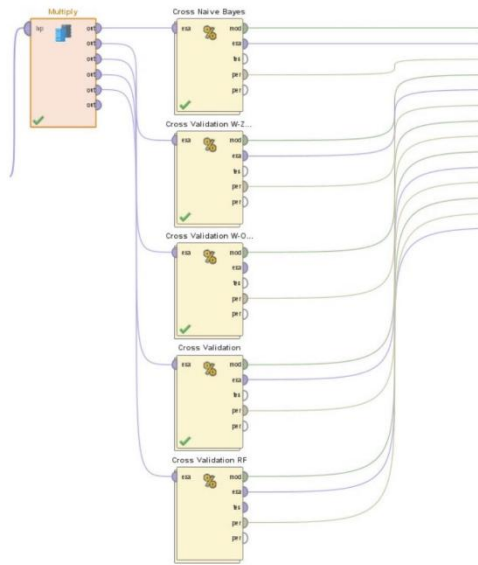
```

3.12.3 Clasificación no supervisada

En la fase no supervisada se aplicó K-Medoids con $K = 5$ para segmentar los registros. El análisis identificó un cluster dominante y varios clusters menores, con solapamiento entre grupos, pero con utilidad para describir tipos de clientes según recencia de contacto y comportamiento financiero.

Desde una perspectiva gerencial, el clustering permite complementar la predicción individual con estrategias de segmentación. Por ejemplo, un banco podría diseñar mensajes diferenciados, frecuencias de contacto y canales específicos según el cluster al que pertenezca cada cliente.

Tabla 2.*Modelo no supervisado*



3.13 Interpretación gerencial de resultados

El principal aprendizaje del proyecto es que un modelo predictivo no debe evaluarse únicamente por su desempeño técnico, sino por su utilidad para mejorar decisiones comerciales. En campañas bancarias, el objetivo no es clasificar correctamente por clasificar, sino priorizar clientes con mayor probabilidad de aceptar, reducir costos y mejorar la experiencia del cliente.

El Árbol de Decisión aparece como la alternativa más adecuada porque combina buen desempeño e interpretabilidad. En un entorno bancario, la interpretabilidad es especialmente importante porque los equipos comerciales y de cumplimiento deben comprender por qué un cliente fue priorizado o no. Modelos más complejos, como redes neuronales, pueden tener valor técnico, pero pueden ser menos adecuados si no ofrecen claridad suficiente para explicar decisiones.

El trabajo de este módulo concluye que los modelos no son un fin en sí mismos, sino herramientas para apoyar decisiones como priorización de clientes, reducción de contactos improductivos y mejora de la efectividad de campañas. Esta visión es esencial para la gestión de proyectos Big Data, porque el éxito se mide por el impacto en el negocio, no solo por la sofisticación del algoritmo.

3.14 Aplicación práctica empresarial

En una entidad bancaria real, el proyecto podría aplicarse mediante un piloto controlado en tres etapas.

3.14.1 Etapa 1: Piloto analítico

El modelo se usaría para clasificar clientes según probabilidad de aceptación. El equipo comercial recibiría una lista priorizada, pero la decisión final de contacto seguiría bajo supervisión humana.

3.14.2 Etapa 2: Campaña segmentada

Los clientes con mayor probabilidad de aceptación serían contactados primero. Los clientes con baja probabilidad podrían recibir mensajes menos costosos, como comunicación digital automatizada, en lugar de llamadas telefónicas.

3.14.3 Etapa 3: Monitoreo y mejora

Después de la campaña, se compararía la tasa de conversión real frente al escenario anterior. También se revisarían falsos positivos, falsos negativos, reclamos, costos y percepción del cliente.

Este enfoque permite implementar la analítica de manera responsable, gradual y medible.

3.15 Consideraciones éticas

El uso de machine learning en banca requiere especial cuidado. Aunque el proyecto académico no trabaja con datos personales identificables, una implementación real sí podría involucrar información sensible.

Las principales consideraciones éticas son:

- Proteger la confidencialidad de los clientes.
- Evitar discriminación por edad, estado civil, educación, ocupación u otras variables sensibles.
- Mantener supervisión humana sobre las decisiones comerciales.
- Explicar los criterios generales del modelo a los responsables de negocio.
- Revisar periódicamente sesgos y errores.
- No excluir clientes de oportunidades financieras sin justificación adecuada.
- Garantizar que el modelo sea usado para mejorar eficiencia, no para afectar derechos del cliente.

La ética no debe verse como una restricción externa, sino como una condición para que la analítica bancaria sea sostenible, confiable y aceptada por la organización.

3.16 Recomendaciones para la gestión del proyecto

A partir del análisis, se proponen las siguientes recomendaciones:

1. Implementar el Árbol de Decisión como modelo inicial por su equilibrio entre desempeño e interpretabilidad.
2. Complementar la predicción con clustering para diseñar campañas diferenciadas por segmento.
3. Medir el éxito comercial con tasa de aceptación, reducción de contactos innecesarios y costo por conversión.
4. No usar únicamente accuracy; evaluar precision, recall y AUC, especialmente para la clase “yes”.
5. Validar el modelo con datos recientes antes de llevarlo a producción.
6. Incorporar revisión ética y control humano antes de automatizar decisiones.
7. Capacitar al equipo comercial para interpretar los resultados del modelo.
8. Monitorear el desempeño después de cada campaña para detectar degradación o cambios en el comportamiento de clientes.
9. Documentar cada transformación de datos para asegurar trazabilidad.
10. Ajustar hiperparámetros y tratar el desbalance de clases en futuras iteraciones.

Estas recomendaciones se alinean con las conclusiones del documento base, que propone profundizar la implementación, evaluar desempeño con datos más recientes, explorar ajuste de hiperparámetros, tratar el desbalance de clases y complementar predicción con segmentación.

Conclusión del capítulo

El capítulo de Gestión y Diseño de Proyectos Big Data permitió estructurar el caso de campañas de marketing bancario como un proyecto analítico gestionable, orientado a resultados de negocio y no únicamente a resultados técnicos. A través de la metodología CRISP-DM, se organizó el proceso desde la comprensión del problema hasta la evaluación de modelos, considerando objetivos comerciales, preparación de datos, modelado, riesgos, KPIs, recursos y aplicación práctica.

El proyecto demuestra que la analítica predictiva puede mejorar la eficiencia de campañas bancarias al priorizar clientes con mayor probabilidad de aceptar depósitos a plazo. Los resultados muestran que el Árbol de Decisión fue el modelo con mejor desempeño general, al alcanzar una accuracy de 78,79 % y un AUC de 0,829, además de ofrecer interpretabilidad para el negocio.

Sin embargo, el valor del proyecto no está solo en seleccionar un modelo. Su mayor aporte está en conectar la analítica con decisiones gerenciales: reducir contactos innecesarios, mejorar conversión, optimizar recursos comerciales y apoyar campañas con evidencia. También se reconoce que una implementación real debe considerar riesgos de privacidad, sesgo, sobreajuste, baja adopción e integración tecnológica.

En síntesis, este capítulo evidencia que un proyecto Big Data bien gestionado requiere metodología, planificación, control de riesgos, métricas adecuadas y alineación con la estrategia empresarial. La tecnología genera valor cuando se convierte en una herramienta para decidir mejor.

Perspectiva gerencial

Desde la perspectiva gerencial, este capítulo aporta una visión práctica sobre cómo administrar proyectos Big Data en entornos bancarios. El modelo predictivo permite pasar de campañas masivas y poco diferenciadas a campañas segmentadas, priorizadas y medibles. Esto tiene impacto directo en eficiencia operativa, reducción de costos y mejora de resultados comerciales.

Para la gerencia comercial, el proyecto permite identificar clientes con mayor probabilidad de aceptación y diseñar estrategias de contacto más efectivas. Para la gerencia financiera, permite reducir costos de campaña y mejorar el retorno de los recursos invertidos. Para la dirección de riesgos, obliga a establecer controles éticos, privacidad y monitoreo de sesgos. Para la alta dirección, demuestra que la analítica puede convertirse en una capacidad estratégica para competir en un mercado financiero cada vez más orientado a datos.

El principal aporte gerencial es que el proyecto transforma datos históricos de campañas en conocimiento accionable. En lugar de depender únicamente de intuición o experiencia comercial, la entidad puede tomar decisiones basadas en evidencia, asignar mejor sus recursos y construir una cultura de gestión apoyada en datos.

Capítulo 4: Economía y Finanzas para la Gestión de Proyectos

Análisis financiero comparativo de Indurama Ecuador S.A., Mabe Ecuador S.A. y Fibroacero S.A.

4.1 Introducción del capítulo

El presente capítulo desarrolla el módulo de Economía y Finanzas para la Gestión de Proyectos, aplicado al análisis financiero comparativo de tres empresas del sector manufacturero ecuatoriano: Indurama Ecuador S.A., Mabe Ecuador S.A. y Fibroacero S.A. El propósito del capítulo es evaluar el desempeño económico-financiero de estas organizaciones a partir de indicadores clave como ventas, utilidad bruta, margen bruto, margen neto, liquidez y sostenibilidad financiera.

En el trabajo de este módulo se construyó a partir de información de estados financieros descargados de la Superintendencia de Compañías, Valores y Seguros del Ecuador, considerando el periodo 2021–2024. A partir de esta información se elaboró una base de datos financiera para analizar patrones de desempeño entre las tres empresas, principalmente en ventas, utilidad bruta y margen neto.

Desde una perspectiva gerencial, el análisis financiero permite comprender no solo cuánto vende una empresa, sino qué tan eficiente es para convertir sus ingresos en utilidad, qué tan sólida es su estructura de costos, qué riesgos enfrenta en términos de liquidez y qué tan sostenible puede ser su operación en el tiempo. En este sentido, el capítulo no se limita a describir cifras, sino que interpreta los resultados financieros como insumos para la toma de decisiones estratégicas.

La evaluación financiera de proyectos y empresas utiliza herramientas como ROI, VAN y TIR para estimar la conveniencia económica de una inversión. El VAN mide la diferencia entre el valor presente de los flujos de entrada y salida de efectivo, considerando el valor del dinero en el tiempo, mientras que la TIR representa la tasa de descuento que hace que el VAN sea igual a cero (Association for Financial Professionals [AFP], 2024). Estas herramientas son relevantes porque permiten evaluar si una decisión de inversión, mejora operativa o proyecto empresarial genera valor económico real.

4.2 Planteamiento del problema financiero

El sector manufacturero ecuatoriano enfrenta retos asociados con costos de producción, eficiencia operativa, liquidez, rentabilidad, competencia, variabilidad de la demanda y necesidad de inversión tecnológica. En este contexto, las empresas no pueden evaluarse únicamente por su volumen de ventas. Una compañía puede vender mucho, pero generar márgenes netos bajos si sus costos operativos, financieros o administrativos son elevados. De igual manera, una empresa puede tener menor escala, pero mayor eficiencia si logra convertir mejor sus ingresos en utilidad.

El documento base del módulo identifica tres realidades empresariales distintas. Indurama Ecuador S.A. se posiciona como la empresa con mayor volumen de ventas durante el periodo analizado, con cifras superiores a 230 millones y aproximadamente 260 millones en 2024. Sin embargo, presenta volatilidad en su margen neto, lo que evidencia desafíos en control de costos y eficiencia operativa. Mabe Ecuador S.A., por su parte, muestra ventas más inestables, pero mantiene márgenes netos más altos en varios periodos. Fibroacero S.A. presenta menor volumen de ventas, márgenes netos débiles e incluso valores negativos en 2022 y 2024, lo que evidencia problemas estructurales de rentabilidad.

El problema financiero del capítulo puede formularse de la siguiente manera:

¿Cómo evaluar el desempeño financiero comparativo de Indurama Ecuador S.A., Mabe Ecuador S.A. y Fibroacero S.A. para identificar fortalezas, debilidades, riesgos económicos y oportunidades de mejora en la gestión empresarial?

Esta pregunta es relevante porque permite conectar los indicadores financieros con decisiones gerenciales. Por ejemplo, una caída de margen neto puede exigir revisión de costos, gastos financieros o eficiencia operativa; una baja liquidez puede requerir ajustes en capital de trabajo; y una rentabilidad negativa puede alertar sobre la sostenibilidad del modelo de negocio.

4.3 Objetivo general del capítulo

Evaluar el desempeño económico-financiero de Indurama Ecuador S.A., Mabe Ecuador S.A. y Fibroacero S.A. durante el periodo 2021–2024, mediante el análisis de ventas, utilidad bruta, márgenes, liquidez y riesgos financieros, con el fin de generar información útil para la toma de decisiones gerenciales.

4.4 Objetivos específicos

1. Analizar la evolución de ventas, utilidad bruta y margen neto de las tres empresas estudiadas.
2. Comparar la eficiencia financiera de Indurama, Mabe y Fibroacero a partir de sus márgenes.
3. Evaluar la posición de liquidez de cada empresa y su capacidad para cubrir obligaciones de corto plazo.
4. Identificar riesgos financieros relacionados con rentabilidad, liquidez, costos y sostenibilidad operativa.
5. Proponer indicadores financieros para el seguimiento gerencial.
6. Relacionar los resultados financieros con criterios de evaluación de proyectos como ROI, VAN y TIR.

4.5 Soporte teórico financiero

La gestión financiera permite evaluar la capacidad de una empresa para generar ingresos, controlar costos, producir utilidad, administrar activos y cubrir obligaciones. En proyectos empresariales, esta evaluación se vuelve crítica porque toda iniciativa debe demostrar no solo viabilidad técnica, sino también conveniencia económica.

El ROI o retorno sobre la inversión mide el beneficio generado en relación con la inversión realizada. Es útil para comparar alternativas de inversión y evaluar si los recursos asignados producen retornos suficientes. La TIR y el VAN son herramientas ampliamente utilizadas para evaluar proyectos de inversión, especialmente cuando existen flujos de caja en distintos periodos. El VAN permite determinar si un proyecto crea valor al descontar los flujos futuros a una tasa que refleja riesgo, inflación, costo de oportunidad y costo de capital. La TIR, en cambio, representa la tasa de rendimiento esperada que iguala el valor presente de entradas y salidas de efectivo (AFP, 2024).

En términos gerenciales, el VAN es especialmente importante porque expresa la creación de valor en términos monetarios. Un VAN positivo indica que el proyecto genera valor por encima del costo de capital; un VAN negativo indica destrucción de valor. La TIR complementa el análisis porque permite comparar la rentabilidad esperada del proyecto con la tasa mínima requerida por la empresa. En proyectos de mejora operativa, estos indicadores ayudan a decidir si conviene invertir en automatización, expansión, reducción de costos, modernización tecnológica o fortalecimiento financiero.

Para el análisis de empresas, los márgenes financieros también son fundamentales. El margen bruto permite evaluar la relación entre ventas y costo de ventas; el margen operativo permite analizar la eficiencia de la operación antes de gastos financieros e impuestos; y el margen neto refleja la capacidad final de convertir ventas en utilidad. Una empresa con alto volumen de ventas pero bajo margen neto puede tener problemas de eficiencia, costos elevados o gastos financieros significativos.

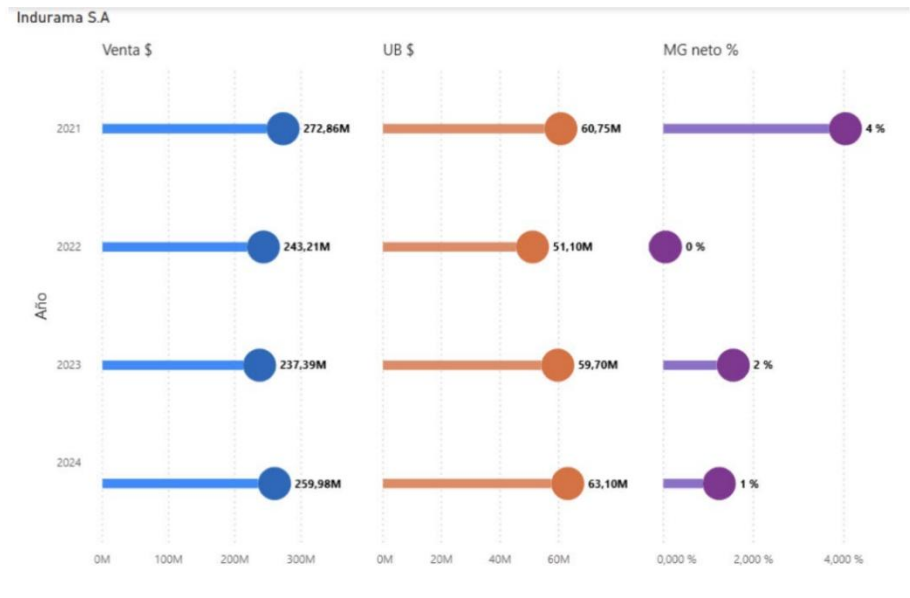
4.6 Descripción de las empresas analizadas

4.6.1 Indurama Ecuador S.A.

Indurama Ecuador S.A. es una empresa manufacturera con fuerte presencia en el mercado de línea blanca. En el análisis del periodo 2021–2024, se identifica como la empresa con mayor volumen de ventas entre las tres compañías estudiadas. El documento base señala que sus ventas se mantienen por encima de los 230 millones y alcanzan aproximadamente 260 millones en 2024.

Su principal fortaleza es la escala comercial. Sin embargo, el análisis también evidencia volatilidad en el margen neto, lo que sugiere que el liderazgo en ventas no necesariamente se traduce en rentabilidad proporcional. Esta situación puede estar asociada con presión en costos, gastos operativos, costos financieros, inversiones, variaciones de demanda o condiciones competitivas.

Tabla 3.
Estado financiero Indurama



4.6.2 Mabe Ecuador S.A.

Mabe Ecuador S.A. presenta un comportamiento más variable en términos de ventas. Según el documento base, registra una fuerte disminución en 2022, una recuperación en 2023 y una nueva caída en 2024. No obstante, destaca por mantener márgenes netos más altos en varios periodos, especialmente en 2021 y 2023.

Esto sugiere que Mabe, aunque no necesariamente lidera en escala frente a Indurama, muestra una mejor capacidad relativa para convertir sus ingresos en rentabilidad final. Desde la perspectiva gerencial, esta característica es relevante porque refleja mayor eficiencia en el control de costos o una estructura operativa más equilibrada.

Tabla 4.
Estado financiero Mabe

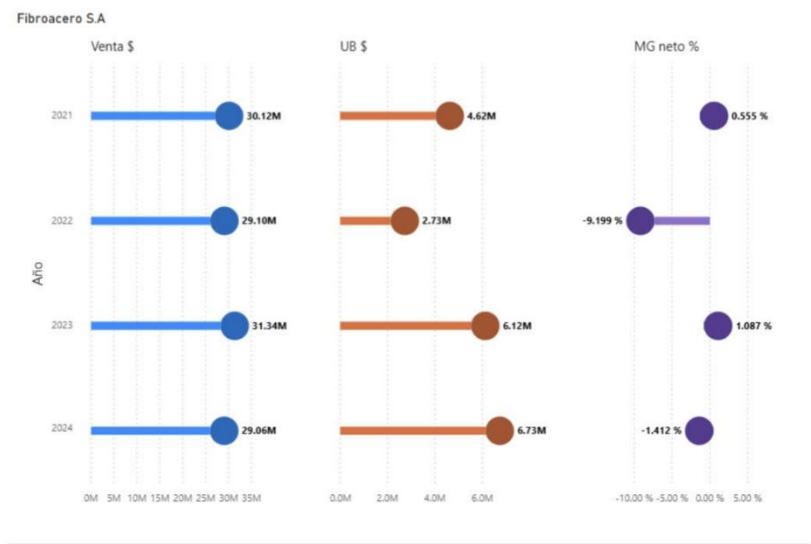


4.6.3 Fibroacero S.A.

Fibroacero S.A. presenta un volumen de ventas considerablemente menor, con valores cercanos a los 30 millones. A pesar de mostrar una evolución positiva en utilidad bruta entre 2022 y 2024, sus márgenes netos son los más débiles del grupo, incluso con valores negativos en 2022 y 2024.

Esta situación evidencia riesgos financieros relevantes. Una empresa puede sostener ventas y mejorar utilidad bruta, pero si el margen neto sigue siendo negativo, existen presiones en gastos operativos, financieros, administrativos o de estructura. Desde la gestión financiera, Fibroacero requiere atención estratégica para recuperar sostenibilidad.

Tabla 5.
Estado financiero Fibroacero



4.7 Análisis de ventas

Las ventas representan el punto de partida del análisis financiero porque muestran la capacidad comercial de una empresa para colocar productos en el mercado. Sin embargo, deben interpretarse junto con la rentabilidad, ya que vender más no siempre implica ganar más.

4.7.1 Indurama: liderazgo en escala

Indurama se posiciona como la empresa con mayor volumen de ventas durante todo el periodo analizado. Esta escala le otorga ventajas competitivas potenciales, como mayor poder de negociación con proveedores, capacidad de absorción de costos fijos, presencia de marca y cobertura de mercado. El documento base señala que sus ventas superan los 230 millones y llegan aproximadamente a 260 millones en 2024.

No obstante, el liderazgo en ventas debe analizarse con prudencia. La escala puede ser una fortaleza si se acompaña de rentabilidad, eficiencia y control de costos. Si el crecimiento en ventas no se traduce en mayor margen neto, la empresa podría estar enfrentando presión competitiva, costos crecientes o gastos que reducen el beneficio final.

4.7.2 Mabe: variabilidad comercial

Mabe muestra un comportamiento más inestable, con disminución en 2022, recuperación en 2023 y nueva caída en 2024. Esta variabilidad podría indicar sensibilidad ante condiciones de mercado, cambios en demanda, competencia, estrategias de precios o condiciones macroeconómicas.

Desde una perspectiva gerencial, la volatilidad en ventas obliga a revisar la estrategia comercial y la capacidad de planificación. Una empresa con ventas variables necesita mecanismos de control presupuestario, pronósticos de demanda y flexibilidad operativa para evitar que la caída en ingresos afecte de manera desproporcionada la rentabilidad.

4.7.3 Fibroacero: menor escala y presión competitiva

Fibroacero opera con una escala menor frente a Indurama y Mabe. Sus ventas cercanas a 30 millones reflejan una posición más limitada en volumen. Esta menor escala puede afectar su capacidad para distribuir costos fijos, negociar insumos y sostener inversiones.

Sin embargo, una empresa de menor tamaño puede competir mediante especialización, eficiencia, diferenciación de producto o nichos de mercado. El problema surge cuando la menor escala se combina con márgenes netos negativos, como ocurre en algunos años del periodo analizado.

4.8 Análisis de utilidad bruta

La utilidad bruta permite evaluar la diferencia entre ventas y costo de ventas. Este indicador es clave porque muestra si la empresa logra producir o comercializar sus bienes con una estructura de costos razonable.

4.8.1 Indurama: recuperación progresiva

Indurama presenta una recuperación progresiva de la utilidad bruta después de una caída en 2022, alcanzando su mejor desempeño en 2024. Esto sugiere una mejora en la capacidad de generar margen desde la operación productiva o comercial.

Desde una perspectiva gerencial, esta recuperación puede ser resultado de ajustes en precios, eficiencia productiva, control de costos de materiales, mejora en mix de productos o recuperación de demanda. Sin embargo, la utilidad bruta debe compararse con el margen neto para determinar si la mejora se sostiene después de gastos operativos y financieros.

4.8.2 Mabe: utilidad bruta dependiente del volumen

En Mabe, la utilidad bruta sigue una tendencia similar a las ventas, lo que evidencia dependencia directa del volumen comercial. Cuando las ventas caen, la utilidad bruta también se reduce; cuando las ventas se recuperan, la utilidad bruta mejora.

Esto indica que la empresa necesita sostener volumen para proteger su utilidad bruta. No obstante, su capacidad para mantener márgenes netos relativamente altos muestra que cuenta con una estructura financiera más eficiente que la de sus competidores en ciertos periodos.

4.8.3 Fibroacero: mejora bruta con debilidad neta

Fibroacero muestra evolución positiva en utilidad bruta entre 2022 y 2024. Este dato es favorable porque indica que la empresa puede estar mejorando su relación entre ventas y costos directos. Sin embargo, la persistencia de márgenes netos negativos revela que la mejora bruta no es suficiente para garantizar rentabilidad final.

Este hallazgo es importante porque permite identificar dónde se ubica el problema financiero. Si la utilidad bruta mejora pero el margen neto sigue débil, la atención debe dirigirse a gastos administrativos, gastos de ventas, costos financieros, depreciaciones, estructura de deuda o eficiencia operativa.

4.9 Análisis de márgenes financieros

Los márgenes permiten comparar empresas de distinto tamaño porque expresan la rentabilidad como porcentaje de las ventas. Esto es especialmente útil en este caso, ya que Indurama tiene una escala mucho mayor que Fibroacero.

4.9.1 Margen bruto

El margen bruto refleja la eficiencia en la producción o adquisición de bienes vendidos. Según el documento base, Indurama presenta el mayor margen bruto, lo que evidencia una sólida estructura de costos de producción.

Esta ventaja es relevante porque una empresa con margen bruto fuerte tiene mayor capacidad para absorber gastos operativos, invertir en innovación o enfrentar presiones de precios. Sin embargo, si el margen neto no acompaña, se debe revisar qué ocurre después del nivel bruto.

4.9.2 Margen neto

El margen neto mide la utilidad final generada por cada unidad monetaria de venta. En este indicador, Mabe destaca por mantener un mejor equilibrio entre margen bruto y margen neto, convirtiéndose en la empresa con mejor capacidad relativa para generar rentabilidad final.

Indurama, aunque lidera en ventas y margen bruto, presenta un margen neto relativamente bajo y volátil. Esto evidencia que la escala comercial no elimina la necesidad de controlar gastos, costos financieros y eficiencia administrativa.

Fibroacero presenta el margen neto más débil del grupo, incluso con valores negativos en 2022 y 2024. Esta situación representa una alerta porque una empresa con margen neto negativo destruye valor si no corrige sus costos, gastos o estructura financiera.

4.10 Análisis de liquidez

La liquidez mide la capacidad de una empresa para cubrir obligaciones de corto plazo con activos corrientes. Este análisis es fundamental porque una empresa puede ser rentable en el estado de resultados, pero enfrentar dificultades si no tiene suficiente efectivo o activos líquidos para cumplir compromisos inmediatos.

El documento base indica que Mabe posee los niveles más altos de activos corrientes en relación con sus obligaciones, lo que sugiere mayor capacidad para cubrir compromisos de corto plazo. Indurama mantiene una posición intermedia, mientras que Fibroacero presenta menores niveles de liquidez relativa, lo que podría implicar riesgos en su capacidad de pago inmediato.

4.10.1 Mabe: mayor holgura financiera

Mabe presenta la mejor posición relativa de liquidez. Esto le permite operar con mayor flexibilidad, negociar con proveedores, enfrentar variaciones de demanda y sostener compromisos inmediatos sin depender excesivamente de financiamiento externo.

Desde el punto de vista gerencial, una buena liquidez no solo reduce riesgo financiero, sino que también permite tomar decisiones con mayor margen de maniobra.

4.10.2 Indurama: posición intermedia

Indurama mantiene una posición intermedia de liquidez. Esto significa que, aunque no presenta la mayor holgura del grupo, tampoco se identifica como la empresa con mayor vulnerabilidad inmediata. Sin embargo, por su escala operativa, requiere una gestión rigurosa del capital de trabajo.

Una empresa de gran volumen necesita controlar inventarios, cuentas por cobrar, cuentas por pagar y flujo de caja operativo. Un desajuste en cualquiera de estos componentes puede generar presión financiera significativa.

4.10.3 Fibroacero: riesgo de liquidez

Fibroacero presenta menor liquidez relativa, lo que podría comprometer su capacidad de pago inmediato. Esta situación, combinada con márgenes netos negativos, aumenta el riesgo financiero.

Desde la gestión de proyectos, una empresa con baja liquidez debe priorizar inversiones de rápida recuperación, reducción de gastos improductivos, renegociación de deuda y mejora del ciclo de conversión de efectivo.

4.11 Evaluación financiera del proyecto empresarial

Aunque el documento base se enfoca en el análisis financiero comparativo de empresas, este capítulo incorpora criterios de evaluación de proyectos para traducir los hallazgos en decisiones gerenciales. En este contexto, se puede plantear que cada empresa necesita evaluar proyectos de mejora financiera orientados a fortalecer rentabilidad, liquidez y eficiencia.

4.11.1 ROI aplicado al caso

El ROI permite medir el rendimiento de una inversión frente al costo de implementarla. Para este caso, puede aplicarse a proyectos como:

- Automatización de procesos productivos.
- Reducción de desperdicios.
- Mejora de eficiencia energética.
- Optimización de inventarios.
- Renegociación logística.
- Implementación de analítica financiera.
- Reducción de costos administrativos.

La fórmula general es:

$$\text{ROI} = (\text{Beneficio neto del proyecto} / \text{Inversión total del proyecto}) \times 100$$

En Indurama, el ROI debería enfocarse en proyectos de eficiencia operativa que ayuden a convertir su alto volumen de ventas en mayor margen neto. En Mabe, debería orientarse a sostener

rentabilidad y estabilizar ventas. En Fibroacero, debería priorizar iniciativas de recuperación financiera y reducción de pérdidas.

4.11.2 VAN aplicado al caso

El VAN permite evaluar si un proyecto genera valor económico considerando el valor del dinero en el tiempo. Según AFP (2024), el VAN representa la diferencia entre el valor presente de los flujos de entrada y el valor presente de los flujos de salida de efectivo. Esta herramienta es útil para decidir si una inversión debe aprobarse, compararse o rechazarse.

La fórmula general es:

$$\text{VAN} = \sum [\text{Flujo de caja neto } t / (1 + r)^t] - \text{Inversión inicial}$$

Donde:

- t = periodo.
- r = tasa de descuento.
- Flujo de caja neto = entradas menos salidas de efectivo.
- Inversión inicial = desembolso requerido al inicio.

Aplicado al caso, una empresa debería aprobar proyectos con VAN positivo. Por ejemplo, si Fibroacero evalúa un proyecto de reducción de costos, este debería generar ahorros suficientes para superar la inversión inicial y el costo de capital.

4.11.3 TIR aplicada al caso

La TIR representa la tasa de rendimiento que hace que el VAN sea igual a cero. AFP (2024) explica que la TIR permite estimar la tasa de crecimiento esperada de una inversión y compararla con el costo de capital o tasa mínima requerida.

La regla general es:

- Si la TIR > costo de capital, el proyecto puede aceptarse.
- Si la TIR < costo de capital, el proyecto debería rechazarse.
- Si la TIR = costo de capital, el proyecto no genera valor adicional.

En este caso, Indurama, Mabe y Fibroacero podrían utilizar la TIR para comparar proyectos alternativos de mejora operativa. Sin embargo, debe complementarse con el VAN, porque la TIR puede generar interpretaciones incompletas en proyectos con flujos irregulares.

4.12 Costos del proyecto de análisis financiero

El proyecto de análisis financiero requiere considerar costos asociados a la construcción de una base de datos, análisis de estados financieros, visualización de resultados y elaboración de informes gerenciales.

4.12.1 Costos directos

Concepto	Descripción
Recolección de estados financieros	Descarga, organización y revisión de información financiera
Limpieza de datos	Homologación de cuentas, años, empresas y formatos
Herramientas de análisis	Uso de Excel, Power BI, Tableau u otra herramienta de visualización
Horas de analistas	Tiempo dedicado a procesamiento, interpretación y validación
Elaboración de informe	Redacción gerencial, tablas, gráficos y conclusiones

4.12.2 Costos indirectos

Concepto	Descripción
Tiempo de supervisión	Revisión por parte de docentes, directivos o responsables financieros
Curva de aprendizaje	Tiempo requerido para interpretar indicadores
Reprocesamiento	Ajustes por errores de datos o cambios en criterios
Validación	Comparación de resultados con fuentes originales

4.12.3 Costos de oportunidad

El costo de oportunidad se relaciona con el valor de la mejor alternativa no escogida. En este caso, dedicar tiempo y recursos al análisis financiero implica dejar de destinarlos a otras actividades. Sin embargo, el análisis se justifica porque proporciona información útil para decisiones estratégicas, especialmente cuando se identifican riesgos de rentabilidad y liquidez.

4.13 Impacto económico esperado

El impacto económico esperado del análisis financiero no proviene únicamente de elaborar gráficos o comparar cifras, sino de usar los resultados para mejorar decisiones empresariales. Los principales impactos esperados son:

4.13.1 Mejora en control de costos

Indurama presenta liderazgo en ventas, pero enfrenta desafíos en margen neto. Por ello, el análisis puede orientar decisiones de control de gastos, eficiencia productiva y revisión de estructura financiera.

4.13.2 Fortalecimiento de rentabilidad

Mabe muestra mejor equilibrio entre margen bruto y margen neto, lo que puede servir como referencia comparativa para identificar prácticas financieras más eficientes.

4.13.3 Recuperación financiera

Fibroacero requiere atención prioritaria debido a sus márgenes netos negativos y menor liquidez relativa. El impacto esperado del análisis es identificar áreas críticas para reducir pérdidas, mejorar capital de trabajo y proteger la continuidad del negocio.

4.13.4 Mejor toma de decisiones

El análisis permite a la gerencia comparar escala, rentabilidad y liquidez. Esta comparación ayuda a decidir dónde invertir, qué costos revisar, qué indicadores monitorear y qué riesgos priorizar.

4.14 Análisis de riesgos financieros

4.14.1 Riesgo de rentabilidad

El riesgo de rentabilidad aparece cuando la empresa no logra convertir ventas en utilidad neta suficiente. Este riesgo es evidente en Fibroacero, debido a sus márgenes netos negativos, y también en Indurama, por la volatilidad de su margen neto pese a su alto volumen de ventas.

4.14.2 Riesgo de liquidez

El riesgo de liquidez se relaciona con la capacidad de cubrir obligaciones de corto plazo. Mabe presenta mejor posición relativa, Indurama una posición intermedia y Fibroacero mayor vulnerabilidad.

4.14.3 Riesgo operativo

El riesgo operativo se asocia con ineficiencias en producción, costos elevados, fallas en procesos, desperdicios o baja productividad. En Indurama, el reto es convertir escala en rentabilidad; en Fibroacero, el reto es corregir debilidades estructurales que afectan el resultado neto.

4.14.4 Riesgo financiero

El riesgo financiero puede surgir por gastos de intereses, estructura de deuda, presión de financiamiento o dependencia de activos específicos para sostener liquidez. El documento base señala que la distribución de componentes de liquidez muestra alta concentración en determinadas cuentas, lo que puede indicar dependencia de ciertos activos para sostener la operación.

4.14.5 Riesgo estratégico

El riesgo estratégico aparece cuando la empresa no adapta su modelo de negocio a las condiciones competitivas. Mabe enfrenta volatilidad en ventas, Indurama presión sobre margen neto y Fibroacero sostenibilidad financiera limitada.

4.15 Matriz de riesgos financieros

Riesgo	Empresa más expuesta	Probabilidad	Impacto	Mitigación sugerida
Margen neto bajo o negativo	Fibroacero o Indurama	/ Alta	Alto	Revisión de costos, gastos y estructura financiera
Volatilidad de ventas	Mabe	Media	Medio-Alto	Planeación comercial, diversificación y forecast
Baja liquidez relativa	Fibroacero	Alta	Alto	Gestión de capital de trabajo y renegociación de deuda
Dependencia de activos corrientes específicos	Tres empresas	Media	Medio	Diversificar fuentes de liquidez
Costos operativos elevados	Indurama o Fibroacero	/ Media	Alto	Optimización de procesos y eficiencia productiva
Pérdida de sostenibilidad financiera	Fibroacero	Alta	Alto	Plan de recuperación financiera
Menor conversión de ventas en utilidad	Indurama	Media	Alto	Control presupuestario y análisis de rentabilidad por línea

4.16 KPIs financieros propuestos

Para dar seguimiento al desempeño financiero de las empresas, se proponen los siguientes KPIs:

Dimensión	KPI	Fórmula	Interpretación
Ventas	Crecimiento de ventas	$\frac{(\text{Ventas año actual} - \text{Ventas año anterior})}{\text{Ventas año anterior}} \times 100$	Mide expansión comercial
Rentabilidad bruta	Margen bruto	$\text{Utilidad bruta} / \text{Ventas} \times 100$	Evalúa eficiencia en costos directos
Rentabilidad neta	Margen neto	$\text{Utilidad neta} / \text{Ventas} \times 100$	Mide utilidad final por dólar vendido
Liquidez	Razón corriente	$\frac{\text{Activo corriente}}{\text{Pasivo corriente}}$	Evalúa capacidad de pago a corto plazo
Eficiencia	Gastos operativos sobre ventas	$\frac{\text{Gastos operativos}}{\text{Ventas}} \times 100$	Controla peso de la estructura operativa
Endeudamiento	Pasivo total sobre activo total	$\frac{\text{Pasivo total}}{\text{Activo total}} \times 100$	Mide dependencia financiera
Valor del proyecto	VAN	Valor presente flujos – inversión inicial	Evalúa creación de valor
Retorno	ROI	$\text{Beneficio neto} / \text{inversión} \times 100$	Mide rendimiento sobre inversión
Rentabilidad esperada	TIR	Tasa que hace VAN = 0	Compara rendimiento con costo de capital

4.17 Recomendaciones financieras por empresa

4.17.1 Recomendaciones para Indurama

Indurama debe aprovechar su liderazgo en ventas para mejorar su rentabilidad neta. Su prioridad no debería ser únicamente crecer en volumen, sino optimizar la conversión de ingresos en utilidad final.

Se recomienda:

- Analizar costos por línea de producto.
- Revisar gastos operativos y administrativos.
- Evaluar rentabilidad por canal.

- Fortalecer control presupuestario.
- Implementar proyectos de eficiencia productiva.
- Usar VAN y ROI para priorizar inversiones.
- Monitorear margen neto de forma mensual.

4.17.2 Recomendaciones para Mabe

Mabe presenta fortaleza en rentabilidad relativa, pero debe gestionar su volatilidad comercial. Su reto es sostener ventas sin perder eficiencia.

Se recomienda:

- Mejorar pronósticos de demanda.
- Diversificar estrategias comerciales.
- Mantener disciplina de costos.
- Proteger margen neto.
- Analizar causas de caída en ventas.
- Fortalecer capital de trabajo.
- Evaluar proyectos de expansión solo con VAN positivo.

4.17.3 Recomendaciones para Fibroacero

Fibroacero requiere un plan de recuperación financiera. Sus márgenes netos negativos y menor liquidez relativa representan señales de alerta.

Se recomienda:

- Reducir gastos no estratégicos.
- Revisar estructura de deuda.
- Mejorar rotación de inventarios.
- Renegociar condiciones con proveedores.
- Evaluar rentabilidad por producto.
- Priorizar proyectos de rápida recuperación.
- Diseñar un plan de flujo de caja mensual.
- Monitorear liquidez y margen neto como indicadores críticos.

4.18 Aplicación gerencial del análisis financiero

El análisis financiero comparativo permite a la gerencia tomar decisiones con mayor evidencia. Para Indurama, el foco debe estar en eficiencia y margen neto; para Mabe, en estabilidad comercial y sostenimiento de rentabilidad; para Fibroacero, en recuperación financiera y liquidez.

A nivel estratégico, los resultados permiten identificar que cada empresa enfrenta un desafío diferente:

- Indurama: liderazgo en escala, pero con reto de rentabilidad neta.
- Mabe: mejor eficiencia relativa, pero con ventas variables.
- Fibroacero: menor escala, debilidad en margen neto y riesgo de liquidez.

Estas diferencias muestran que no existe una única estrategia financiera aplicable a todas las empresas. Cada una requiere decisiones específicas según su estructura de ingresos, costos, liquidez y rentabilidad.

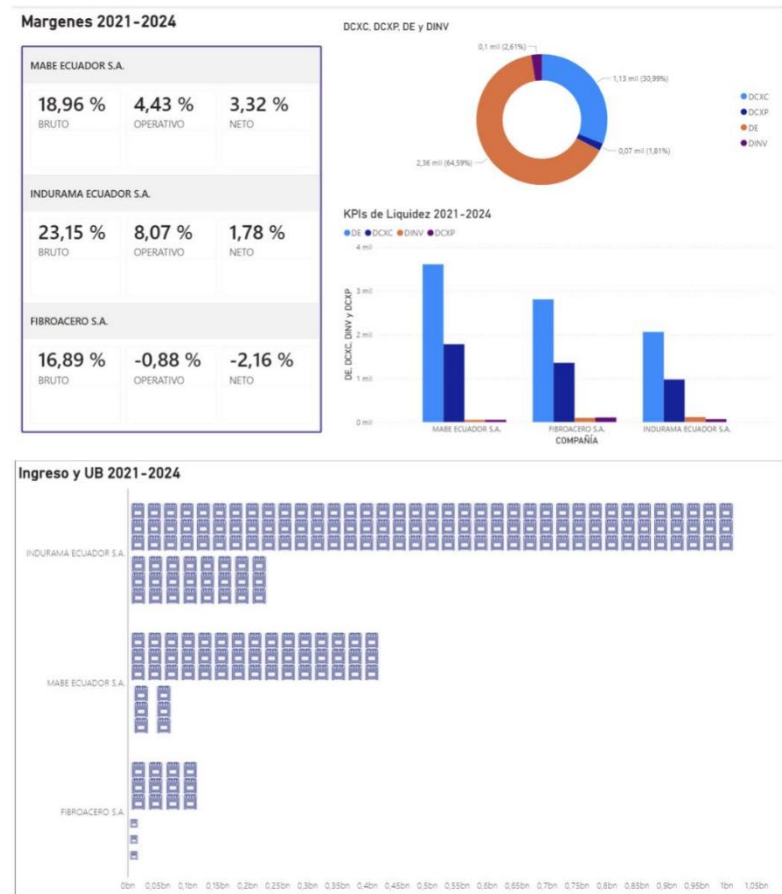
Conclusión del capítulo

El capítulo de Economía y Finanzas para la Gestión de Proyectos permitió analizar el desempeño financiero de Indurama Ecuador S.A., Mabe Ecuador S.A. y Fibroacero S.A. durante el periodo 2021–2024, a partir de variables como ventas, utilidad bruta, margen neto y liquidez. El análisis evidencia tres realidades empresariales distintas: Indurama lidera en escala de ventas, pero enfrenta retos en rentabilidad neta; Mabe muestra mejor capacidad relativa para generar utilidad final, aunque con ventas más variables; y Fibroacero presenta debilidades importantes en margen neto y liquidez que requieren atención estratégica.

La importancia del capítulo radica en que demuestra cómo los indicadores financieros permiten transformar estados contables en información gerencial. No basta con saber qué empresa vende más; es necesario comprender cuál convierte mejor sus ingresos en utilidad, cuál tiene mayor capacidad de pago y cuál enfrenta mayor riesgo financiero.

Desde la evaluación de proyectos, herramientas como ROI, VAN y TIR permiten priorizar inversiones orientadas a mejorar eficiencia, rentabilidad y sostenibilidad. En este sentido, el análisis financiero no es únicamente una revisión histórica, sino una base para decidir qué proyectos deben implementarse, cuáles deben postergarse y qué riesgos deben mitigarse.

Tabla 6.
Márgenes e ingresos 2021-2024



Perspectiva gerencial

Desde una perspectiva gerencial, este capítulo aporta directamente a la toma de decisiones estratégicas, eficiencia operativa y generación de valor. Para la alta dirección, el análisis permite identificar fortalezas y debilidades financieras de cada empresa. Para los gerentes financieros, ofrece una base para controlar márgenes, liquidez, costos y rentabilidad. Para los responsables de proyectos, permite justificar inversiones mediante criterios como VAN, TIR y ROI.

El principal aporte gerencial es que el análisis financiero revela que el tamaño de una empresa no garantiza rentabilidad. Indurama muestra liderazgo en ventas, pero necesita fortalecer su margen neto. Mabe demuestra que la eficiencia financiera puede compensar parcialmente la volatilidad comercial. Fibroacero evidencia que la rentabilidad negativa y baja liquidez pueden comprometer la sostenibilidad si no se ejecutan acciones correctivas.

En síntesis, la gestión financiera permite pasar de una lectura descriptiva de estados financieros a una gestión estratégica basada en evidencia. Esto ayuda a priorizar recursos, anticipar riesgos, mejorar eficiencia y asegurar que los proyectos empresariales generen valor económico real.

Capítulo 5: Competitividad y Transformación Digital

Proyecto de actualización SAP S/4HANA 1809 a SAP S/4HANA 2023 para una empresa manufacturera de Cuenca

5.1 Introducción del capítulo

El presente capítulo desarrolla el módulo de Competitividad y Transformación Digital, aplicado al proyecto de actualización de SAP S/4HANA versión 1809 a SAP S/4HANA 2023 en una empresa manufacturera de la ciudad de Cuenca. Este proyecto se analiza como una iniciativa estratégica de transformación digital, orientada a garantizar la continuidad operativa, reducir riesgos tecnológicos, fortalecer la competitividad empresarial y preparar a la organización para un entorno industrial cada vez más digitalizado.

El trabajo base del módulo plantea que la empresa enfrenta un riesgo crítico debido a que su ERP SAP S/4HANA 1809 se encuentra próximo a quedar sin soporte oficial, lo cual compromete la estabilidad, seguridad, integraciones clave y continuidad de procesos como finanzas, logística, producción, comercial y auditoría. Además, el documento identifica que operar con un core empresarial desactualizado genera presión sobre TI, riesgo operativo y exposición financiera por posibles costos adicionales de soporte.

Desde una perspectiva gerencial, este proyecto no debe interpretarse únicamente como una actualización técnica del ERP. Se trata de una decisión estratégica que impacta toda la cadena de valor de la organización. La actualización permite fortalecer procesos críticos como planificación de materiales, producción, compras, inventarios, ventas, facturación, finanzas, trazabilidad, integraciones y toma de decisiones. En este sentido, la transformación digital se convierte en un mecanismo para sostener la competitividad, reducir deuda tecnológica y mejorar la capacidad de respuesta de la empresa frente a cambios del mercado.

La metodología utilizada en el trabajo base se apoya en Design Thinking, a través de las etapas de empatizar, definir, idear, prototipar y testear. Este enfoque es pertinente porque permite comprender las necesidades reales de las áreas usuarias, definir el problema central, generar alternativas de solución, construir prototipos mediante ambientes de prueba y validar la solución antes del Go-Live. La Stanford d.school reconoce estas cinco etapas como modos fundamentales del pensamiento de diseño: empatizar, definir, idear, prototipar y testear (Stanford d.school, s. f.).

5.2 Planteamiento del problema

La empresa manufacturera analizada depende de SAP como sistema central para ejecutar procesos críticos. Finanzas lo utiliza para cierres contables y fiscales; logística y supply chain dependen del ERP para compras, inventarios, pedidos y trazabilidad; comercial lo necesita para facturación,

pedidos y entregas; producción lo usa para órdenes de fabricación, listas de materiales, rutas, consumos y reportes; y TI debe sostener la disponibilidad e integridad del sistema. El trabajo resume este hallazgo señalando que la organización opera con un core empresarial en “cuenta regresiva”, con dependencia absoluta y sin margen para fallas.

El problema central puede formularse de la siguiente manera:

La empresa enfrenta un riesgo crítico de continuidad operativa y pérdida de competitividad debido a la obsolescencia de su ERP SAP S/4HANA 1809, por lo que requiere ejecutar una actualización hacia SAP S/4HANA 2023 de forma ágil, controlada y con mínimo impacto en la operación.

Esta situación genera varios riesgos:

- Riesgo de interrupción operativa en procesos críticos.
- Riesgo de vulnerabilidades de seguridad.
- Riesgo de incompatibilidad con integraciones externas.
- Riesgo de costos adicionales por soporte extendido.
- Riesgo de pérdida de eficiencia en producción, logística y finanzas.
- Riesgo de baja capacidad de innovación tecnológica.
- Riesgo de incumplimiento frente a exigencias de auditoría, control interno y protección de datos.

Desde el enfoque de transformación digital, la actualización del ERP representa una oportunidad para rediseñar procesos, depurar desarrollos innecesarios, fortalecer datos maestros, mejorar controles, modernizar la experiencia del usuario y construir una base tecnológica más estable para futuras capacidades analíticas.

5.3 Objetivo general del capítulo

Analizar el impacto estratégico de la actualización SAP S/4HANA 1809 a SAP S/4HANA 2023 como proyecto de transformación digital, evaluando su aporte a la competitividad empresarial, continuidad operativa, cambio organizacional, cultura data-driven y posicionamiento estratégico de una empresa manufacturera de Cuenca.

5.4 Objetivos específicos

1. Evaluar cómo la actualización SAP S/4HANA contribuye a la ventaja competitiva de la empresa.
2. Analizar el impacto del proyecto en la cadena de valor organizacional.

3. Describir el proceso de cambio organizacional requerido para una transición exitosa.
4. Identificar cómo el proyecto fortalece la cultura data-driven y la toma de decisiones basada en información confiable.
5. Relacionar la actualización tecnológica con el posicionamiento estratégico de la empresa manufacturera.
6. Proponer indicadores de seguimiento para medir competitividad, adopción, estabilidad y valor generado.

5.5 Transformación digital como eje de competitividad

La transformación digital implica mucho más que incorporar tecnología. Supone modificar procesos, capacidades organizacionales, modelos de gestión y formas de tomar decisiones. En este proyecto, la actualización de SAP S/4HANA representa una transformación del core operativo de la organización, ya que afecta procesos productivos, financieros, logísticos, comerciales y tecnológicos.

La Agenda de Transformación Digital del Ecuador 2022–2025 plantea la transformación digital como un proceso transversal a todos los sectores, con líneas de acción relacionadas con infraestructura digital, economía digital, tecnologías emergentes, interoperabilidad, tratamiento de datos y seguridad digital (Ministerio de Telecomunicaciones y de la Sociedad de la Información [MINTEL], 2022). Esta visión respalda la importancia de que las empresas ecuatorianas modernicen sus sistemas, procesos y capacidades tecnológicas para sostener productividad y competitividad.

En el caso de una empresa manufacturera, el ERP es el núcleo que conecta materiales, producción, inventarios, ventas, finanzas y control. Cuando este núcleo opera sobre una versión desactualizada, la empresa no solo enfrenta un problema tecnológico, sino una limitación estratégica. Un ERP actualizado permite reducir errores, mejorar integración, optimizar flujos de información, acelerar procesos y generar datos más confiables para decisiones gerenciales.

SAP presenta la metodología SAP Activate como un marco de adopción que permite acelerar implementaciones de SAP S/4HANA mediante mejores prácticas, configuración guiada, procesos listos para usar y roadmaps estructurados (SAP, s. f.). Este enfoque es coherente con el proyecto, especialmente porque el documento base propone una actualización controlada, con priorización de procesos, ambientes de prueba, Fit-to-Standard y gestión del cambio.

Aplicación de Design Thinking al proyecto

El documento base estructura el proyecto bajo las etapas de Design Thinking: empatizar, definir, idear, prototipar y testear. Esta metodología resulta pertinente porque el problema no afecta únicamente a TI, sino a usuarios de finanzas, logística, producción, comercial, auditoría y dirección.

Empatizar

En la etapa de empatía se identificaron las preocupaciones de cada área:

Finanzas: riesgo en cierres contables y fiscales.

Logística / Supply Chain: riesgo en pedidos, compras, inventarios y trazabilidad.

Comercial: afectación en facturación, pedidos y entregas.

TI: presión por mantener operativo un sistema desactualizado.

Auditoría: riesgo crítico por operar con un core sin soporte.

Producción: dependencia total del ERP para órdenes, BOM, rutas, consumos y reportes.

Esta etapa permite comprender que la actualización SAP no es un requerimiento técnico aislado, sino una necesidad transversal de continuidad de negocio.

Definir

La etapa de definición permitió establecer el problema central: la empresa enfrenta un riesgo crítico de continuidad operativa porque su versión SAP S/4HANA 1809 quedará sin soporte oficial, comprometiendo estabilidad, seguridad e integraciones clave.

Esta definición es clave porque evita que el proyecto se entienda como una simple migración de software. El verdadero problema es la exposición operativa, financiera y competitiva que surge de operar con un core empresarial obsoleto.

Idear

En la etapa de ideación se propusieron alternativas como:

Programa Fast-Track con PMO estricta.

Reuniones diarias y war room.

Fit-to-Standard agresivo.

Depuración de desarrollos Z obsoletos.

Priorización por criticidad de procesos.

Migración Brownfield optimizada.

Ambientes Sandbox, QA y Preproducción.

Gestión del cambio y capacitación.

Estas ideas buscan reducir tiempo, riesgo y complejidad, evitando personalizaciones innecesarias y concentrando esfuerzos en procesos vitales para la operación.

Prototipar

El documento base utiliza el storytelling del “campamento base” para explicar el Sandbox de SAP 2023. Este ambiente permite probar una copia del sistema actual sin afectar la operación diaria. Allí se evalúan procesos, integraciones, desarrollos Z, rendimiento, datos maestros y experiencia de usuario.

La lógica del prototipo es reducir incertidumbre. Antes de ejecutar el Go-Live, la empresa puede identificar fallas, corregir configuraciones, validar integraciones y preparar a los usuarios.

Testear

La etapa de testeo incluye pruebas funcionales End-to-End, pruebas de integraciones críticas, pruebas de carga, pruebas de usuario final, validación de datos maestros, simulación de Día Cero y Día Uno, matriz de hallazgos y comité Go / No-Go.

Esta etapa es fundamental porque permite validar si la solución es operable antes de exponer a la organización a un cambio productivo.

Impacto en la ventaja competitiva

La ventaja competitiva de una empresa manufacturera depende de su capacidad para producir, entregar, facturar, controlar costos, responder al mercado y gestionar información de manera eficiente. En este contexto, el ERP actualizado fortalece la competitividad en varias dimensiones.

Continuidad operativa

La actualización reduce el riesgo de interrupciones en procesos críticos. Producción, compras, inventarios, ventas, facturación y finanzas dependen del ERP para operar diariamente. Si el sistema falla, la empresa puede enfrentar retrasos, errores, incumplimientos, costos adicionales y pérdida de confianza.

La continuidad operativa se convierte así en una ventaja competitiva porque permite cumplir con clientes, mantener eficiencia productiva y evitar paros que afecten ingresos.

Reducción de deuda tecnológica

La deuda tecnológica aparece cuando una organización mantiene sistemas, configuraciones o desarrollos obsoletos que dificultan la innovación. El documento base propone eliminar desarrollos Z innecesarios y aplicar un enfoque Fit-to-Standard agresivo, con el objetivo de reducir personalizaciones que complican el upgrade.

Esta decisión mejora la flexibilidad futura, disminuye costos de mantenimiento y facilita nuevas integraciones.

Eficiencia de procesos

SAP Activate promueve mejores prácticas y procesos estandarizados, lo que permite reducir complejidad y acelerar adopción (SAP, s. f.). En el caso del proyecto, esto se traduce en procesos más limpios para compras, producción, inventarios, ventas, finanzas y logística.

La eficiencia de procesos impacta directamente en costos, tiempos de respuesta, calidad de información y productividad del personal.

Mejor trazabilidad

La actualización fortalece la trazabilidad de datos, materiales, órdenes, inventarios, documentos contables y procesos logísticos. En manufactura, la trazabilidad es clave para controlar calidad, responder auditorías, gestionar reclamos y mejorar planificación.

Mayor capacidad de innovación

Una plataforma actualizada permite adoptar nuevas funcionalidades, integraciones, analítica avanzada, automatización y tecnologías emergentes. El documento base identifica que el proyecto habilita capacidades futuras como inteligencia de decisiones, fuerza de trabajo conectada aumentada y modelos de IA generativa específicos del dominio.

Impacto en la cadena de valor

El documento base aplica la Cadena de Valor de Porter para identificar dónde impacta el proyecto SAP. La conclusión principal es que la actualización no afecta solo a TI, sino a toda la cadena de valor de la empresa.

Logística interna

Esto permite evitar quiebres de inventario, duplicidades, errores de abastecimiento y retrasos en producción.

Operaciones

Producción depende del ERP para órdenes de fabricación, consumos, confirmaciones, rutas, BOM y reportes. La actualización fortalece la integración con terminales de planta, mejora la confiabilidad de datos maestros y reduce el riesgo de paros por fallas del sistema.

En términos competitivos, una operación más estable mejora cumplimiento, productividad y control de costos.

Logística externa

La logística externa incluye distribución, entregas y facturación. El proyecto fortalece el flujo de pedidos, la estabilidad de facturación electrónica y la reducción de errores por problemas de inventario.

Esto impacta directamente en satisfacción del cliente y cumplimiento comercial.

Marketing y ventas

El área comercial se beneficia de información más confiable sobre stock, pedidos, precios, promociones y disponibilidad. Esto permite responder mejor al mercado, reducir errores en pedidos y mejorar la experiencia del cliente.

Servicio postventa

El servicio postventa se beneficia de datos más consistentes sobre números de serie, historial de productos, garantías, reclamos y repuestos. La actualización puede mejorar tiempos de respuesta y calidad del servicio.

Actividades de apoyo

En tecnología, la actualización reduce deuda técnica y mejora seguridad. En talento humano, exige capacitación y adopción. En compras, permite procesos más limpios y mejores validaciones. En dirección estratégica, reduce riesgos de continuidad y costos futuros.

Cambio organizacional

La transformación digital no tiene éxito únicamente por la instalación de un sistema. Requiere gestión del cambio, comunicación, capacitación, liderazgo y apropiación de los usuarios. En este proyecto, el cambio organizacional es crítico porque SAP atraviesa múltiples áreas y cualquier resistencia puede afectar el Go-Live.

Resistencia al cambio

La resistencia puede surgir por temor a nuevas pantallas, cambios en procesos, pérdida de control, aumento de carga laboral durante pruebas o incertidumbre frente al nuevo sistema. El documento base reconoce la importancia de la gestión del cambio y capacitación para reducir resistencia y alinear a la organización.

Capacitación por roles

La capacitación debe ser diferenciada. Un usuario de planta no necesita el mismo entrenamiento que un analista contable o un usuario de compras. El documento base propone formación por roles, manuales cortos, red de key users y mesa de ayuda reforzada en hiper care.

Key users y champions

Los key users son esenciales porque funcionan como puente entre TI, consultores y usuarios finales. Ayudan a validar procesos, detectar errores, capacitar equipos y resolver dudas operativas.

Comunicación ejecutiva

La dirección debe comunicar que el proyecto no es opcional ni exclusivamente tecnológico. Debe presentarse como una iniciativa de continuidad, eficiencia y competitividad. Esta comunicación reduce incertidumbre y ayuda a priorizar recursos.

Hiper care

Después del Go-Live, el periodo de hiper care debe reforzar soporte, monitoreo, resolución de incidentes y acompañamiento a usuarios. Esto permite estabilizar la operación y evitar que pequeños problemas se conviertan en crisis.

Cultura data-driven

Una cultura data-driven implica que las decisiones se fundamenten en datos confiables, oportunos y trazables. En este proyecto, la actualización SAP fortalece esta cultura porque mejora la calidad del dato maestro, estandariza procesos y reduce inconsistencias.

Datos maestros confiables

El documento base señala que deben validarse BOMs, rutas, materiales, proveedores, unidades de medida, determinaciones de cuentas, condiciones de precio, tipos de documento, series y numeraciones.

Estos datos son la base para producción, inventarios, compras, ventas, costos y finanzas. Si los datos maestros son incorrectos, los reportes, decisiones y procesos también lo serán.

Indicadores de gestión

El proyecto define KPIs para monitorear estabilidad del sistema, integraciones, calidad de datos maestros, cumplimiento del cronograma, riesgos mitigados, Fit-to-Standard, hallazgos críticos, adopción de usuarios y disponibilidad durante el Go-Live.

Estos indicadores promueven una cultura de seguimiento basada en evidencia.

Mejora de información para decisiones

Con SAP actualizado, la empresa puede contar con información más consistente para analizar costos, inventarios, productividad, cumplimiento, trazabilidad, ventas y desempeño financiero.

Base para analítica avanzada

Una vez estabilizado el ERP, la empresa puede avanzar hacia analítica avanzada, inteligencia de decisiones, automatización de reportes, predicción de demanda, optimización de inventarios y modelos de IA aplicados a manufactura.

Posicionamiento estratégico

La actualización SAP S/4HANA 2023 fortalece el posicionamiento estratégico de la empresa manufacturera en varios niveles.

Posicionamiento operativo

La empresa puede operar con mayor estabilidad, menor riesgo de fallas y procesos más estandarizados. Esto mejora cumplimiento, productividad y confiabilidad ante clientes y proveedores.

Posicionamiento tecnológico

Actualizar el ERP posiciona a la empresa como una organización que reduce deuda técnica y se prepara para integrar nuevas capacidades digitales. Esto es relevante en un entorno manufacturero donde la competitividad depende cada vez más de datos, automatización e integración.

Posicionamiento frente a auditoría y cumplimiento

El documento base menciona que auditoría identifica riesgo crítico por operar con un core sin soporte. La actualización reduce esta exposición y mejora el control interno.

Además, la Ley Orgánica de Protección de Datos Personales del Ecuador establece obligaciones relacionadas con el tratamiento adecuado de datos personales y la protección de la información. En proyectos ERP, este marco es relevante porque el sistema gestiona datos de empleados, clientes, proveedores y transacciones. La existencia de una normativa específica fortalece la necesidad de sistemas seguros, controlados y gobernados (Asamblea Nacional del Ecuador, 2021).

Posicionamiento competitivo

Una empresa con procesos integrados, información confiable y sistemas actualizados puede responder mejor al mercado. Esto le permite competir no solo por producto o precio, sino por eficiencia, cumplimiento, trazabilidad y capacidad de adaptación.

Tecnologías emergentes vinculadas al proyecto

El documento base identifica tecnologías emergentes presentes y futuras en el proyecto. Entre las presentes se encuentran fuerza de trabajo conectada aumentada, inteligencia de decisiones y soberanía tecnológica parcial. Entre las futuras se mencionan modelos de IA generativa específicos del dominio.

Fuerza de trabajo conectada aumentada

La empresa ya opera con áreas conectadas al ERP: finanzas, producción, supply chain, comercial y TI. La actualización puede fortalecer esa conexión mediante datos integrados, flujos más estables y procesos más visibles.

Inteligencia de decisiones

SAP actualizado permite construir una base más sólida para decisiones basadas en reglas, indicadores y analítica. Aunque el proyecto no implementa inteligencia artificial de forma directa, sí prepara la arquitectura para decisiones más informadas.

Soberanía tecnológica parcial

La eliminación de desarrollos Z obsoletos y la adopción de estándares SAP reducen dependencia de personalizaciones difíciles de mantener. Esto mejora el control interno sobre la arquitectura tecnológica.

IA generativa futura

En una etapa posterior, la empresa podría explorar modelos generativos para documentación, soporte a usuarios, optimización de rutas de producción, análisis de BOMs, generación de reportes y asistencia en procesos.

KPIs del proyecto de transformación digital

Los KPIs permiten medir si la actualización SAP genera valor y reduce riesgos. El documento base propone indicadores asociados con estabilidad, integraciones, calidad de datos, gestión del proyecto, testing, adopción y Go-Live.

Dimensión	KPI	Propósito	Meta sugerida
Estabilidad	Tasa de errores en Sandbox, QA y Preproducción	Medir fallas técnicas antes del Go-Live	Tendencia decreciente
Procesos críticos	% de procesos críticos ejecutados sin fallas	Validar operación de producción, MRP, ventas, compras y finanzas	$\geq 95 \%$
Integraciones	% de integraciones operativas sin incidentes	Evaluar facturación electrónica, bancos, MES, e-commerce y sistemas externos	$\geq 95 \%$
Datos maestros	% de datos maestros validados	Garantizar calidad de BOMs, rutas, materiales y proveedores	$\geq 98 \%$
Cronograma	Cumplimiento del cronograma	Controlar avance del proyecto	$\geq 90 \%$

Dimensión	KPI	Propósito	Meta sugerida
Riesgos	% de riesgos mitigados	Medir gestión del tablero de riesgos	≥ 80 %
Fit-to-Standard	% de actividades completadas bajo estándar	Reducir personalizaciones innecesarias	≥ 85 %
Testing	Número de hallazgos críticos pendientes	Controlar bloqueos antes del Go-Live	0
Adopción	% de usuarios capacitados	Medir preparación de usuarios finales	≥ 95 %
Go-Live	Disponibilidad del sistema Día Cero y Día Uno	Medir continuidad operativa inicial	≥ 99 %
Hiper care	Número de incidentes post Go-Live	Medir estabilidad posterior	Tendencia decreciente

5.6 Riesgos estratégicos del proyecto

Riesgo de interrupción operativa

Si el Go-Live falla, la empresa puede enfrentar afectaciones en producción, ventas, facturación, compras y cierres financieros. Se mitiga con pruebas E2E, simulación Día Cero/Día Uno y comité Go / No-Go.

Riesgo de mala calidad de datos

Datos maestros incorrectos pueden afectar MRP, producción, costos, inventarios y facturación. Se mitiga con validación previa de BOMs, rutas, materiales, proveedores y cuentas contables.

Riesgo de resistencia del usuario

Usuarios no capacitados pueden cometer errores o rechazar el nuevo sistema. Se mitiga con capacitación por roles, key users, manuales cortos y soporte hiper care.

Riesgo de alcance excesivo

Intentar resolver todos los problemas durante el upgrade puede retrasar el proyecto. Se mitiga con Fit-to-Standard, priorización por criticidad y control de alcance.

Riesgo de integraciones fallidas

El ERP interactúa con sistemas externos como facturación electrónica, bancos, terminales de planta, MES o e-commerce. Se mitiga con pruebas específicas de integración y ambientes paralelos.

Riesgo financiero

Retrasos pueden generar costos adicionales de consultoría, soporte, licencias o dedicación interna. Se mitiga con PMO estricta, war room y seguimiento ejecutivo.

Matriz de riesgos

Riesgo	Probabilidad	Impacto	Nivel	Mitigación
Falla en Go-Live	Media	Alto	Alto	Simulación Día Cero/Día Uno y comité Go / No-Go
Datos maestros inconsistentes	Alta	Alto	Alto	Validación previa y responsables por dominio
Integraciones críticas fallidas	Media	Alto	Alto	Pruebas específicas por integración
Resistencia de usuarios	Media	Medio-Alto	Alto	Capacitación, key users y soporte hiper care
Retrasos del proyecto	Media	Alto	Alto	PMO estricta, war room y control semanal
Alcance excesivo	Alta	Medio	Alto	Fit-to-Standard y control de cambios
Sobrecostos	Media	Alto	Alto	Seguimiento financiero y priorización
Baja adopción post Go-Live	Media	Medio	Medio	Comunicación, acompañamiento y medición de uso

Aplicación práctica empresarial

La aplicación práctica del proyecto consiste en ejecutar una actualización SAP S/4HANA 2023 bajo un modelo Brownfield optimizado, tomando como base las configuraciones actuales de la compañía, pero depurando desarrollos innecesarios y priorizando procesos críticos.

Fase 1: Preparación y diagnóstico

Incluye levantamiento de procesos, inventario de desarrollos Z, revisión de integraciones, análisis de datos maestros y definición de criticidad.

Fase 2: Sandbox

Permite instalar SAP 2023 en un ambiente seguro, probar compatibilidad y detectar fallas tempranas.

Fase 3: QA y pruebas integrales

Incluye pruebas funcionales, regresión, integraciones, rendimiento, seguridad y UAT.

Fase 4: Preproducción

Simula condiciones reales, valida cargas, usuarios, tiempos de corte y plan de rollback.

Fase 5: Go-Live

Ejecuta la actualización productiva bajo monitoreo intensivo.

Fase 6: Hiper care

Refuerza soporte, resolución de incidentes, estabilización y transferencia a operación.

Recomendaciones estratégicas

1. Tratar la actualización SAP como un proyecto de negocio, no solo de TI.
2. Mantener patrocinio ejecutivo permanente.
3. Priorizar procesos vitales de producción, logística, ventas, finanzas y compras.
4. Aplicar Fit-to-Standard para reducir personalizaciones innecesarias.
5. Validar datos maestros antes de pruebas críticas.
6. Usar Sandbox, QA y Preproducción como ambientes obligatorios.
7. Capacitar usuarios por rol y no con entrenamientos genéricos.
8. Implementar una red de key users por área.
9. Mantener una matriz viva de riesgos y hallazgos.
10. No aprobar Go-Live con hallazgos críticos abiertos.
11. Medir adopción y estabilidad durante hiper care.
12. Documentar lecciones aprendidas para futuras iniciativas digitales.

Conclusión del capítulo

El capítulo de Competitividad y Transformación Digital demuestra que la actualización de SAP S/4HANA 1809 a SAP S/4HANA 2023 constituye una iniciativa estratégica para garantizar la continuidad operativa, reducir riesgos tecnológicos y fortalecer la competitividad de una empresa manufacturera de Cuenca. El proyecto no se limita a una mejora técnica del ERP, sino que impacta directamente en producción, logística, finanzas, ventas, compras, auditoría, TI y dirección estratégica.

El análisis evidencia que el ERP es el core operativo de la organización. Por ello, mantener una versión desactualizada expone a la empresa a fallas, costos adicionales, vulnerabilidades, pérdida de eficiencia y menor capacidad de innovación. En cambio, actualizar el sistema permite reducir

deuda tecnológica, mejorar datos maestros, fortalecer trazabilidad, estandarizar procesos y construir una base sólida para futuras capacidades digitales.

La aplicación de Design Thinking permitió comprender las necesidades de las áreas usuarias, definir el problema, idear soluciones, prototipar mediante Sandbox y validar la solución con pruebas integrales. Además, el enfoque Fit-to-Standard, la priorización por criticidad, la migración Brownfield optimizada y la gestión del cambio aparecen como elementos clave para reducir riesgos y asegurar adopción.

En síntesis, este capítulo muestra que la transformación digital genera valor cuando se conecta con continuidad operativa, eficiencia, datos confiables, cultura organizacional y posicionamiento competitivo.

Perspectiva gerencial

Desde una perspectiva gerencial, el proyecto de actualización SAP aporta directamente a la toma de decisiones estratégicas, la eficiencia operativa y la generación de valor. Para la alta dirección, reduce riesgos de continuidad y mejora el control del negocio. Para operaciones, fortalece producción, MRP, inventarios y trazabilidad. Para finanzas, mejora cierres, controles y confiabilidad contable. Para comercial, asegura facturación, pedidos y entregas. Para TI, reduce deuda técnica y mejora capacidad de soporte.

El principal aporte gerencial del capítulo es evidenciar que la competitividad no depende únicamente de vender más o producir más, sino de contar con una plataforma tecnológica estable, integrada y preparada para el futuro. Una empresa manufacturera que opera con un ERP actualizado puede responder con mayor rapidez, controlar mejor sus procesos, reducir errores, generar datos confiables y sostener decisiones basadas en evidencia.

Por tanto, la actualización SAP S/4HANA 2023 debe entenderse como una inversión estratégica en resiliencia, productividad y competitividad. Su éxito dependerá de la capacidad de la organización para gestionar el cambio, capacitar a los usuarios, controlar riesgos, validar procesos críticos y convertir la tecnología en una ventaja real para el negocio.

Bibliografía

Banco Mundial. (s. f.). *Índice de Gini (estimación del Banco Mundial)*. DataBank. <https://databank.worldbank.org/metadataglossary/jobs/series/SI.POV.GINI>

Chang, W. L., Grady, N., Boyd, D., & National Institute of Standards and Technology. (2019). *NIST Big Data Interoperability Framework: Volume 1, Definitions* (NIST Special Publication 1500-1r2). National Institute of Standards and Technology. <https://www.nist.gov/publications/nist-big-data-interoperability-framework-volume-1-definitions>

Chang, W. L., Grady, N., Boyd, D., & National Institute of Standards and Technology. (2019). *NIST Big Data Interoperability Framework: Volume 6, Reference Architecture* (NIST Special Publication 1500-6r2). National Institute of Standards and Technology. <https://www.nist.gov/publications/nist-big-data-interoperability-framework-volume-6-reference-architecture>

DAMA International. (2017). *DAMA-DMBOK: Data management body of knowledge* (2.^a ed.). Technics Publications.

Datosmacro. (s. f.). *Ecuador - Índice de Gini*. Expansión/Datosmacro. <https://datosmacro.expansion.com/demografia/indice-gini/ecuador>

Google Cloud. (s. f.-a). *Dataflow documentation*. <https://docs.cloud.google.com/dataflow/docs>

Google Cloud. (s. f.-b). *Managed Service for Apache Airflow documentation*. <https://docs.cloud.google.com/composer/docs>

Google Cloud. (s. f.-c). *Managed Service for Apache Spark documentation*. <https://docs.cloud.google.com/managed-spark/docs>

Google Cloud. (s. f.-d). *Pub/Sub documentation*. <https://docs.cloud.google.com/pubsub/docs>

Google Cloud. (s. f.-e). *What is MLOps?* <https://cloud.google.com/discover/what-is-mlops>

Kaggle. (s. f.). *Tourism and Economic Impact Dataset*. <https://www.kaggle.com/datasets/bushraqurban/tourism-and-economic-impact>

Marr, B. (2015). *Big data: Using SMART big data, analytics and metrics to make better decisions and improve performance*. John Wiley & Sons.

United Nations Conference on Trade and Development. (2017). *Economic development in Africa report 2017: Tourism for transformative and inclusive growth*. United Nations.
https://unctad.org/system/files/official-document/aldcafrica2017_en.pdf

World Bank. (2025). *Nature-based tourism*.
<https://www.worldbank.org/en/topic/environment/brief/nature-based-tourism>