UIDE
Powered by
Arizona State University®

Westfield
Business
School

eig

Maestría en

ADMINISTRACIÓN DE EMPRESAS

Tesis previa a la obtención de título de
Magister en Administración de Empresas

AUTORES:

Lcdo. Diego Sebastián Merino Bermeo

Ing. David Ricardo Márquez Coca

Lcdo. Danny Rodrigo Martínez Suárez

Lcda. Karol Mishel Cueva Muñoz

Ing. Edison Mauricio Torres Orellana

TUTOR: Mgtr. Carlos Gonzalez Luca de Tena

Sistema de Inteligencia Comercial para el Cumplimiento de Metas y Gestión de Clientes Potenciales en RM (Retail Moda)

REINVENTEMOS

PROYECTO DE TITULACIÓN – FIN DE MÁSTER
MAESTRÍA EN ADMINISTRACIÓN DE EMPRESA ENL

Título del trabajo de titulación

Sistema de Inteligencia Comercial para el Cumplimiento de Metas y Gestión de Clientes Potenciales en RM (Retail Moda)

Por

Nombres completos de los autores

Lcdo. Diego Sebastián Merino Bermeo

Ing. David Ricardo Márquez Coca

Lcdo. Danny Rodrigo Martínez Suárez

Lcda. Karol Mishel Cueva Muñoz

Ing. Edison Mauricio Torres Orellana

Octubre 2026

Aprobado



Cristian Melo
Presidente(a) del Tribunal
Universidad Internacional del Ecuador

Yo, Cristian Javier Melo González e Ignacio Maroto, declaramos que, personalmente conocemos que los graduandos: Lcdo. Diego Sebastián Merino Bermeo, Ing. David Ricardo Márquez Coca, Lcdo. Danny Rodrigo Martínez Suárez, Lcda. Karol Mishel Cueva Muñoz y Ing. Edison Mauricio Torres Orellana, son los autores exclusivos de la presente investigación y que ésta es original, auténtica y personal de ellos.



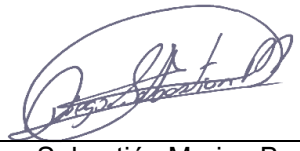
Cristian Melo
Coordinador MBA UIDE



Ignacio Maroto Provost WBS

Autoría del Trabajo de Titulación

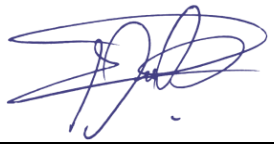
Nosotros, Lcdo. Diego Sebastián Merino Bermeo, Ing. David Ricardo Márquez Coca, Lcdo. Danny Rodrigo Martínez Suárez, Lcda. Karol Mishel Cueva Muñoz y Ing. Edison Mauricio Torres Orellana, declaramos bajo juramento que el trabajo de titulación titulado **Sistema de Inteligencia Comercial para el Cumplimiento de Metas y Gestión de Clientes Potenciales en RM (Retail Moda)** es de nuestra autoría y exclusiva responsabilidad legal y académica; que no ha sido presentado anteriormente para ningún grado o calificación profesional, habiéndose citado las fuentes correspondientes y respetando las disposiciones legales que protegen los derechos de autor vigentes.



Lcdo. Diego Sebastián Merino Bermeo
Correo electrónico:
diegomb.150499@gmail.com



Ing. David Ricardo Márquez Coca
Correo electrónico:
david.marquez.eng@gmail.com



Lcdo. Danny Rodrigo Martínez Suárez
Correo electrónico:
dannyrodrigo98@gmail.com



Lcda. Karol Mishel Cueva Muñoz
Correo electrónico:
lorak_14@hotmail.com



Ing. Edison Mauricio Torres Orellana
Correo electrónico:
edytho89@gmail.com

Autorización de Derechos de Propiedad Intelectual

Nosotros, Lcdo. Diego Sebastián Merino Bermeo, Ing. David Ricardo Márquez Coca, Lcdo. Danny Rodrigo Martínez Suárez, Lcda. Karol Mishel Cueva Muñoz y Ing. Edison Mauricio Torres Orellana, en calidad de autores del trabajo de investigación titulado Sistema de Inteligencia Comercial para el Cumplimiento de Metas y Gestión de Clientes Potenciales en RM (Retail Moda), autorizo a la Universidad Internacional del Ecuador (UIDE) para hacer uso de todos los contenidos que me pertenecen o de parte de los que contiene esta obra, con fines estrictamente académicos o de investigación. Los derechos que como autor me corresponden, lo establecido en los artículos 5, 6, 8, 19 y demás pertinentes de la Ley de Propiedad Intelectual y su Reglamento en Ecuador.

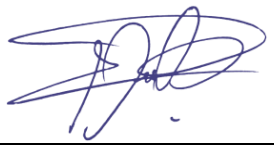
D. M. Quito, octubre de 2026



Lcdo. Diego Sebastián Merino Bermeo
Correo electrónico:
diegomb.150499@gmail.com



Ing. David Ricardo Márquez Coca
Correo electrónico:
david.marquez.eng@gmail.com



Lcdo. Danny Rodrigo Martínez Suárez
Correo electrónico:
dannyrodrigo98@gmail.com



Lcda. Karol Mishel Cueva Muñoz
Correo electrónico:
lorak_14@hotmail.com



Ing. Edison Mauricio Torres Orellana
Correo electrónico:
edytho89@gmail.com

Dedicatorias y Agradecimientos

Dedicatoria

A mi familia, por enseñarme con su ejemplo el valor del esfuerzo y la constancia, y por sostenerme en cada etapa de mi camino.

A mi pareja, por su paciencia, su amor y su fe en mí, incluso en los días en que yo dudaba.

Y a mi futuro hijo, que aún no llega pero que ya inspira cada paso que doy. Que este logro sea un recordatorio de que los sueños se alcanzan con perseverancia.

Agradecimiento

Quiero expresar mi más sincero agradecimiento a todas las personas que hicieron posible este trabajo.

A mis docentes, por compartir su conocimiento y experiencia, por exigirme y orientarme con generosidad, y por sembrar en mí el deseo de seguir aprendiendo.

A mis compañeros de posgrado, por las jornadas compartidas, los debates, las risas y el apoyo mutuo que hicieron más llevadero este camino.

A mis padres, a mi hermano y a mi pareja, por ser mi red de apoyo constante, y a mi futuro hijo, por ser la motivación que me impulsó a concluir esta etapa con esperanza.

Lcdo. Diego Sebastián Merino Bermeo

Dedicatoria

Dedico este logro a mi esposa, por su amor, paciencia y comprensión durante las largas horas de estudio y sacrificio. Tu apoyo fue el pilar fundamental que me mantuvo firme en cada etapa.

A mi familia, por ser mi motor constante, por comprender mis ausencias y por brindarme siempre el impulso y las palabras de aliento necesarias para no rendirme. Este triunfo es tan suyo como mío.

Agradecimiento

Agradezco profundamente a 3AComposites, por la confianza depositada en mi capacidad, el patrocinio de este programa de maestría y el respaldo institucional brindado para hacer posible mi desarrollo profesional.

A mi esposa y a mi familia, por su amor incondicional y por siempre mantenerme motivado.

A mis profesores, por guiarnos con su conocimiento y experiencia, y a mis compañeros, por las jornadas de debate, el trabajo en equipo y por enriquecer este proceso de aprendizaje continuo.

A todos los que formaron parte de este camino, mi más sincero reconocimiento.

Ing. David Ricardo Márquez Coca

Dedicatoria

Dedico este trabajo a mi familia, por acompañarme durante este proceso, por comprender mis momentos de dedicación y esfuerzo, y por brindarme siempre palabras de ánimo para seguir adelante.

Este logro representa no solo la culminación de una etapa académica, sino también el resultado del apoyo y la confianza de quienes han estado a mi lado en cada paso del camino.

Agradecimiento

Agradezco a mis padres y familia, por su amor, apoyo incondicional y por motivarme siempre a seguir adelante.

A mis profesores, por compartir sus conocimientos y experiencias, y a mis compañeros, por hacer de este proceso una experiencia enriquecedora de aprendizaje y crecimiento.

A todos quienes fueron parte de este camino, mi sincero agradecimiento por su apoyo y confianza.

Lcdo. Danny Rodrigo Martínez Suárez

Dedicatoria

Dedico este logro a mi hijo, mi mayor inspiración y la razón que me impulsa cada día a seguir creciendo y superándome. Que algún día pueda mirar este camino y entender que cada esfuerzo también fue una forma de enseñarle, con el ejemplo, que los sueños se construyen con perseverancia.

A mi familia, por su amor, paciencia y apoyo incondicional durante este proceso.

Este logro representa mucho más que una meta académica cumplida; representa esfuerzo, sacrificio, crecimiento y la decisión de nunca dejar de avanzar.

Agradecimiento

Agradezco a Dios por darme la fortaleza, sabiduría y perseverancia para llegar hasta aquí.

A mi hijo y a mi familia, gracias por ser mi motivación, mi apoyo y mi refugio en cada etapa de este camino.

A mis profesores, por compartir sus conocimientos y experiencias, y a mis compañeros, por cada aprendizaje y experiencia compartida.

Finalmente, agradezco a todas las personas que, de una u otra manera, fueron parte de este proceso. Hoy cierro esta etapa con gratitud y orgullo, sabiendo que cada esfuerzo valió la pena y que este logro es también el comienzo de nuevos sueños.

Lcda. Karol Mishel Cueva Muñoz

Dedicatoria

A mis padres, raíz de todo lo que soy.
A mis hermanos, compañeros de vida.
A mi familia, por creer en mí siempre.
Y a mí, por cada sacrificio que hizo posible esta meta.

Agradecimiento

Agradezco a mi familia por su paciencia, su apoyo y su amor a lo largo de este proceso; su compañía hizo más ligero cada esfuerzo. Me agradezco también a mí mismo, por la disciplina y la voluntad de terminar lo que empecé. Extiendo mi gratitud a la Universidad Internacional del Ecuador y a Westfield Business School, a los docentes que compartieron su conocimiento durante la maestría y a mis compañeros, con quienes el aprendizaje se convirtió en trabajo en equipo.

Ing. Edison Mauricio Torres Orellana

INDICE DEL DOCUMENTO

INTRODUCCIÓN Y OBJETO DEL PROYECTO 15
PARTE I: Diagnóstico y Diseño del Sistema de Inteligencia Comercial..... 17
CAPITULO 1 Contexto y Problema de Negocio 17
CAPITULO 2 Origen de los Datos y Procesos de Carga 21
CAPITULO 3 Modelo Dimensional y Calidad de Datos 26
CAPITULO 4 Visualización y Gobierno de la Información 30
CAPITULO 5 La Unidad de Inteligencia Comercial y Datos 35
PARTE II: Extensión Analítica y Capacidades Predictivas en Tiempo Real 41
CAPITULO 1 Incorporación de Datos No Estructurados 41
CAPITULO 2 Selección de la Arquitectura Cloud 46
CAPITULO 3 Arquitectura Alternativa Open Source..... 50
CAPITULO 4 Ciencia de Datos y Machine Learning 55
CAPITULO 5 Caso de Uso en Tiempo Real 60
CONCLUSIONES Y APLICACIONES..... 65
REFERENCIAS 70

INDICE DE FIGURAS Y TABLAS

Tabla 1. Estructura del trabajo: capítulos, pregunta que responde cada uno y resultado principal 16
Tabla 2. Problema de negocio de RM y sus efectos sobre la gestión comercial 18
Tabla 3. Áreas que demandan información en RM y uso actual de la misma 19
Tabla 4. Objetivos y necesidades de información por área 20
Tabla 5. Indicadores clave de desempeño propuestos para el Sistema de Inteligencia Comercial 21
Tabla 6. Fuentes internas y externas de información y su recurrencia de carga 22
Tabla 7. Problemas del proceso actual de carga de datos 23
Tabla 8. Sistemas origen, tipo de carga, destino y frecuencia de actualización 24
Tabla 9. Necesidades de información no cubiertas por los informes actuales 24
Tabla 10. Plan general de implantación por fases, aproximadamente 30 semanas 25
Tabla 11. Capas de la arquitectura de datos propuesta para RM 26
Tabla 12. Variables disponibles por tabla de hechos y dimensión del modelo dimensional preliminar 27
Tabla 13. Dimensiones de calidad de datos y hallazgos del diagnóstico de RM 28
Tabla 14. Recursos de TI internos y externos empleados en el proyecto 30
Tabla 15. Comparación de herramientas de visualización 31
Tabla 16. Tableros, usuarios, nivel de decisión y frecuencia de consulta 32
Tabla 17. Matriz de permisos propuesta por perfil 33
Tabla 18. Catálogo de indicadores, fórmulas y tablero de destino 34
Tabla 19. Revisión del cumplimiento de los objetivos iniciales 35
Tabla 20. Matriz GAP de competencias de RM, escala de 1 a 5 36
Tabla 21. Plantilla y costo anual estimado de la Unidad, con aportes patronales 37
Tabla 22. Matriz make or buy de la Unidad de Inteligencia Comercial y Datos 38

Tabla 23. Costes, beneficios y resultado del business case a tres años, escenario base 39
Tabla 24. Problema de negocio, señal faltante y decisión que apoya 42
Tabla 25. Inventario consolidado de fuentes estructuradas y no estructuradas 43
Tabla 26. Matriz de priorización de los casos de uso en escala de 1 a 5 44
Tabla 27. Criterios iniciales de aceptación de calidad del dato enriquecido 45
Tabla 28. Matriz de riesgos del tratamiento de datos no estructurados 46
Tabla 29. Estimación de volumen, crecimiento y usuarios, supuestos a validar 47
Tabla 30. Matriz de decisión ponderada de proveedores cloud, escala 1–5 48
Tabla 31. Costo anual estimado de la plataforma por capa, en USD 48
Tabla 32. Riesgos de la decisión tecnológica y su mitigación 49
Tabla 33. Modelos de servicio en la nube y su aplicación en RM 50
Tabla 34. Criterios de evaluación y pesos, fijados antes de puntuar 52
Tabla 35. Matriz comparativa por capa con puntuación ponderada y justificación de descarte 52
Tabla 36. Flujo de datos tramo por tramo de la arquitectura open source 53
Tabla 37. Costo total estimado de la plataforma open source por componente 54
Tabla 38. Matriz de riesgos de la arquitectura open source y su mitigación 54
Tabla 39. Matriz de priorización de decisiones candidatas, escala 1–5 56
Tabla 40. Ficha del modelo: definición, línea base, técnica y criterios de éxito 56
Tabla 41. Fuentes requeridas por el modelo, calidad y propiedad 57
Tabla 42. Departamentos implicados y matriz RACI (R: responsable; A: aprueba; C: consultado; I: informado) 58
Tabla 43. Beneficio anual por palanca, con fórmula explícita, en USD 59
Tabla 44. Matriz de priorización de casos de uso en tiempo real, escala 1–5 61
Tabla 45. Descomposición de la latencia y elección del nivel tecnológico 61
Tabla 46. Fuentes del caso de uso: flujos de eventos frente a datos de referencia 62
Figura 1. Arquitectura Lambda ligera para el caso de uso en tiempo real de RM: tramos de la vía rápida 62
Tabla 47. Departamentos y matriz RACI del caso de uso en tiempo real (R: responsable; A: aprueba; C: consultado; I: informado) 63

Tabla 48. Costes de las dos velocidades comparadas: micro-lotes frente a flujo continuo 63
Tabla 49. Business case a tres años de la alternativa seleccionada y escenarios, en USD 64
Tabla 50. Comparación de los tres business case cuantificados del proyecto 65
Tabla 51. Riesgos transversales del proyecto y su mitigación común 66
Tabla 52. Hoja de ruta integrada del proyecto, por fases 67
Tabla 53. Validaciones pendientes antes de ejecutar, por responsable y capítulo de origen 68

INTRODUCCIÓN Y OBJETO DEL PROYECTO

TIENDEC S.A. opera RM (Retail Moda), una cadena de 12 tiendas físicas y un canal de comercio electrónico dedicada a la venta de prendas de vestir y accesorios de moda. Cuando iniciamos el proyecto, la empresa gestionaba sus ventas, su inventario, sus clientes potenciales y el cumplimiento de sus metas comerciales con sistemas aislados y hojas de cálculo. La consolidación manual consumía entre tres y cuatro horas diarias, cerca del 8 % de los reportes mensuales contenía errores de digitación y la información llegaba a la Gerencia Comercial con uno a dos días de desfase (TIENDEC S.A., 2026).

Este trabajo de titulación recoge el proyecto que, como Grupo N.° 2, desarrollamos en la asignatura de Big Data y Business Intelligence de la Maestría en Administración de Empresas de Westfield Business School. Su resultado es un Sistema de Inteligencia Comercial que resuelve ese problema en capas sucesivas. Primero definimos el problema y los KPI; luego diseñamos el origen de los datos y el plan de implantación, así como el modelo dimensional del Data Warehouse, y construimos la capa de visualización en Power BI. A continuación, creamos y justificamos financieramente la Unidad de Inteligencia Comercial y Datos, que gobierna el sistema (Parte I). Más adelante extendimos el alcance a los datos no estructurados de origen digital, validamos formalmente la plataforma cloud frente a sus alternativas y diseñamos una arquitectura abierta equivalente. Por último, incorporamos una primera capacidad predictiva de aprendizaje automático para la reposición de inventario y una capacidad de procesamiento en tiempo real para detectar huecos en sala y roturas de curva de tallas (Parte II).

En este documento presentamos el proyecto como una narrativa única, organizada en dos partes de cinco capítulos cada una. Conservamos las cifras que sostienen cada decisión y omitimos el detalle metodológico que cada etapa del proyecto ya documentó por separado. El resultado del proyecto se resume en tres conclusiones transversales. La primera es que la arquitectura tecnológica (Azure como plataforma seleccionada, con una alternativa open source documentada) es sólida, proporcional al volumen real de RM y no exige sobredimensionamiento. La segunda, que el sistema solo entrega valor si se gobierna como una capacidad organizativa, con una unidad responsable de 4,5 personas, y no como un proyecto de implementación que concluye al encender los tableros. La tercera, que el caso de negocio es desigual entre capacidades. La reposición de inventario con aprendizaje automático y la detección de huecos en sala en tiempo real recuperan la inversión dentro de tres años, aunque con holguras muy distintas. En cambio, la Unidad de Inteligencia Comercial y Datos requiere un horizonte más largo, por tratarse de una capacidad habilitante y no de una aplicación analítica puntual. Esa heterogeneidad no se aprecia con la misma claridad cuando cada capacidad se analiza de forma aislada.

Todas las cifras que presentamos provienen de los análisis de cada etapa del proyecto y mantienen sus supuestos y advertencias. La facturación anual de RM

(USD 8.000.000) es hipotética, y así la declaramos desde el business case de la Unidad de Inteligencia Comercial y Datos. Los datos de volumen, calidad y costo del origen de los datos, del modelo dimensional y de la selección de la arquitectura cloud son estimaciones del equipo, pendientes de validación con TI y Finanzas, y los business case de la reposición de inventario y del caso de uso en tiempo real evitan de forma explícita contabilizar dos veces un mismo beneficio.

Desarrollamos el proyecto con la metodología de aprendizaje basado en proyectos, a lo largo de diez entregas semanales agrupadas en dos bloques. Cada capítulo de este trabajo recoge una de esas entregas: la Parte I corresponde al Bloque 1 y la Parte II, al Bloque 2. En las tablas empleamos una notación compacta en la que «I.3» designa el Capítulo 3 de la Parte I y «II.2», el Capítulo 2 de la Parte II. La Tabla 1 resume la estructura del trabajo.

Capítulo	Pregunta que responde	Resultado principal
I.1	¿Cuál es el problema y quién necesita qué información?	Diagnóstico inicial del problema comercial y diez KPI propuestos para el sistema.
I.2	¿De dónde provienen los datos y cómo se cargan hoy?	Inventario de fuentes internas y externas y plan de implantación en seis fases.
I.3	¿Cómo se estructura el almacenamiento y quién accede?	Arquitectura en tres capas (Landing Zone, ETL, Data Warehouse/Data Marts), modelo en estrella y control de acceso basado en roles.
I.4	¿Con qué herramienta se visualiza y para quién?	Selección de Power BI y diseño de cuatro tableros diferenciados por perfil de usuario.
I.5	¿Quién es responsable del sistema y se justifica la inversión?	Creación de la Unidad de Inteligencia Comercial y Datos (4,5 FTE) y su business case fundacional.
II.1	¿Cómo se incorporan los datos no estructurados?	Modelo de captura y gobierno de texto libre; dos casos de uso priorizados para el piloto.
II.2	¿La plataforma cloud elegida es la correcta?	Comparación formal de proveedores: Azure (4,55) frente a Google Cloud (4,20) y AWS (3,85).
II.3	¿Existe una alternativa sin dependencia de un proveedor?	Arquitectura equivalente con software abierto, documentada como alternativa portable.
II.4	¿Dónde aplicar la primera capacidad predictiva?	Modelo de aprendizaje automático para la reposición semanal de inventario por tienda y referencia.
II.5	¿Cómo procesar eventos a medida que ocurren?	Detección de huecos en sala y rotura de curva de tallas mediante arquitectura Lambda ligera, con micro-lotes de 15 minutos.

Tabla 1. *Estructura del trabajo: capítulos, pregunta que responde cada uno y resultado principal*

PARTE I:

Diagnóstico y Diseño del Sistema de Inteligencia Comercial

Esta primera parte reúne los cinco primeros capítulos del proyecto (Bloque 1), en los que definimos el problema de negocio, el origen y el almacenamiento de los datos, la capa de visualización y la unidad organizativa responsable del sistema.

CAPITULO 1 Contexto y Problema de Negocio

En este capítulo establecemos el marco de referencia sobre el que se construyó el Sistema de Inteligencia Comercial para RM (Retail Moda). Presentamos la empresa y delimitamos el problema de negocio y los objetivos iniciales del proyecto. Además, identificamos las áreas y personas que demandan información, describimos el uso que hoy hacen de ella y propusimos los indicadores clave de desempeño con los que se mediría la gestión comercial. El aporte central de esta primera etapa del proyecto fue mostrar que a la compañía no le faltan datos: le faltan un repositorio centralizado y tableros interactivos que los conviertan en decisiones oportunas y confiables.

1.1 Presentación de la empresa

TIENDEC S.A., que opera con la marca comercial RM (Retail Moda), es una empresa del sector retail dedicada a la comercialización de prendas de vestir y accesorios de moda mediante una red de tiendas físicas. En el desarrollo de sus actividades administra a diario un volumen importante de información sobre ventas, clientes, inventarios, metas comerciales y desempeño de los asesores de ventas. Sostuvimos entonces que esa información constituye un recurso estratégico solo en la medida en que pueda transformarse en conocimiento útil para la toma de decisiones (Sharda et al., 2023; TIENDEC S.A., 2026).

La empresa opera en un contexto marcado por la transformación digital y por una competencia creciente, que obliga a las compañías de retail a integrar la información generada por sus operaciones para mejorar la eficiencia de sus procesos y su competitividad (Laudon & Laudon, 2022). Cuando los datos permanecen dispersos en distintos sistemas o en registros manuales, resulta complejo obtener una visión integral del negocio y disminuye la capacidad de respuesta ante los cambios del mercado (Davenport, 2014). Sobre esa base, propusimos un sistema orientado a centralizar la información comercial, automatizar la generación de reportes y facilitar el seguimiento de indicadores estratégicos mediante herramientas de Business Intelligence. Con ello se buscaba apoyar la planificación comercial, el monitoreo del desempeño y la toma de decisiones basada en evidencia (Chen et al., 2012).

1.2 Problema de negocio y objetivos iniciales

Definimos el problema a partir de dos procesos críticos: el seguimiento del cumplimiento de metas comerciales y la gestión de clientes potenciales. En RM ambos dependen de información distribuida entre diferentes sistemas o administrada en hojas de cálculo, lo que dificulta realizar análisis oportunos que apoyen la gestión estratégica (Laudon & Laudon, 2022). La elaboración de reportes exige consolidaciones manuales que alargan los tiempos de procesamiento, generan duplicidad de información y restan confiabilidad a los datos que llegan a la gerencia (TIENDEC S.A., 2026). A ello se suma la ausencia de tableros interactivos, que impide monitorear los indicadores de forma continua y priva a la organización de oportunidades para optimizar su gestión comercial (Sharda et al., 2023).

Problema identificado	Efecto sobre la gestión comercial
Información dispersa en sistemas y hojas de cálculo	Dificultad para realizar análisis oportunos y obtener una visión integral del negocio
Consolidación manual de reportes	Mayores tiempos de procesamiento y duplicidad de información
Datos poco confiables para la gerencia	Menor confianza en las decisiones comerciales
Ausencia de dashboards interactivos	Limitado monitoreo de indicadores y del desempeño de los asesores
Detección tardía de desviaciones	Dificultad para identificar tendencias y responder a tiempo frente a las metas

Tabla 2. *Problema de negocio de RM y sus efectos sobre la gestión comercial*

A partir de esta problemática, fijamos seis objetivos iniciales para el proyecto, que orientaron las etapas siguientes del trabajo (Chen et al., 2012):

- Diseñar un Sistema de Inteligencia Comercial que centralice la información comercial.
- Automatizar la generación de reportes ejecutivos.
- Mejorar el seguimiento de clientes potenciales.
- Monitorear el cumplimiento de metas individuales y por tienda.
- Implementar tableros interactivos para visualizar indicadores comerciales.
- Fortalecer la toma de decisiones mediante herramientas de Business Intelligence.

1.3 Áreas que demandan información y uso actual

El proyecto involucra a cinco áreas funcionales que requieren información específica para cumplir sus objetivos y decidir en distintos niveles de la organización (Kotler & Keller, 2016). En todas ellas observamos un patrón común: la información existe, pero llega tarde, fragmentada o a costa de consolidaciones adicionales. En el caso de los asesores comerciales, la falta de

una plataforma integrada reduce, además, la trazabilidad del proceso comercial con cada cliente (Han et al., 2012).

Área	Información que demanda	Situación actual
Gerencia Comercial	Ventas, cumplimiento de metas, rentabilidad e indicadores estratégicos	Reportes elaborados manualmente; sin información actualizada en tiempo real
Jefes de tienda	Ventas diarias, productividad de asesores, cumplimiento de objetivos y comportamiento de clientes	Información distribuida en distintos registros; difícil comparación entre tiendas
Asesores comerciales	Clientes potenciales, oportunidades de venta y metas individuales	Sin plataforma integrada; seguimiento tardío y baja trazabilidad
Marketing	Segmentación, comportamiento de compra y resultados de campañas	Consolidaciones adicionales que alargan el análisis
Gerencia General	Indicadores ejecutivos del desempeño global	Sin dashboards ejecutivos con información integrada y confiable

Tabla 3. *Áreas que demandan información en RM y uso actual de la misma*

La Gerencia General se sitúa en el nivel más alto de esta demanda, pues necesita información confiable, integrada y disponible en tiempo real, a través de tableros ejecutivos, para evaluar el desempeño de la organización y apoyar la planificación estratégica (Sharda et al., 2023). Esta segmentación de usuarios es el antecedente de los tableros por perfil y de la seguridad a nivel de fila que desarrollamos en el Capítulo 4.

1.4 Objetivos de cada área y necesidades detectadas

Para cada área contrastamos el objetivo que persigue con la información disponible y con las necesidades que justifican el sistema. La Gerencia Comercial busca supervisar el cumplimiento de metas y evaluar a los equipos comerciales; los jefes de tienda, controlar el rendimiento de su establecimiento; y los asesores, incrementar la conversión de clientes potenciales y alcanzar sus metas individuales (Kotler & Keller, 2016). Por su parte, Marketing procura comprender el comportamiento del consumidor, y la Gerencia General, identificar oportunidades de crecimiento con base en datos confiables (Chen et al., 2012).

Área	Objetivo	Necesidades detectadas
Gerencia Comercial	Supervisar metas, analizar ventas y evaluar a los equipos comerciales	Dashboards ejecutivos actualizados, alertas automáticas y reportes consolidados
Jefes de tienda	Controlar el rendimiento de la tienda y la productividad de los asesores	Reportes automáticos por tienda, comparativos históricos y análisis de tendencias
Asesores comerciales	Incrementar la conversión y alcanzar metas individuales	Historial de cada cliente, programación de seguimientos y estado de oportunidades

Área	Objetivo	Necesidades detectadas
Marketing	Comprender al consumidor y diseñar campañas más efectivas	Segmentación automática, análisis de campañas y medición de conversión
Gerencia General	Evaluar el desempeño y planificar estrategias comerciales	Indicadores consolidados en herramientas de Business Intelligence

Tabla 4. *Objetivos y necesidades de información por área*

En términos generales, las necesidades detectadas se agrupan en cinco ejes: centralización de la información, automatización de reportes, eliminación de procesos manuales, disponibilidad de datos confiables y fortalecimiento de una cultura organizacional basada en el análisis de datos (Laudon & Laudon, 2022). Posteriormente, en el Capítulo 2, tradujimos estos ejes en un inventario de fuentes y en una cuantificación del costo del proceso manual.

1.5 Indicadores clave de desempeño

Con base en los objetivos planteados, propusimos un conjunto de indicadores clave de desempeño orientados a medir la eficiencia de la gestión comercial y a apoyar la toma de decisiones estratégicas (Parmenter, 2020). Cada indicador se definió con su fórmula de cálculo, su objetivo y una meta sugerida. En conjunto, cubren los dos procesos críticos del problema (metas comerciales y clientes potenciales) y el impacto de la propia automatización, medido por el tiempo de generación de reportes.

KPI	Fórmula	Objetivo	Meta sugerida
Cumplimiento de metas comerciales (%)	(Ventas reales / Meta de ventas) × 100	Medir el cumplimiento de los objetivos comerciales	≥ 90 %
Ventas por tienda	Total de ventas por establecimiento	Comparar el desempeño entre tiendas	Según meta
Ventas por asesor	Total de ventas individuales	Evaluar la productividad de cada asesor	Según meta
Conversión de clientes potenciales (%)	(Clientes convertidos / Clientes potenciales) × 100	Medir la efectividad del proceso comercial	≥ 25 %
Clientes potenciales activos	Total de clientes registrados sin conversión	Determinar el volumen de oportunidades disponibles	Monitoreo continuo
Tiempo promedio de conversión	Total de días hasta la venta / Número de ventas	Medir la eficiencia del proceso de ventas	≤ 10 días
Ticket promedio	Ventas totales / Número de transacciones	Evaluar el valor promedio de compra por cliente	Según meta

KPI	Fórmula	Objetivo	Meta sugerida
Crecimiento mensual de ventas (%)	((Ventas actuales − Ventas anteriores) / Ventas anteriores) × 100	Analizar la evolución de las ventas	≥ 5 %
Cumplimiento de metas por tienda (%)	(Ventas de la tienda / Meta asignada) × 100	Comparar el rendimiento de cada establecimiento	≥ 90 %
Tiempo de generación de reportes	Horas requeridas para emitir reportes	Medir el impacto de la automatización	≤ 15 min

Tabla 5. *Indicadores clave de desempeño propuestos para el Sistema de Inteligencia Comercial*

Concebimos estos indicadores para visualizarse en tableros interactivos que permitan monitorear el comportamiento comercial, identificar desviaciones respecto a las metas y evaluar el desempeño individual y organizacional (Parmenter, 2020; Sharda et al., 2023). Los KPI constituyeron así la base sobre la cual se diseñarían, en las etapas siguientes del proyecto, el modelo de datos y los tableros de control, con el propósito de fortalecer una cultura organizacional basada en evidencia y mejorar el cumplimiento de las metas comerciales. Desde entonces, fueron estos indicadores, antes que una lista de tecnologías, el criterio con el que evaluamos en cada etapa del proyecto si una decisión agregaba valor al negocio.

CAPITULO 2 Origen de los Datos y Procesos de Carga

En esta etapa del proyecto analizamos el origen de los datos, su recurrencia de carga, los sistemas implicados y los informes que hoy produce RM; además, fijamos los objetivos y el plan con los que se implantaría el sistema. Con ello pasamos del problema de negocio planteado en el Capítulo 1 a un inventario verificable de fuentes y a una cuantificación del costo actual del proceso manual.

2.1 Fuentes de información y recurrencia de carga

La información que debe integrar el sistema proviene de fuentes internas, generadas por los procesos operativos de la propia compañía, y de fuentes externas, procedentes de plataformas digitales y de terceros. Las fuentes internas constituyen el insumo principal de los KPI definidos en el planteamiento del problema (cumplimiento de metas, conversión de clientes potenciales y ticket promedio). Las externas, en cambio, enriquecen el análisis con el contexto del mercado y del comportamiento digital de los clientes y, por su carácter complementario, tienen una recurrencia de carga menor (Laudon & Laudon, 2022; Sharda et al., 2023).

Fuente	Tipo	Información que aporta	Propósito	Recurrencia de carga
Punto de venta (POS)	Interna	Transacciones de venta: producto, monto, medio de pago, tienda y asesor	Operativo / Comercial	Diaria (cierre de jornada)
CRM de clientes potenciales	Interna	Prospectos, historial de contacto y estado de la oportunidad	Comercial	Diaria
Módulo de metas comerciales	Interna	Metas asignadas por tienda y por asesor	Corporativo	Mensual, con ajustes eventuales
Sistema de inventarios	Interna	Existencias, traslados y disponibilidad de producto por tienda	Operativo	Diaria
Recursos humanos	Interna	Asesores activos, tienda asignada y esquema de comisiones	Operativo / Corporativo	Mensual
Servicio al cliente	Interna	Reclamos, devoluciones y postventa	Comercial	Semanal
Plataforma de e-commerce	Externa	Tráfico, navegación y conversión de ventas en línea	Comercial	Semanal
Redes sociales y campañas digitales	Externa	Interacciones, alcance y generación de clientes potenciales	Comercial	Semanal
Estudios de mercado y benchmarking	Externa	Comparativos del sector y tendencias de consumo	Corporativo	Trimestral
Información de proveedores	Externa	Tiempos de entrega y disponibilidad de nuevas colecciones	Operativo	Mensual

Tabla 6. *Fuentes internas y externas de información y su recurrencia de carga*

Agrupamos los datos de estas fuentes en seis familias, con el código de cliente o el código de asesor como variable común: ventas diarias (incluidas variables transformadas como el ticket promedio), clientes potenciales, metas comerciales, inventario, desempeño de asesores y variables demográficas y de comportamiento del cliente. Esta clasificación es el antecedente directo de las tablas de hechos y dimensiones que formalizamos en el modelo dimensional (Capítulo 3).

2.2 Situación actual y costo del proceso manual

Cada fuente opera hoy de forma independiente. Los datos del POS, del CRM y del módulo de metas se descargan manualmente en hojas de cálculo, y el personal de la Gerencia Comercial los combina mediante fórmulas y tablas

dinámicas, sin un procedimiento estandarizado entre tiendas y con dependencia del conocimiento tácito de una o dos personas. Según la propia Gerencia Comercial, esta consolidación demanda entre tres y cuatro horas diarias por persona (cerca de 60 horas mensuales). Asimismo, alrededor del 8 % de los reportes mensuales contiene errores de digitación o de fórmulas, y la información llega a la Gerencia General con uno a dos días de retraso. Ese desfase tiene un efecto comercial directo: en el último semestre, la falta de información oportuna impidió reasignar inventario a tiempo en al menos tres quiebres de stock, con una pérdida de venta estimada cercana a USD 15.000 (TIENDEC S.A., 2026). En adelante denominamos a esta pérdida «reacción tardía» ante quiebres ya producidos, para distinguirla de la pérdida total por rotura de stock que cuantificamos en el Capítulo 4 de la Parte II.

Problema	Causa	Consecuencia	Nivel
Duplicidad de datos	Varias fuentes para un mismo dato	Indicadores inconsistentes	Alto
Archivos manuales	Uso de Excel para consolidar	Errores humanos	Alto
Información desactualizada	No existe un proceso ETL	Decisiones tardías	Alto
Baja integración	Sistemas aislados	Reportes incompletos	Alto
Sin validaciones	No existe gobierno de datos	Datos incorrectos	Medio
Retrasos	Consolidación manual	Pérdida de productividad	Alto

Tabla 7. *Problemas del proceso actual de carga de datos*

2.3 Proceso propuesto y sistemas de carga

Propusimos sustituir la consolidación manual por un esquema propio de los proyectos de Data Warehouse en la nube (Davenport, 2014). Al final de la jornada, los sistemas de tienda cierran su actualización y cada uno genera un archivo plano o una interfaz con la información del día, que se envía a un repositorio transitorio (Landing Zone). Un proceso ETL automático, ejecutado en horario nocturno, recoge esa información y actualiza el Data Warehouse; si una tienda no logra transmitir, el proceso toma el último archivo disponible para no interrumpir los tableros. De este modo, la información queda disponible de forma consistente y con trazabilidad diaria, sin que los usuarios de negocio deban conocer las tablas en las que reside.

Sistema origen	Tipo de archivo / interfaz	Destino de la carga	Frecuencia
POS (ventas)	Interfaz automática	Landing Zone – Data Warehouse RM	Diaria
CRM de clientes potenciales	Archivo plano / interfaz	Landing Zone – Data Warehouse RM	Diaria

Sistema origen	Tipo de archivo / interfaz	Destino de la carga	Frecuencia
Módulo de metas comerciales	Archivo plano	Data Mart Comercial	Mensual
Inventarios	Interfaz automática	Data Mart de Producto	Diaria
Recursos humanos	Archivo plano	Data Mart Comercial	Mensual
Servicio al cliente	Archivo plano	Data Mart de Clientes	Semanal
E-commerce y redes sociales	Exportación / API	Data Mart de Marketing	Semanal
Estudios de mercado	Archivo / informe	Data Mart de Marketing	Trimestral

Tabla 8. *Sistemas origen, tipo de carga, destino y frecuencia de actualización*

Para garantizar la calidad de lo que llega al almacén, cada archivo debe cumplir una estructura mínima de campos y quedar documentado en tipo de dato, fuente y frecuencia, con una certificación previa a su calendarización. A partir de ese punto, la carga es automática e incremental (solo la información nueva o modificada) y deja un registro diario que permite identificar fallas o tiendas sin transmisión (Pérez Galán, 2025a). Recomendamos, además, que el gobierno de la estructura de archivos lo liderara un responsable de Inteligencia Comercial con autoridad sobre definiciones y accesos (Laudon & Laudon, 2022); más adelante en el diseño, esa recomendación se concretó en la Unidad de Inteligencia Comercial y Datos (Capítulo 5).

2.4 Informes actuales y necesidades no cubiertas

La Gerencia Comercial elabora hoy cinco informes a partir de archivos descargados de los sistemas origen: ventas por tienda y ventas por asesor (diarios), metas comerciales (mensual), inventario (diario) y clientes potenciales (diario). Se trata de informes elaborados a mano con datos dispersos, que en el mejor de los casos se actualizan una vez al día. Su alcance, además, es esencialmente descriptivo: muestran qué pasó, pero no ofrecen comparación histórica ni proyección.

Necesidad de información	Actualmente	Con el sistema
Dashboard actualizado durante el día	No	Sí
Alertas automáticas ante desviaciones	No	Sí
Pronóstico de ventas	No	Sí
Rentabilidad por tienda	No	Sí
Identificación de clientes de alto valor	No	Sí
Indicadores históricos y estacionalidad	Parcial	Sí
KPI calculados automáticamente	No	Sí

Tabla 9. *Necesidades de información no cubiertas por los informes actuales*

2.5 Objetivos SMART y plan de implantación

Formulamos los objetivos del sistema con el modelo SMART, de modo que puedan verificarse durante y después de la implementación:

- Reducir en al menos 80 % el tiempo de elaboración de los reportes comerciales (de un promedio de cuatro horas a menos de 45 minutos) durante los primeros seis meses posteriores a la implementación del Data Warehouse.
- Disminuir los errores de consolidación de aproximadamente 8 % a menos de 1 % de los reportes mensuales en un plazo de cuatro meses desde la puesta en marcha del ETL automatizado.
- Incrementar la tasa de conversión de clientes potenciales en al menos 5 puntos porcentuales durante el primer año, mediante el seguimiento habilitado por el Data Mart de Clientes.
- Disponer de tableros actualizados diariamente para el 100 % de las tiendas en un plazo máximo de nueve meses desde el inicio del proyecto.
- Reducir en 3 % las pérdidas comerciales asociadas a quiebres de stock no gestionados a tiempo, en un horizonte de doce meses.

Fase	Actividades principales	Producto	Responsable	Duración
1. Diagnóstico y gobierno de datos	Levantamiento de fuentes, estándares y área responsable	Documento de gobierno y catálogo de fuentes	Intel. Comercial / TI	4–6 sem.
2. Diseño del modelo de datos	Modelado dimensional y diseño de la Landing Zone	Modelo dimensional validado	TI / Intel. Comercial	4 sem.
3. Construcción del Data Warehouse y ETL	Procesos ETL, carga inicial y pruebas de integración	Data Warehouse y ETL en pruebas	TI / proveedor	8–10 sem.
4. Data Marts y dashboards	Data Marts y tableros por perfil en Power BI	Data Marts y dashboards	TI / Intel. Comercial	6 sem.
5. Pruebas, capacitación y cambio	Calidad de datos, capacitación y comunicación	Informe de pruebas y personal capacitado	Intel. Comercial / RR. HH.	4 sem.
6. Puesta en marcha y estabilización	Producción, monitoreo de cargas y ajustes	Sistema en producción y bitácora	Intel. Comercial / TI	4 sem. (continuo)

Tabla 10. *Plan general de implantación por fases, aproximadamente 30 semanas*

Por último, identificamos cinco riesgos de implementación con su respectiva mitigación: calidad inconsistente del histórico a migrar; resistencia del personal de tienda al nuevo proceso, el único calificado con probabilidad alta; dificultades

de integración entre POS, CRM y el ETL; ausencia de un responsable claro del gobierno del dato; y dependencia de un único proveedor tecnológico. Los dos últimos reaparecen en capítulos posteriores, por lo que los tratamos como riesgos transversales en las conclusiones de este trabajo.

CAPITULO 3 Modelo Dimensional y Calidad de Datos

En este capítulo profundizamos el diagnóstico del Sistema de Inteligencia Comercial de RM a partir de seis ejes de análisis. Estos abarcaron los procesos de carga (ETL y almacenamiento tipo data lake), la información disponible y su calidad, los horarios de carga y la accesibilidad, las variables disponibles y el acceso al Data Warehouse, y los sistemas implicados junto con los recursos de TI o externos empleados. Partiendo del inventario de fuentes del Capítulo 2, el aporte central consistió en definir cómo se almacena y modela la información, qué grado de confiabilidad tiene hoy y quién debe acceder a ella.

3.1 Arquitectura en capas y proceso ETL

Un proceso ETL (extracción, transformación y carga) toma datos de fuentes heterogéneas, los homogeneiza y depura, y los deposita en un repositorio de destino. Para RM propusimos un esquema de tres capas instrumentado sobre servicios de Microsoft Azure, con Azure Data Factory como servicio de integración y orquestación del ETL (Microsoft, 2025a). El diseño es coherente con el enfoque estándar de arquitecturas de Business Intelligence en la nube (Sharda et al., 2023) y con el concepto de repositorio transitorio previo a la carga descrito por Davenport (2014).

Capa	Componente	Función
Aterrizaje (Landing Zone)	Azure Blob Storage	Recibe a diario el archivo plano o la interfaz de cada sistema origen, en su formato original y sin transformar
Transformación (ETL)	Azure Data Factory, en horario nocturno	Extrae de la Landing Zone, limpia, estandariza y valida, y convierte los datos al modelo dimensional
Carga y almacenamiento	Data Warehouse corporativo (Azure SQL Database)	Recibe la carga incremental y concentra el modelo de hechos y dimensiones
Consumo por área	Data Marts Comercial, Clientes, Marketing e Inventario	Subconjuntos especializados que alimentan los tableros de cada área funcional

Tabla 11. *Capas de la arquitectura de datos propuesta para RM*

El proceso ETL nocturno incluye, como mínimo, las siguientes tareas. Primero, extrae los archivos disponibles al cierre de cada jornada y verifica su estructura y sus campos obligatorios según la documentación de gobierno de cada fuente. Luego estandariza los formatos de fecha y los códigos de tienda y de producto, y convierte los datos al modelo de hechos y dimensiones. La carga es

incremental, es decir, se limita a la información nueva o modificada respecto del día anterior, y cada ejecución deja un registro (log) que permite identificar fallas puntuales o tiendas sin transmisión. Como contingencia, si una tienda no transmite su actualización, el proceso toma el último archivo disponible para no interrumpir los tableros (TIENDEC S.A., 2026).

En esta etapa del proyecto hicimos una precisión conceptual: lo que la arquitectura denomina Landing Zone es, en sentido estricto, un área de almacenamiento transitorio (staging area) sobre Azure Blob Storage. No se trata necesariamente de un data lake corporativo completo, con capacidades de gobierno, catalogación y consulta analítica directa sobre datos crudos. Dado que no contábamos con información confirmada de que RM planificara un data lake formal, recomendamos verificarlo con el área de TI antes de incluirlo en el alcance del proyecto.

3.2 Modelo dimensional: hechos y dimensiones

El modelo dimensional preliminar sigue el esquema en estrella de Kimball y Ross (2013): las tablas de hechos concentran las variables numéricas del negocio y las dimensiones aportan los atributos descriptivos que permiten filtrarlas y agruparlas. Optamos por tres tablas de hechos (ventas, metas e inventario) y seis dimensiones, de las cuales tres son conformadas, es decir, compartidas por varias tablas de hechos: Dim_Tiempo, Dim_Tienda y Dim_Producto.

Tabla	Tipo	Variables principales
Hechos_Ventas	Hechos	fecha_id, tienda_id, asesor_id, producto_id, cliente_id, medio_pago_id, monto_venta, cantidad
Hechos_Metas	Hechos	fecha_id, tienda_id, asesor_id, meta_asignada, meta_acumulada, pct_cumplimiento
Hechos_Inventario	Hechos	fecha_id, tienda_id, producto_id, existencia, traslados, quiebre_stock
Dim_Tiempo	Dimensión	fecha, día, mes, trimestre, año, temporada
Dim_Tienda	Dimensión	tienda_id, nombre, ciudad, zona, formato
Dim_Producto	Dimensión	producto_id, categoría, línea, temporada, precio_lista
Dim_Asesor	Dimensión	asesor_id, nombre, tienda_asignada, esquema_comisión
Dim_Cliente	Dimensión	cliente_id, canal_captación, rango_edad, ciudad, estado_oportunidad
Dim_MedioPago	Dimensión	medio_pago_id, tipo, entidad

Tabla 12. *Variables disponibles por tabla de hechos y dimensión del modelo dimensional preliminar*

Estas variables permiten construir los KPI comerciales definidos en el Capítulo 1 (cumplimiento de metas, ticket promedio, tasa de conversión, rotación de inventario, ROI de campañas y pérdida estimada por quiebre de stock) mediante

el cruce de las tablas de hechos con las dimensiones conformadas (TIENDEC S.A., 2026).

La información que alimenta el modelo abarca seis dominios de negocio con grados de disponibilidad desiguales. Las ventas (POS) y las metas comerciales están disponibles, aunque las primeras se cargan de forma manual y las segundas tienen corte mensual. Los clientes potenciales (CRM) están disponibles pero dispersos, y el inventario cuenta con un corte diario. En cambio, el desempeño de asesores (CRM, POS y RR. HH.) y las variables demográficas y de comportamiento (CRM y e-commerce) solo están disponibles parcialmente, lo que condiciona el detalle que pueden alcanzar Dim_Asesor y Dim_Cliente en una primera versión.

3.3 Calidad de la información disponible

El diagnóstico previo había mostrado que la consolidación manual en hojas de cálculo genera duplicidad de datos, errores de digitación y de fórmulas en aproximadamente el 8 % de los reportes mensuales revisados, e inconsistencias entre áreas por la ausencia de una única versión de la verdad (TIENDEC S.A., 2026). En el diseño interpretamos estos hallazgos a la luz de las dimensiones de calidad habituales en la disciplina de gobierno de datos: completitud, exactitud, consistencia, oportunidad, validez y unicidad (DAMA International, 2017). Asimismo, asociamos cada problema con una mitigación incorporada a la arquitectura.

Dimensión	Problema identificado en RM	Evidencia	Mitigación propuesta
Exactitud	Errores de digitación y de fórmulas en Excel	≈ 8 % de reportes mensuales con error (estimación interna preliminar)	Validaciones automáticas de campo en el ETL
Consistencia	Cifras distintas para un mismo indicador entre áreas	Duplicidad de archivos de trabajo	Única versión de la verdad en el Data Warehouse
Oportunidad	Retraso de uno a dos días en la disponibilidad de la información	Consolidación manual diaria o semanal	Carga ETL automática nocturna
Completitud	Omisión de cifras si una tienda no transmite su cierre	Dependencia de archivos manuales por tienda	Último archivo disponible y alertas de no transmisión

Tabla 13. *Dimensiones de calidad de datos y hallazgos del diagnóstico de RM*

Advertimos, no obstante, que las cifras de error (≈ 8 %) y de pérdida de venta por reacción tardía (cercana a USD 15.000 en un semestre) provienen de una estimación preliminar de la propia Gerencia Comercial y no de una auditoría externa de datos. Por ello, recomendamos verificarlas mediante una medición formal de calidad antes de reportarlas como indicadores definitivos del proyecto. Con la misma cautela, precisamos que las estadísticas de volumen de datos y

ciberseguridad del material de la cátedra son cifras globales de la industria y no mediciones propias de RM (Pérez Galán, 2025b).

3.4 Horarios de carga, accesibilidad y seguridad de acceso

La frecuencia de carga al Data Warehouse sigue la recurrencia de cada fuente descrita en el Capítulo 2. Es diaria, en proceso nocturno, para POS, CRM e inventarios; mensual para metas comerciales y Recursos Humanos; semanal para servicio al cliente, e-commerce y redes sociales, y trimestral para los estudios de mercado. Ejecutar el ETL de noche evita interferir con la operación diurna de los sistemas transaccionales de tienda (Sharda et al., 2023). Con este esquema, la información de ventas, CRM e inventario llegaría a los tableros al inicio del día siguiente, con un desfase máximo de un día frente a los uno a dos días del proceso manual.

Para el acceso al Data Warehouse y a los Data Marts planteamos un esquema de control de acceso basado en roles (RBAC), conforme al principio de otorgar a cada usuario solo el acceso que necesita para su trabajo. Aplicado a RM, el esquema comprende:

- Acceso administrativo al Data Warehouse completo, restringido al equipo de TI e Inteligencia Comercial, responsable del modelo dimensional y de la certificación de los archivos de carga.
- Acceso de solo lectura a los Data Marts, diferenciado por perfil: Comercial (Gerencia Comercial), Clientes (asesores), Marketing (equipo de Marketing) e Inventario (jefes de tienda).
- Acceso a los tableros de Power BI mediante autenticación corporativa de inicio de sesión único, durante el horario laboral y desde la red corporativa o VPN, sin que el usuario deba conocer las tablas subyacentes.
- Registro de auditoría de las cargas y, de forma deseable, de los accesos a los datos sensibles del cliente.

Al no tener confirmación de que RM contara con un proveedor de identidad corporativo (por ejemplo, un directorio activo en la nube), recomendamos validarlo con TI antes de detallar las asignaciones de acceso. En el Capítulo 4 se desarrolla la traducción de este esquema a tableros por perfil y a seguridad a nivel de fila en Power BI.

3.5 Sistemas implicados y recursos de TI

En cuanto a los sistemas, distinguimos entre los transaccionales que alimentan la solución (POS, CRM, módulo de metas, inventarios, Recursos Humanos, e-commerce y redes sociales) y los componentes de integración, almacenamiento y visualización que la conforman. La distinción responde a la separación entre sistemas operacionales y de apoyo a la decisión (Laudon & Laudon, 2022). Hoy los sistemas transaccionales operan aislados, sin interfaces automáticas entre sí; la arquitectura los integra a través de la Landing Zone y el ETL sin modificar su funcionamiento interno.

Recurso	Tipo	Función en el proyecto
Azure Blob Storage	Almacenamiento en la nube (externo)	Landing Zone para los archivos diarios de los sistemas origen
Azure Data Factory	Servicio ETL y orquestación en la nube (externo)	Extracción, transformación y carga automática nocturna
Azure SQL Database	Base de datos relacional en la nube (externo)	Alojamiento del Data Warehouse corporativo y los Data Marts
Microsoft Power BI	Herramienta de visualización (externo)	Construcción de los dashboards por perfil de usuario
Equipo de Inteligencia Comercial / TI	Recurso humano interno	Gobierno del dato, certificación de archivos y mantenimiento del modelo
Proveedor o consultor de BI	Recurso humano externo	Apoyo en el Data Warehouse y el ETL durante la fase 3 de implantación
Recursos Humanos	Recurso interno	Capacitación y gestión del cambio hacia el nuevo proceso

Tabla 14. *Recursos de TI internos y externos empleados en el proyecto*

Optamos por servicios en la nube de pago por uso, coherentes con el modelo de plataforma como servicio, en el que el proveedor administra la infraestructura subyacente. La decisión se justificó por la agilidad para disponer de recursos en minutos, el pago únicamente por la capacidad consumida y la escalabilidad sin inversión previa en hardware (Pérez Galán, 2025b). Reconocimos, sin embargo, que este enfoque introduce dependencia de un único proveedor, Microsoft Azure; para mitigar ese riesgo recomendamos cláusulas contractuales de portabilidad y respaldo periódico de la información. La alternativa de herramientas de código abierto se evaluó más adelante, en la Parte II.

CAPITULO 4 Visualización y Gobierno de la Información

En esta etapa del proyecto desarrollamos la estrategia de visualización del Sistema de Inteligencia Comercial a partir de la arquitectura definida en el Capítulo 3: Landing Zone en Azure Blob Storage, procesos ETL en Azure Data Factory, Data Warehouse en Azure SQL Database, Data Marts funcionales y una capa de consumo analítico. El trabajo respondió a cuatro requerimientos: seleccionar la herramienta de visualización más compatible con ese ecosistema, definir usuarios y perfiles de uso para cada tablero, establecer las necesidades de información y los accesos por departamento, y revisar el cumplimiento de los objetivos iniciales. La decisión central fue recomendar Microsoft Power BI como plataforma de autoservicio gobernado, con una condición expresa: que cada usuario visualice únicamente la información autorizada según su función.

4.1 Análisis y selección de la herramienta

Para la evaluación consideramos seis criterios. Los primeros atendieron la integración nativa con Azure SQL Database, Azure Data Factory, Excel y los servicios corporativos de identidad; la capacidad para construir tableros

ejecutivos, comerciales y operativos con filtros interactivos; y el gobierno, la seguridad, la auditoría y el control de acceso basado en roles. Los restantes cubrieron la actualización programada, el acceso web y móvil y la distribución controlada; la curva de aprendizaje, el soporte y el costo total; y la escalabilidad para incorporar nuevos indicadores, tiendas, canales y fuentes. Con esos criterios comparamos cuatro alternativas.

Herramienta	Integración con Azure	Visualización	Gobierno	Costo relativo	Valoración para RM
Power BI	Muy alta	Alta	Alta	Media	Integración natural con la arquitectura Microsoft propuesta
Tableau	Alta	Alta	Alta	Media-alta	Potente visualización, pero añade otra plataforma
Qlik Sense	Alta	Alta	Alta	Media-alta	Buen motor asociativo; menor alineación con el stack
Excel	Media	Baja-media	Baja	Baja	Útil para análisis puntual; insuficiente como plataforma corporativa

Tabla 15. *Comparación de herramientas de visualización*

Tableau y Qlik Sense ofrecen capacidades analíticas comparables; sin embargo, incorporarían una plataforma adicional con nuevas necesidades de integración, licenciamiento y soporte. Excel destaca por familiaridad y costo, aunque sus limitaciones de gobierno, colaboración y escalabilidad reproducirían parte del problema actual de archivos dispersos descrito en el Capítulo 2. Optamos por Power BI por su ajuste con el ecosistema Azure, sin atribuirle una superioridad absoluta: se conecta a los Data Marts, aplica un modelo semántico común y permite seguridad a nivel de fila por zona, tienda o responsabilidad funcional (Microsoft, 2025b, 2025c). La solución se compone de Power BI Desktop para diseñar modelos y prototipos, Power BI Service para publicar y distribuir, espacios de trabajo separados para desarrollo, pruebas y producción, y aplicaciones para entregar contenido certificado a cada audiencia. Un gateway de datos solo sería necesario si alguna fuente permaneciera instalada localmente. Dejamos deliberadamente abierta la modalidad de licenciamiento hasta confirmar el número de usuarios, la capacidad requerida y la infraestructura de identidad disponible.

4.2 Tableros por perfil de usuario

Propusimos cuatro tableros principales, diferenciados por la decisión que cada uno apoya, para evitar que todos los usuarios reciban la misma cantidad de información. Su estructura responde al principio de mínimo privilegio, es decir, acceso suficiente para cumplir la función sin exposición innecesaria (Sharda et al., 2023).

Tablero	Usuarios	Nivel	Decisión principal	Alcance de datos	Frecuencia
Ejecutivo	Dirección General	Estratégico	Desempeño global, desviaciones y rentabilidad comercial	Resumen nacional; todas las zonas y tiendas	Mensual, con revisión semanal de alertas
Comercial	Gerencia Comercial y analistas	Táctico	Metas, ventas, conversión, asesores y oportunidades	Zona, tienda, asesor, canal y período	Diaria
Marketing y Clientes	Marketing y CRM	Táctico	Captación, conversión, campañas y comportamiento del cliente	Agregados y segmentados; datos personales limitados	Diaria / semanal
Operativo de Tienda	Jefes de tienda	Operativo	Metas diarias, inventario, quiebres y productividad	Solo la tienda asignada; comparación regional agregada	Diaria

Tabla 16. *Tableros, usuarios, nivel de decisión y frecuencia de consulta*

El tablero Ejecutivo privilegia pocos indicadores, tendencias, semáforos y excepciones: ventas netas, cumplimiento de meta, ticket promedio, conversión, rotación de inventario, pérdida estimada por quiebre de stock y ROI de campañas. Para ello emplea tarjetas KPI, ranking de zonas, mapa geográfico y matriz de desviaciones. El Comercial añade venta acumulada, brecha y proyección de cierre, con velocímetro de cumplimiento, embudo de clientes potenciales y ranking de asesores. El de Marketing y Clientes integra CRM, e-commerce, redes sociales y campañas para medir clientes potenciales, costo por lead, recompra y conversión por canal. Por su parte, el Operativo de Tienda ofrece la vista diaria del establecimiento: ventas del día, unidades por transacción, existencias críticas, quiebres y productos de baja rotación. Los cuatro comparten una estructura visual común (cabecera con fecha de actualización y responsable del dato, filtros, tarjetas KPI, tendencias, desagregación y bloque de alertas y acciones). En todos ellos el color comunica estado, con verde para cumplimiento, ámbar para riesgo y rojo para desviación crítica.

4.3 Necesidades de información y matriz de permisos

La matriz de necesidades mostró que una misma fuente puede atender decisiones distintas sin exponerse con igual granularidad a todos los usuarios. Ventas y metas son transversales; el CRM aporta mayor valor a Comercial y Marketing, y los registros técnicos de carga corresponden solo a TI e Inteligencia Comercial. Como la granularidad disminuye conforme el perfil se aleja de la operación directa, tradujimos ese criterio en reglas concretas dentro de la matriz de permisos. Estas reglas abarcan también la exportación, ya que el riesgo está tanto en visualizar la información como en descargarla y redistribuirla.

Perfil	Ventas	Metas e inventario	Clientes / CRM	Restricción principal	Exportación
Dirección General	Lectura	Lectura agregada	Lectura agregada	Corporativo	No
Gerencia Comercial	Lectura detallada	Lectura detallada	Lectura operativa	Zona / tienda / asesor	Controlada
Marketing	Lectura agregada	No	Lectura limitada	Campaña / segmento	Controlada
Jefe de tienda	Solo su tienda	Solo su tienda	Leads de su tienda	Tienda asignada	No
Asesor comercial	Ventas propias	Meta propia	Leads asignados	Usuario	No
TI / BI	Administración	Administración	Administración	Técnico	Sí, auditada

Tabla 17. *Matriz de permisos propuesta por perfil*

Las reglas de seguridad y gobierno que acompañan la matriz se agrupan en cuatro frentes:

- Autenticación corporativa y asignación de usuarios a grupos, evitando permisos individuales difíciles de mantener.
- Seguridad a nivel de fila para restringir zona, tienda, asesor o cartera asignada, y separación de ambientes de desarrollo, pruebas y producción.
- Conjuntos de datos y KPI certificados por TI e Inteligencia Comercial como única versión de la verdad, con registro de cargas, errores, accesos sensibles y cambios en modelos o permisos.
- Revisión trimestral de accesos, retiro inmediato de permisos por cambio de cargo o desvinculación, y minimización o seudonimización de datos personales en los análisis de marketing.

4.4 Catálogo de indicadores y navegación

Con el catálogo de indicadores convertimos el modelo dimensional del Capítulo 3 en productos analíticos accionables. Buscamos equilibrar indicadores de resultado (ventas, cumplimiento y ROI) con indicadores explicativos (conversión,

ticket, rotación y quiebres), de modo que el tablero muestre si se alcanzó la meta y también por qué (Parmenter, 2020).

KPI	Fórmula propuesta	Frecuencia	Tablero
% cumplimiento de meta	Venta acumulada / Meta asignada × 100	Diaria y mensual	Comercial / Ejecutivo
Brecha de meta	Meta asignada − Venta acumulada	Diaria	Comercial / Tienda
Ticket promedio	Venta neta / Número de transacciones	Diaria	Todos
Tasa de conversión	Compradores / Clientes atendidos o clientes potenciales gestionados × 100	Diaria / semanal	Comercial / Marketing
Rotación de inventario	Costo de ventas / Inventario promedio	Mensual	Ejecutivo / Tienda
Quiebre de stock	Productos sin existencia con demanda / Productos monitoreados × 100	Diaria	Tienda / Ejecutivo
Conversión de clientes potenciales	Leads convertidos / Leads gestionados × 100	Semanal / mensual	Marketing / Comercial
ROI de campaña	(Beneficio atribuible − Inversión) / Inversión × 100	Por campaña	Marketing / Ejecutivo

Tabla 18. *Catálogo de indicadores, fórmulas y tablero de destino*

Varias fórmulas dependen todavía de reglas de negocio no formalizadas. La tasa de conversión requiere un denominador oficial y el ROI, una metodología de atribución de ingresos a campañas; por tanto, esas reglas deben aprobarse y documentarse antes de automatizar el cálculo. La navegación sigue una ruta de lo general a lo específico (resumen corporativo, zona, tienda y asesor o producto, según permisos). Incluye filtros comunes de tiempo y ubicación y filtros propios de cada perfil (comerciales, de producto y de marketing); en cada tablero se muestran solo los pertinentes, con lo que se evita complejidad innecesaria y se impiden consultas que excedan los permisos.

4.5 Revisión de los objetivos iniciales

Al revisar los objetivos del diagnóstico distinguimos entre cumplimiento de diseño y cumplimiento operativo verificable, dado que la solución todavía tenía carácter de propuesta.

Objetivo inicial	Respuesta de la propuesta	Estado	Evidencia pendiente
Centralizar la información comercial	DW y Data Marts integran POS, CRM, metas, inventario y fuentes digitales	Cumplido en diseño	Validar interfaces y ejecutar piloto ETL

Objetivo inicial	Respuesta de la propuesta	Estado	Evidencia pendiente
Crear una única versión de la verdad	Modelo semántico, KPI certificados y diccionario de indicadores	Parcial	Aprobar definiciones con dueños del dato
Dar seguimiento oportuno a metas	Tablero diario con meta, resultado, brecha y proyección	Cumplido en diseño	Probar actualización nocturna y alertas
Gestionar clientes potenciales	Embudo por canal, estado, tienda y asesor	Cumplido en diseño	Depurar CRM y establecer responsables
Reducir trabajo manual y errores	ETL automatizado y tableros reutilizables	Parcial	Medir tiempo y error antes y después
Controlar inventario y quiebres	KPI de existencia, rotación y quiebre por tienda y producto	Cumplido en diseño	Validar demanda y regla de quiebre
Proteger el acceso a la información	RBAC, seguridad por fila, auditoría y mínimo privilegio	Parcial	Confirmar identidad corporativa y matriz final

Tabla 19. *Revisión del cumplimiento de los objetivos iniciales*

Los estados parciales muestran que un diseño completo no equivale a un resultado alcanzado; el piloto es la etapa que convierte el cumplimiento teórico en evidencia. Para medirlo propusimos indicadores de implementación con metas referenciales, que deberán ajustarse tras levantar la línea base. En cuanto a oportunidad y estabilidad, se fijó una disponibilidad del tablero en 12 horas o menos desde el cierre para ventas, CRM e inventario, y una disponibilidad de al menos 99 % en horario de negocio. En cuanto al uso, las metas son una adopción de al menos 80 % de usuarios activos después del tercer mes, una reducción de al menos 60 % de las horas de consolidación manual y el uso del tablero en el 100 % de las reuniones comerciales periódicas. La hoja de ruta de visualización abarca doce semanas (definición, construcción, pruebas, piloto en una zona o grupo de tiendas y despliegue), seguidas de mejora continua mensual a cargo de un comité de gobierno. Entre los riesgos identificados, los de mayor impacto son los datos incompletos, los accesos excesivos y las definiciones contradictorias de indicadores. En consecuencia, un tablero aislado no resuelve por sí mismo los problemas de calidad y consolidación manual; la implementación debe combinar tecnología, gobierno del dato, seguridad, capacitación y gestión del cambio (Laudon & Laudon, 2022).

CAPITULO 5 La Unidad de Inteligencia Comercial y Datos

En este capítulo abordamos una pregunta que los capítulos anteriores habían dejado abierta: quién gestionará el sistema de Business Intelligence de RM, con qué estructura, personas y recursos, cuánto costará y por qué le conviene económicamente a la compañía. Partimos de que ninguna arquitectura técnica se sostiene por sí sola y seguimos la secuencia Organización BI → Recursos → Costes → Beneficios → Viabilidad. Las cifras de cargos, costes y beneficios son

estimaciones ilustrativas, construidas con rangos de mercado razonables para Ecuador; por tanto, deben validarse con Talento Humano, Finanzas y TI antes de adoptarse como definitivas.

5.1 Diagnóstico organizativo y brechas de competencias

El diagnóstico de la situación actual (AS-IS) mostró que las funciones relacionadas con datos están dispersas. TI mantiene el POS, el CRM, el sistema de inventarios y el módulo de RR. HH. como sistemas independientes, y administra infraestructura y no información. La Gerencia Comercial consolida ventas y metas en hojas de Excel, con errores en cerca del 8 % de los reportes mensuales (TIENDEC S.A., 2026). Marketing extrae a mano los datos de e-commerce y redes sociales y calcula el retorno de sus campañas de forma aislada; Logística detecta tarde los quiebres, y Recursos Humanos no cruza las comisiones con el desempeño real de cada asesor. A ello se suma un vacío organizativo: no existe un responsable único del dato ni un Data Steward, de modo que cada área puede calcular el mismo KPI de forma distinta.

De este diagnóstico derivamos cinco problemas organizativos de fondo (ausencia de gobierno del dato, falta de un data owner por dominio, coordinación inexistente entre áreas, especialización técnica insuficiente y propiedad del dato dispersa) y un conjunto de carencias que, conforme a una clasificación de uso extendido, agrupamos en personas, conocimiento, procesos y tecnología (Pérez Galán, 2026a). Entre estas carencias sobresalen la falta de catálogo y diccionario de datos, la ausencia de SLA para el soporte de los tableros y la falta de separación operativa entre los ambientes de desarrollo, pruebas y producción de Power BI. Concluimos, en consecuencia, que a RM no le basta con incorporar más tecnología: necesita una estructura capaz de gobernar y explotar la información que la arquitectura ya prevé centralizar.

Competencia	Nivel actual	Nivel necesario	Brecha (GAP)
Power BI (modelado y DAX)	2/5	4/5	2
SQL	1/5	4/5	3
Azure Data Factory / ETL	1/5	4/5	3
Gobierno del dato	1/5	4/5	3
Analítica de negocio	2/5	3/5	1
Gestión de proyectos BI	2/5	4/5	2

Tabla 20. *Matriz GAP de competencias de RM, escala de 1 a 5*

Las brechas más amplias se concentran en SQL, integración de datos y gobierno del dato. De ahí que se justificaran, de forma simultánea, la incorporación de nuevos perfiles, un plan de formación para el personal existente y el apoyo externo durante la implantación.

5.2 Estructura y perfiles de la Unidad

Propusimos crear la Unidad de Inteligencia Comercial y Datos como puente entre Negocio, Datos y Tecnología, responsable del gobierno de los datos y de la certificación de los archivos de carga. Sus funciones abarcan el diccionario de KPI y las reglas de calidad, la administración de las cargas en Azure Data Factory y la validación de la calidad según sus dimensiones estándar: exactitud, consistencia, completitud, oportunidad, validez y unicidad (DAMA International, 2017). Le corresponden también el mantenimiento del modelo dimensional (Kimball & Ross, 2013), el desarrollo de los cuatro tableros por perfil, la seguridad por roles y a nivel de fila, el soporte, la formación y el control de los proveedores externos.

Optamos por una estructura deliberadamente compacta, para evitar la sobreespecialización prematura. La Unidad depende de la Gerencia Comercial, que es el área con mayor necesidad de información y la que impulsa el proyecto. Mantiene, además, una relación matricial con la Gerencia General para el gobierno corporativo del dato y una relación funcional, no jerárquica, con Marketing, Logística, Recursos Humanos y TI; esta última conserva la administración de la infraestructura Azure. Cada puesto se describió con sus funciones, competencias, herramientas y decisiones autónomas. Así, el Data Engineer responde por un desfase del dato no mayor a 12 horas desde el cierre de cada jornada, y el Data Steward puede rechazar la certificación de un indicador o fuente antes de su publicación en producción.

Perfil	FTE	Costo anual (USD)	Responsabilidad principal
Responsable de Inteligencia Comercial y Datos	1	28.000	Prioriza proyectos, controla el presupuesto y gobierna el diccionario de KPI
Data Engineer	1	22.000	Procesos ETL, Landing Zone, Data Warehouse y cargas incrementales
BI Analyst / Power BI Developer	1	18.000	Cuatro tableros, modelo semántico y seguridad a nivel de fila
Data Analyst	1	16.000	Análisis ad hoc de conversión, segmentación y rotación de inventario
Data Steward	0,5	9.000	Catálogo de datos y única versión de la verdad por KPI
Total personal	4,5	93.000	—

Tabla 21. *Plantilla y costo anual estimado de la Unidad, con aportes patronales*

5.3 Modelo de operación: RACI, solicitudes y SLA

Para evitar que la Unidad se convierta en un cuello de botella o en una isla, definimos una matriz RACI, conforme a las buenas prácticas de asignación de responsabilidades en la gestión de proyectos (Project Management Institute, 2021). Según esta, la Unidad es responsable de definir los KPI comerciales y de desarrollar los tableros, pero la aprobación corresponde a la Gerencia Comercial.

TI, por su parte, ejecuta y aprueba las cargas ETL y la gestión de accesos, mientras que la Unidad controla la calidad de los datos y certifica las fuentes, en este último caso con aprobación de TI. Cada indicador nuevo sigue un flujo formal: la necesidad nace en el negocio, el BI Analyst la traduce en requerimientos, el Data Steward valida la definición y el Data Engineer prepara el dato; luego, el BI Analyst ajusta el tablero y el usuario lo valida antes de su publicación.

Toda solicitud debe expresar el objetivo de negocio, el usuario, la decisión que apoyará, los KPI, la prioridad y la frecuencia requerida. Las solicitudes se priorizan con una matriz de impacto, urgencia y esfuerzo; con ella, los tableros Comercial y Operativo de Tienda obtuvieron prioridad alta, el Ejecutivo media-alta y el de Marketing y Clientes media. Las incidencias siguen un procedimiento de ocho pasos que concluye con la documentación de la causa raíz. El SLA fija una hora de respuesta y menos de cuatro horas de solución para las críticas, cuatro horas y menos de un día para las altas, un día y menos de tres días para las medias, y planificación según prioridad para las bajas.

5.4 Make or buy, formación y recursos tecnológicos

Como criterio de abastecimiento, decidimos conservar dentro de RM el conocimiento crítico del negocio y el gobierno del dato, y externalizar las actividades que exigen especialización puntual o economías de escala que RM no justifica mantener de forma permanente.

Actividad	Interna	Externa	Justificación
Definir KPI comerciales	Sí	No	Requiere conocimiento del negocio de RM
Gobierno del dato	Sí	Apoyo puntual	Función estratégica; apoyo externo solo en el diseño inicial de políticas
Implantación inicial del Data Warehouse y ETL	Parcial	Sí	Especialización técnica en Azure que RM no tiene hoy
Uso continuo de Power BI	Sí	Formación externa	Uso diario del negocio; solo la capacitación inicial se contrata
Ciberseguridad avanzada	Parcial	Sí	Experticia especializada de bajo uso recurrente
Soporte de infraestructura cloud	No	Sí	Escalabilidad y disponibilidad 24/7 propias de Azure

Tabla 22. *Matriz make or buy de la Unidad de Inteligencia Comercial y Datos*

El outsourcing del primer año, de carácter no recurrente, suma USD 45.000: implantación inicial del Data Warehouse y ETL (USD 25.000), diseño de la arquitectura en Azure (USD 9.000), formación especializada (USD 6.000) y auditoría de seguridad (USD 5.000). Sus riesgos (dependencia de Microsoft Azure, vendor lock-in y pérdida de conocimiento al cerrar el contrato) se mitigan con documentación completa, cláusulas de portabilidad, estándares SQL abiertos y transferencia formal de conocimiento. Organizamos la formación en

tres fases, con 68 horas y USD 11.000: un taller de cuatro horas para dirección, jefes de tienda y asesores; 24 horas de Power BI y SQL para los analistas de Comercial y Marketing, y 40 horas especializadas para el equipo de la Unidad. Por último, estimamos en USD 18.000 anuales los recursos tecnológicos, que comprenden almacenamiento y Data Factory, Azure SQL Database, licencias de Power BI, seguridad y respaldos.

5.5 Business case fundacional

Cuantificamos los beneficios a partir del diagnóstico de TIENDEC S.A. (2026), que reporta unas 60 horas mensuales de consolidación manual, errores en cerca del 8 % de los reportes y una pérdida de venta cercana a USD 15.000 por semestre por reacción tardía ante al menos tres quiebres de stock. Como base de ingresos se tomó una facturación referencial de USD 8 millones anuales, cifra hipotética que debe sustituirse por el dato real de Finanzas.

Concepto	Año 1 (USD)	Año 2 (USD)	Año 3 (USD)
Personal	93.000	95.000	97.000
Tecnología	18.000	19.000	20.000
Implantación / outsourcing	45.000	—	—
Formación	11.000	3.000	3.000
Total costes	167.000	117.000	120.000
Ahorro de horas de consolidación	8.100	13.000	14.500
Reducción de errores y reprocesos	3.000	6.000	7.500
Mejora en conversión de clientes potenciales	30.000	58.000	65.000
Menores pérdidas por quiebre de stock (reacción tardía)	10.000	20.000	23.000
Mayor productividad de asesores	8.900	13.000	15.000
Total beneficios	60.000	110.000	125.000
Resultado neto anual	−107.000	−7.000	+5.000
Resultado neto acumulado	−107.000	−114.000	−109.000

Tabla 23. *Costes, beneficios y resultado del business case a tres años, escenario base*

Cada beneficio se proyectó sobre una base de cálculo referencial. El ahorro de horas supone pasar de unas 60 horas mensuales de consolidación a 15–20 horas con el proceso automatizado, valoradas a USD 15 por hora. La reducción de errores parte de la disminución progresiva del 8 % de reportes con error. La mejora en conversión asume entre 5 y 8 puntos porcentuales adicionales sobre el embudo comercial gestionado en el Data Mart de Clientes. La reducción de pérdidas por quiebre se calcula sobre una pérdida de venta estimada en USD 30.000 anuales por reacción tardía. Por último, la productividad de asesores supone entre un 10 % y un 15 % más de tiempo disponible para la venta al reducirse las tareas de consolidación manual.

Distinguimos, asimismo, la naturaleza de cada beneficio. El ahorro de horas y de reprocesos es ahorro contable; la reducción de quiebres mejora el flujo de caja; la conversión y la productividad de asesores constituyen ingresos incrementales, y el capital liberado en inventario se excluyó del cálculo para no contar dos veces el mismo efecto. Con costes acumulados de USD 404.000 y beneficios de USD 295.000, el ROI a tres años del escenario base es de −27 % y el payback se estima entre 4 y 4,5 años. Por su lado, el VAN al 12 % resulta negativo y la TIR queda por debajo de esa tasa. Los escenarios permiten acotar la sensibilidad del resultado:

- Optimista: adopción alta desde el primer semestre y facturación real superior a la referencial; el beneficio acumulado alcanza USD 440.000, con lo que el proyecto sí se recupera en el horizonte analizado.
- Base: adopción conforme al plan de implantación; beneficio acumulado de USD 295.000.
- Pesimista: retrasos en la implantación y menor adopción; beneficio acumulado de USD 175.000.

A nuestro juicio, este resultado no invalida el proyecto, pero obliga a validar con Finanzas la facturación real de RM y a vigilar la velocidad de adopción. Con ese fin fijamos indicadores de gestión de la propia Unidad, entre ellos más del 90 % de proyectos entregados a tiempo, más del 80 % de usuarios activos mensuales, satisfacción superior a 4 sobre 5 y más del 98 % de registros válidos.

PARTE II:

Extensión Analítica y Capacidades Predictivas en Tiempo Real

Esta segunda parte reúne los cinco capítulos finales del proyecto (Bloque 2), en los que extendimos el sistema hacia los datos no estructurados, validamos la arquitectura cloud y su alternativa abierta, e incorporamos las dos primeras capacidades analíticas avanzadas: una predictiva por lotes y otra de procesamiento en tiempo real.

CAPITULO 1 Incorporación de Datos No Estructurados

En esta etapa del proyecto ampliamos la propuesta de la Parte I hacia la información no estructurada y semiestructurada que RM genera y que hoy no aprovecha: notas de asesores, conversaciones de WhatsApp Business y chat web, reseñas y búsquedas del comercio electrónico, reclamos, encuestas y redes sociales. La pregunta de desarrollo fue cómo aprovechar esos datos para anticipar oportunidades comerciales y mejorar las decisiones de inventario y marketing sin vulnerar la privacidad ni perder la trazabilidad. Nuestro propósito consistió en seleccionar fuentes útiles y convertirlas, con controles de privacidad, en variables sobre intención de compra, satisfacción, objeciones, tendencias y desempeño de campañas, antes que en acumular mensajes, imágenes o archivos (Pérez Galán, 2026b).

1.1 Diagnóstico de las necesidades de información

Partimos de las decisiones que toman las áreas de RM (Comercial, Marketing, Inventario, Operaciones de Tienda, Servicio al Cliente y Gerencia) y de las consecuencias de tomarlas con información insuficiente: oportunidades atendidas tarde, presupuesto de campañas asignado sin conocer las causas de la respuesta, ventas perdidas por falta de talla o color y problemas recurrentes de tienda sin causa identificada. La plataforma definida para la Unidad de Inteligencia Comercial y Datos integra datos estructurados de ventas, metas, inventario, clientes y campañas. Esos datos muestran qué ocurrió, dónde y cuándo, pero explican de forma limitada por qué el cliente se comporta como lo hace (Sharda et al., 2023).

Problema	Información que falta	Decisión apoyada
Baja conversión de clientes potenciales	Intención, objeción, urgencia y categoría buscada	Priorizar seguimiento y personalizar la oferta
Metas incumplidas	Motivos de no compra y fricción por tienda	Coaching y acciones correctivas
Quiebres de stock	Consultas por talla, color y producto no encontrado	Reposición, transferencia o ajuste de surtido

Problema	Información que falta	Decisión apoyada
ROI de campañas incierto	Sentimiento, tema y producto mencionado	Optimizar contenido, audiencia y presupuesto
Experiencia inconsistente	Tema, severidad y causa del reclamo	Activar alertas y mejorar procesos

Tabla 24. *Problema de negocio, señal faltante y decisión que apoya*

Del diagnóstico concluimos que los sistemas tradicionales son indispensables, aunque insuficientes: el ERP y el CRM describen transacciones y contactos, mientras que chats, reseñas, búsquedas, reclamos y menciones permiten interpretar motivos y detectar señales tempranas. De ahí que el valor de Big Data para RM resida en relacionar ambas perspectivas más que en incrementar el volumen almacenado (Davenport, 2014).

1.2 Identificación y clasificación de las fuentes

Clasificamos las fuentes por tipo de dato, origen y frecuencia de actualización, y las seleccionamos con seis criterios. El primero es la relevancia, es decir, que cada fuente responda a una decisión concreta; le siguen la legitimidad (finalidad, base legitimadora, origen y derechos del titular documentados) y la calidad (exactitud, completitud, oportunidad, representatividad, unicidad y consistencia). Completan el conjunto la factibilidad (API, exportación controlada o mecanismo autorizado y sostenible), el riesgo, entendido como evitar información sensible innecesaria o perfilado injustificado, y la economía, que exige un beneficio superior al costo de acceso, almacenamiento, procesamiento y gobierno. Se priorizaron datos propios, agregados, públicos o licenciados, y se descartó adquirir bases de contactos sin verificar procedencia, consentimiento y compatibilidad de finalidad (DAMA International, 2017).

Fuente	Tipo de dato	Origen	Actualización	Utilidad
ERP y ventas	Estructurado	Interna	Diaria	Cruzar resultados con señales de clientes
CRM	Estructurado y semiestructurado	Interna	Diaria	Segmentar y vincular clientes potenciales autorizados
Inventario	Estructurado	Interna	Intradiaria	Contrastar consultas con existencias por SKU
Notas de asesores	Texto libre	Interna no estructurada	Diaria	Intención, objeciones y próxima acción
WhatsApp Business y chat web	Conversaciones y metadatos	Interna no estructurada	Intradiaria	Motivo, urgencia, producto y resultado

Fuente	Tipo de dato	Origen	Actualización	Utilidad
E-commerce	Reseñas, búsquedas, clics y abandono	Interna mixta	Intradiaria	Interés, fricción y demanda no satisfecha
Encuestas y reclamos	Texto libre y archivos	Interna no estructurada	Diaria	Satisfacción, causa y severidad
Redes sociales propias	Texto, imagen y video	Interna no estructurada	Diaria	Campañas y conversación con la marca
Tendencias de búsqueda	Series agregadas y términos	Externa estructurada	Semanal	Estacionalidad e interés por categoría
Datos demográficos y económicos	Tablas y series oficiales	Externa estructurada	Según fuente	Contextualizar zonas y tiendas
Meteorología pública	Series temporales	Externa estructurada	Diaria	Relación entre clima y demanda
Social listening autorizado	Menciones públicas	Externa no estructurada	Diaria	Reputación y tendencias con licencia válida

Tabla 25. *Inventario consolidado de fuentes estructuradas y no estructuradas*

1.3 Relación entre fuentes e integración con la arquitectura

Cada relación entre fuentes se justificó con una pregunta de negocio, y en la matriz de relaciones propusimos cinco cruces. Ventas, CRM y notas permiten identificar factores asociados con la conversión y ordenar el seguimiento comercial; inventario, chats y reseñas, cuantificar la demanda no atendida por producto y ubicación; y campañas, redes y ventas, vincular la respuesta cualitativa con el resultado comercial. Por su parte, CRM, reclamos y encuestas sirven para detectar causas recurrentes de insatisfacción, mientras que ventas, tendencias y clima permiten explorar patrones estacionales sin asumir causalidad. Cuando exista una llave autorizada, la información se vinculará mediante identificadores seudonimizados; en los demás casos, el análisis se mantendrá agregado por fecha, campaña, tienda, canal o producto.

Ampliamos la arquitectura Azure de la Parte I sin sustituir el Data Warehouse, mediante un flujo por zonas. Azure Data Factory orquesta la ingesta y Blob Storage conserva los objetos crudos en la zona Bronze durante el periodo justificado. En la zona Silver se normaliza el contenido, se trata la información personal (PII) y Azure AI Language extrae variables de texto. Por último, Azure SQL integra los datos certificados en la capa Gold, el Data Warehouse y los Data Marts, y Power BI presenta indicadores y alertas según perfiles de acceso. Con ello, el Data Warehouse se mantiene como la fuente estructurada para el consumo analítico (Kimball & Ross, 2013).

Cada unidad de texto analizada genera un registro enriquecido mínimo que separa el original de las variables derivadas y permite auditar la versión del

modelo. Sus campos son fuente, registro, marca temporal, canal, cliente seudonimizado cuando aplica, tienda, producto, campaña, idioma, tema, intención, sentimiento y su puntuación, frases clave, urgencia, indicador de PII redactada, versión del modelo, fecha de procesamiento y estado de calidad. Las reglas de procesamiento exigen conservar el original solo durante el periodo de retención, separar ambientes y deduplicar por fuente, fecha y huella del contenido. Asimismo, obligan a validar las variantes lingüísticas del mercado de RM, a enviar a revisión humana los registros de baja confianza o posible sensibilidad y a evitar texto literal en los tableros ejecutivos.

1.4 Casos de uso y priorización

Diseñamos cuatro casos de uso vinculados con los problemas diagnosticados, bajo un principio común: la automatización produce señales de apoyo, mientras que las decisiones comerciales, operativas y de servicio permanecen bajo supervisión humana.

- Priorización de clientes potenciales: ordenar el seguimiento y personalizar la oferta con datos del CRM, de las notas de asesores, de WhatsApp Business y del chat web. El caso está a cargo de Comercial y Servicio al Cliente, y su resultado esperado es un mayor número de clientes potenciales atendidos dentro del SLA y una mayor conversión.
- Detección de demanda no satisfecha: cuantificar la demanda no atendida por SKU, tienda y fecha a partir de inventario, chats, reseñas, búsquedas internas, ventas y devoluciones, con el fin de sustentar transferencias y ajustes de surtido.
- Voz del cliente y evaluación de campañas: relacionar temas y sentimiento con interacción, conversión y ticket promedio, como insumo para Marketing, E-commerce y Gerencia Comercial.
- Alertas de servicio: detectar, entre altos volúmenes de mensajes, los reclamos de alta severidad y derivarlos al responsable para reducir el tiempo de respuesta.

Caso	Valor	Factibilidad	Riesgo inverso	Puntaje	Prioridad
Priorización de clientes potenciales	5	4	4	13/15	Alta
Demanda no satisfecha	5	4	5	14/15	Alta
Voz del cliente	4	4	4	12/15	Alta
Alertas de servicio	4	5	4	13/15	Alta

Tabla 26. *Matriz de priorización de los casos de uso en escala de 1 a 5*

En la escala de riesgo inverso, 5 representa menor exposición legal, ética u operativa. Aunque los cuatro casos obtuvieron prioridad alta, recomendamos

comenzar con la priorización de clientes potenciales y la detección de demanda no satisfecha, por su relación directa con ventas e inventario, la disponibilidad de datos propios y la facilidad de supervisión. Los resultados se presentarán en cinco tableros (ejecutivo, comercial, de marketing, operativo y de gestión BI); este último servirá para gobernar calidad, latencia, precisión, costo y adopción.

1.5 Gobierno, calidad y piloto

Incorporamos la Ley Orgánica de Protección de Datos Personales desde el diseño. La ley exige informar la finalidad, la base legal, los tipos de tratamiento, la conservación y el origen cuando los datos no se obtienen directamente del titular. Además, los datos públicamente visibles no pierden su carácter personal, por lo que RM debe diferenciar la escucha agregada del perfilado individual (Asamblea Nacional del Ecuador, 2021). Definimos ocho controles mínimos, cada uno con responsable: registro de actividades, minimización, transparencia, atención de derechos del titular, evaluación de impacto antes del perfilado a escala, seguridad (RBAC, cifrado, secretos, logs y segregación de ambientes), retención y borrado verificable, y gestión de terceros.

Dimensión	Indicador	Meta piloto	Acción si falla
Completitud	Canal, fecha y fuente válidos	≥ 98 %	Cuarentena del lote
Unicidad	Duplicados sobre registros captados	< 2 %	Ajustar llaves
Oportunidad	Tiempo desde origen hasta disponibilidad	< 12 horas	Revisar SLA
Precisión	F1 macro en intención o tema	≥ 0,80	Reetiquetar y ajustar
Cobertura PII	Hallazgos sensibles tratados	100 % revisado	Bloquear publicación
Trazabilidad	Versión y lote informados	100 %	No certificar dataset
Sesgo	Diferencia de error entre canales	≤ 10 p.p.	Ampliar muestra

Tabla 27. *Criterios iniciales de aceptación de calidad del dato enriquecido*

Riesgo	Probabilidad	Impacto	Mitigación
Captura sin base legitimadora	Media	Muy alto	Inventario legal y bloqueo de fuentes no aprobadas
Uso fuera de finalidad	Media	Alto	Catálogo, control de acceso y revisión de nuevos usos
Clasificación incorrecta	Media	Alto	Umbrales, revisión humana y monitoreo
Sesgo por canal o muestra	Alta	Medio	Muestras estratificadas y métricas comparativas
Exposición de texto sensible	Media	Muy alto	Detección PII, enmascaramiento y agregación
Cambio de API o términos	Alta	Medio	Conectores desacoplados y fuentes alternativas

Riesgo	Probabilidad	Impacto	Mitigación
Costo cloud creciente	Media	Medio	Retención, compresión y alertas de consumo
Baja adopción	Media	Alto	Piloto, formación y KPI de uso y acción

Tabla 28. *Matriz de riesgos del tratamiento de datos no estructurados*

Planteamos la implantación como un piloto medible en seis etapas: gobierno (semanas 1–2), diseño (3–4), piloto con CRM y reseñas (5–8), validación con asesores (9–10), despliegue con automatización, SLA y formación (11–14) y escalamiento a nuevos canales (de la 15 en adelante). Esta última etapa queda condicionada a un business case actualizado. Los KPI del piloto fijan una cobertura de al menos 95 % de los registros elegibles, una latencia inferior a 12 horas y un F1 macro de al menos 0,80. A ello se suman 90 % de clientes potenciales prioritarios contactados dentro del SLA, una conversión incremental de 3 a 5 puntos porcentuales, al menos 75 % de alertas confirmadas como útiles y 80 % de adopción. La meta de conversión tiene carácter académico y está sujeta a la línea base real de RM. Por ello, los beneficios económicos solo se incorporarán al business case después de medir el piloto, lo que evita la doble contabilización. La ampliación a redes sociales, audio o imagen queda supeditada a validar calidad, utilidad, costos y cumplimiento legal.

CAPITULO 2 Selección de la Arquitectura Cloud

Desde el diseño del modelo dimensional (Capítulo 3 de la Parte I) asumimos Microsoft Azure y Power BI como plataforma del proyecto sin haberlos comparado formalmente con otros proveedores. En esta etapa del proyecto corregimos esa omisión siguiendo un orden riguroso: primero el diagnóstico y los requisitos, luego los criterios y pesos de evaluación y, solo al final, la puntuación de las alternativas. El ejercicio confirmó la plataforma elegida, aunque por una razón que ningún documento anterior había enunciado: RM opera sobre un tenant corporativo de Microsoft 365, hecho que condiciona de forma decisiva la integración y el costo de licenciamiento.

2.1 Diagnóstico, requisitos y dimensionamiento

Al mapa de sistemas que RM opera sin interfaces automáticas entre sí, el diagnóstico sumó un activo que no habíamos considerado en los documentos previos: Microsoft 365, con ofimática, Excel e identidad corporativa en Entra ID, ya en uso pero sin aprovecharse para distribuir información analítica. Entre los requisitos fijados figuran integrar las siete fuentes en un repositorio único con 24 meses de histórico, disponer del dato en menos de 12 horas desde el cierre y restringir el acceso por tienda o zona. Se exigió además que la solución pudiera operarla un equipo de 4,5 FTE sin especialización en la nube, con costo predecible y en cumplimiento de la Ley Orgánica de Protección de Datos Personales (Asamblea Nacional del Ecuador, 2021).

El dimensionamiento, que hasta entonces no se había cuantificado, resultó la variable más influyente. Sus valores corresponden a una estimación propia que deberá validarse con TI antes de contratar capacidad.

Dimensión	Estimación actual	Crecimiento anual	Implicación
Hechos de ventas	≈ 2,2 millones de filas/año	10–15 %	Volumen bajo para cualquier motor analítico moderno
Snapshots de inventario	≈ 13 millones de filas/año	10 %	Basta con partición por fecha y tienda
Data Warehouse a 24 meses	< 60 GB	≈ 5 GB/año	No requiere clúster dedicado
Objetos no estructurados	20–40 GB/año	Según canales activados	Costo de almacenamiento marginal
Consultas analíticas	< 1 TiB/mes	Proporcional a la adopción	Niveles gratuitos o de consumo mínimo
Usuarios de negocio	≈ 70 (8 autores, 62 consumidores)	Según apertura a asesores	El licenciamiento por usuario domina el costo

Tabla 29. *Estimación de volumen, crecimiento y usuarios, supuestos a validar*

Concluimos que RM no es un caso de Big Data por volumen. Su problema es de integración, gobierno y oportunidad, y frente al licenciamiento por usuario el cómputo y el almacenamiento resultan casi irrelevantes. Por ello descartamos clústeres dedicados, lakehouses o plataformas como Snowflake y Databricks, y concentramos la comparación en las tres nubes generalistas. En términos de Sharda et al. (2023), RM se encuentra aún en una etapa descriptiva; la arquitectura habilita el salto a la analítica predictiva, pero no lo garantiza.

2.2 Criterios y matriz ponderada de proveedores

Adoptamos sin modificaciones los ocho criterios y pesos orientativos de referencia: integración (20 %), costo total a tres años (15 %), escalabilidad (15 %), seguridad (15 %), datos y BI (10 %), gobierno (10 %), operación (10 %) y portabilidad (5 %). Fue una decisión deliberada, pues alterarlos después de conocer las alternativas habría equivalido a ajustar la vara a favor del proveedor ya preferido. Se compararon componentes equivalentes (Azure SQL Database y Power BI; Redshift Serverless y QuickSight; BigQuery y Looker Studio), con puntuaciones de 1 a 5 ajustadas al perfil de RM y no a la potencia absoluta de cada plataforma.

Criterio	Peso	Azure	AWS	Google	Justificación del diferencial
Integración	20 %	5	3	4	Tenant: Entra ID y Excel; AWS exige más ensamblaje
Costo total	15 %	5	4	5	Tenant: con E5, Power BI Pro sin costo adicional
Escalabilidad	15 %	4	5	5	Azure SQL obliga a migrar si el volumen crece mucho

Criterio	Peso	Azure	AWS	Google	Justificación del diferencial
Seguridad	15 %	5	4	3	RLS de Power BI sobre Entra ID, la más madura
Datos y BI	10 %	4	4	5	BigQuery es un warehouse analítico nativo
Gobierno	10 %	4	4	4	Catálogo y linaje equivalentes para este alcance
Operación	10 %	5	3	4	Tenant: el aprovisionamiento de usuarios ya existe
Portabilidad	5 %	3	4	3	El SQL de Redshift, derivado de PostgreSQL, es más portable
Total ponderado	100 %	4,55	3,85	4,20	Diferencia Azure–Google: 0,35 puntos

Tabla 30. *Matriz de decisión ponderada de proveedores cloud, escala 1–5*

En el mismo análisis explicitamos la sensibilidad del resultado. El tenant de Microsoft 365 incide en integración, costo y operación, criterios que suman el 45 % de la ponderación. Si RM no dispusiera de él, la valoración de Azure bajaría de 5 a 3 en costo y de 5 a 4 en operación, mientras que Google subiría de 4 a 5 en operación; con esos valores, Google Cloud encabezaría la tabla con 4,30 frente a 4,15.

2.3 Costo total y riesgos de la decisión

Los costos se calcularon con precios de lista consultados el 14 de septiembre de 2026, sin descuentos, para 70 usuarios y menos de 60 GB. Azure se presentó en dos escenarios, ya que Power BI Pro está incluido en los planes Microsoft 365 E5 y Office 365 E5, pero no en E3 ni en Business Premium.

Componente	Microsoft Azure	Amazon Web Services	Google Cloud
Licencias de BI	Con E5: 0 incremental; sin E5: 14 USD/usuario/mes → 11.760	QuickSight: 8 autores y 62 lectores → 4.536	Looker Studio Pro solo para 8 creadores → 864
Motor analítico	Azure SQL Database → 4.200	Redshift Serverless → ≈ 1.620	BigQuery bajo demanda → ≈ 0–150
Orquestación y ETL	Data Factory → ≈ 400	Glue → ≈ 500	Dataform y Data Transfer → ≈ 300
Almacenamiento	≈ 25	≈ 25	≈ 25
Total anual	≈ 4.600 con E5; ≈ 16.400 sin E5	≈ 6.700	≈ 1.350
Total a tres años	≈ 13.800 con E5; ≈ 49.200 sin E5	≈ 20.100	≈ 4.050

Tabla 31. *Costo anual estimado de la plataforma por capa, en USD*

Las líneas de Data Factory y Azure SQL Database son las menos confiables, pues se apoyan en fuentes secundarias y en nuestra propia estimación. El cuadro, además, caduca pronto: Microsoft elevó Power BI Pro de 10 a 14 dólares en abril de 2025. El punto más frágil es el licenciamiento. En el business case de la Unidad de Inteligencia Comercial y Datos consignamos USD 3.600 anuales en licencias de Power BI sin precisar la modalidad, una partida innecesaria con E5 y que asciende a USD 11.760 con E3 o Business Premium.

Riesgo	Probabilidad	Impacto	Mitigación
Sobrecosto de licencias al extender tableros a 45 asesores	Alta	Alto	Despliegue por fases y revisión semestral de licencias ociosas
El plan de Microsoft 365 no incluye Power BI Pro	Media	Alto	Verificar el SKU con Compras; Premium por Usuario solo para 8 autores
Dependencia reforzada de un único fabricante	Media	Alto	SQL estándar, modelo documentado, exportación a formatos abiertos
Consumo inesperado en Data Factory o Azure SQL	Media	Medio	Cargas incrementales y alertas de presupuesto
Capacidad interna insuficiente (brecha 1/5)	Alta	Alto	Formación por fases y outsourcing inicial (I.5)
Incumplimiento de la LOPDP con fuentes no estructuradas	Media	Muy alto	Controles de II.1: minimización y seudonimización

Tabla 32. *Riesgos de la decisión tecnológica y su mitigación*

2.4 Selección y justificación ejecutiva

Seleccionamos Microsoft Azure con Power BI por el tenant corporativo confirmado por TI y no por continuidad con las decisiones previas del proyecto, argumento que habría resultado circular. De esta elección esperamos tres beneficios:

- Identidad y seguridad sin un plano adicional: la restricción por tienda y zona se apoya en los grupos existentes en Entra ID, sin crear 70 identidades nuevas en un directorio paralelo.
- Adopción más rápida: los usuarios siguen trabajando en Excel y reciben los tableros por los canales corporativos que ya utilizan.
- Costo incremental de plataforma cercano a USD 4.600 anuales si el plan es E5, frente a los USD 18.000 que estimamos en el business case de la Unidad de Inteligencia Comercial y Datos para la capa tecnológica, sin retrabajo sobre el diseño.

Google Cloud quedó como segunda opción: es la más eficiente en términos absolutos, pero exige un segundo plano de identidad. AWS se descartó porque no encabeza ningún criterio decisivo y exige más a una capacidad técnica interna valorada en 1 sobre 5; tampoco propusimos un modelo multinube. Antes de

comprometer presupuesto queda una única verificación: confirmar con Compras el plan exacto de Microsoft 365, del que depende que el costo anual sea de USD 4.600 o de USD 16.400. Para controlar el gasto optamos por el marco FinOps (informar, optimizar y operar) y asignamos a la Unidad de Inteligencia Comercial y Datos la revisión mensual de facturación (FinOps Foundation, 2026).

2.5 Modelo de servicio, gobierno y arquitectura final

Según la definición del NIST, los modelos de servicio se distinguen por la frontera de responsabilidad entre proveedor y cliente (Mell & Grance, 2011). Optamos por una arquitectura predominantemente PaaS, dado que la matriz GAP de la Unidad de Inteligencia Comercial y Datos sitúa a RM en 1 sobre 5 en integración y gobierno del dato, de modo que no puede asumir la administración de infraestructura. A cambio, la empresa acepta una mayor dependencia del proveedor.

Modelo	Responsabilidad del proveedor	Responsabilidad de RM	Aplicación en la arquitectura
SaaS	Aplicación, plataforma e infraestructura	Usuarios, configuración y contenido publicado	Power BI Service y Microsoft 365
PaaS	Infraestructura, ejecución y servicios gestionados	Datos, modelos, pipelines y configuración	Dominante: Azure SQL Database, Blob Storage, Data Factory, AI Language
IaaS	Hardware y virtualización	Sistema operativo, middleware, aplicaciones y datos	No se utiliza: sin capacidad para administrar servidores

Tabla 33. *Modelos de servicio en la nube y su aplicación en RM*

El gobierno abarca accesos por mínimo privilegio sobre Entra ID, cifrado y seudonimización, certificación del Data Steward antes de publicar, retención conforme a la LOPDP y separación de entornos. Mediante una matriz RACI asignamos a la Unidad la operación de la ingesta, la certificación de la calidad y el control de costos, y a TI la gestión de identidades y la aprobación del licenciamiento. Microsoft audita Azure frente a ISO/IEC 27001:2022; ello, sin embargo, no exime a RM de configurar sus propios controles (International Organization for Standardization, 2022).

La arquitectura final consta de seis capas independientes del proveedor (fuentes, ingesta, zona raw, transformación, Data Warehouse y capa BI), cada una comunicada únicamente con la contigua. Cualquiera que sea el proveedor, esta arquitectura lógica, el modelo PaaS y el esquema de gobierno se mantienen sin cambios. En el Capítulo 3 examinamos la alternativa de código abierto.

CAPITULO 3 Arquitectura Alternativa Open Source

En este capítulo no buscamos sustituir la decisión tomada al seleccionar la arquitectura cloud, donde elegimos Microsoft Azure con Power BI sobre el tenant

corporativo de Microsoft 365. Nuestro propósito fue dimensionar su costo de oportunidad: diseñar una arquitectura equivalente construida enteramente con frameworks open source, evaluarla con los mismos criterios de negocio y dejar documentada una alternativa portable por si el tenant cambiara de condiciones comerciales.

3.1 Punto de partida y requisitos priorizados

El diagnóstico de partida proviene de los capítulos anteriores: siete fuentes sin interfaces automáticas entre sí y una consolidación manual costosa y propensa a errores. El volumen, en cambio, es reducido: cerca de 2,2 millones de filas de ventas al año, un almacén que no supera los 60 GB a 24 meses y 70 usuarios de negocio. De ahí concluimos que RM enfrenta un problema de integración, gobierno y oportunidad antes que un caso de Big Data por volumen. Esa magnitud orienta todo el diseño, pues obliga a evitar el sobredimensionamiento.

La arquitectura lógica, sin marcas comerciales, conservó las seis capas validadas en la arquitectura Azure seleccionada (fuentes, ingesta, zona cruda, transformación, almacén de datos y capa de consulta y BI), complementadas con capas transversales de seguridad, calidad, gobierno y observabilidad. Sobre esa base priorizamos nueve requisitos verificables. Entre ellos figuran una sola versión de la verdad para las siete fuentes, 24 meses de histórico y la clasificación de sentimiento y tema de chats y reseñas con F1 macro de al menos 0,80. Se exigió, además, que el dato esté disponible en menos de 12 horas desde el cierre, que la operación sea viable con 4,5 FTE sin especialización previa, que el costo sea predecible, sin licencias por nodo ni por conector, y que se cumpla la Ley Orgánica de Protección de Datos Personales (Asamblea Nacional del Ecuador, 2021).

3.2 Criterios de evaluación y selección de frameworks

Fijamos los criterios y sus pesos antes de puntuar cualquier alternativa y los aplicamos de forma homogénea a todas las capas, para no adaptar la evaluación a una herramienta preferida. El mayor peso recayó en la proporcionalidad frente al volumen y en la facilidad operativa, dado que el propio equipo declaró una brecha de 1 sobre 5 en SQL, ETL y gobierno.

Criterio	Peso	Qué se valora en el caso de RM
Ajuste al volumen y caso de uso	25 %	Proporcionalidad frente a 2,2 M filas/año y menos de 60 GB de almacén
Facilidad operativa / curva de aprendizaje	20 %	Equipo de 4,5 FTE con brecha 1/5 en SQL, ETL y gobierno
Integración con las fuentes de RM	15 %	Conectores hacia POS, CRM, inventarios, e-commerce y redes
Madurez y comunidad	10 %	Soporte, documentación y talento disponible en Ecuador
Escalabilidad	10 %	Margen ante un crecimiento del 10–15 % anual

Criterio	Peso	Qué se valora en el caso de RM
Costo de soporte y operación	10 %	Horas de administración e infraestructura recurrente
Seguridad	10 %	Gestión de roles y aislamiento por tienda o zona

Tabla 34. *Criterios de evaluación y pesos, fijados antes de puntuar*

Con estos criterios, RM se ubicó en el nivel minimalista, ya que un volumen de esta magnitud no justifica mensajería distribuida ni procesamiento distribuido de propósito general (Kleppmann, 2017). La matriz por capa, puntuada en una escala de 1 a 5, confirmó esa lectura en las cinco decisiones evaluadas.

Capa	Alternativas	Puntuación ponderada	Selección	Motivo del descarte
Orquestación / ETL	Airflow vs. NiFi	4,35 vs. 3,20	Apache Airflow	NiFi añade un modelo de flujo visual que el equipo aprendería desde cero
Procesamiento distribuido	Sin Spark/Flink vs. Spark	4,55 vs. 2,80	Python/dbt sobre PostgreSQL	Spark exige clúster, tuning y perfiles especializados
Mensajería de eventos	Sin Kafka vs. Kafka	4,60 vs. 2,55	Sin mensajería distribuida	Ningún caso exige reacción en segundos; bastan cargas cada 2–4 h
Almacén de datos	PostgreSQL vs. lakehouse (MinIO + Iceberg)	4,50 vs. 3,05	PostgreSQL	Un solo motor y menos de 60 GB no requieren lakehouse
Capa de BI	Superset vs. Grafana	4,40 vs. 3,10	Apache Superset	Grafana se orienta a series de tiempo y observabilidad técnica

Tabla 35. *Matriz comparativa por capa con puntuación ponderada y justificación de descarte*

Airflow aporta orquestación como código versionable en Git, aunque exige disciplina de pruebas. Superset, por su parte, ofrece seguridad a nivel de fila por rol y conexión nativa a PostgreSQL, si bien su editor y su exportación a Excel son más limitados que los de Power BI. En lenguajes y bases de datos optamos por PostgreSQL frente a ClickHouse, sobredimensionado para este volumen, y por Python y SQL con dbt frente a Scala o Java, que solo se justifican junto a Spark o Flink. Asimismo, como zona cruda para 20–40 GB anuales de chats, reseñas e imágenes preferimos MinIO, compatible con la API de S3, a HDFS.

3.3 Arquitectura final y flujo de datos

Las seis capas lógicas quedaron materializadas de la siguiente manera. La ingesta se resuelve con DAG de Airflow, en batch nocturno y cada 2–4 horas

para inventario, y MinIO actúa como zona cruda con retención de 90 días. Para limpiar, seudonimizar y calcular el sentimiento se emplean dbt y Python; PostgreSQL aloja el modelo en estrella de hechos de ventas, metas e inventario durante 24 meses (Kimball & Ross, 2013), y Superset publica KPI certificados con restricción por tienda o zona.

Tramo	Dato que circula	Frecuencia	Decisión que habilita
Fuentes → Ingesta	Archivos y respuestas de API de las siete fuentes	Diaria; cada 2–4 h en inventario	Ninguna todavía: habilita el resto del flujo
Ingesta → Zona cruda	Objetos crudos con marca temporal y origen	Al cierre de cada carga	Reprocesar sin volver a pedir el dato al origen
Zona cruda → Transformación	Registros validados, deduplicados y seudonimizados	DAG nocturno de dbt	Ningún tablero expone datos personales sin enmascarar
Transformación → Almacén	Hechos y dimensiones certificados, carga incremental	Diaria	Cumplimiento de metas por tienda/asesor; rotación de inventario
Almacén → Capa BI	Consultas sobre el modelo semántico y KPI certificados	Refresco diario; alertas por umbral	Reponer o transferir producto, priorizar clientes potenciales, ajustar campañas

Tabla 36. *Flujo de datos tramo por tramo de la arquitectura open source*

En las capas transversales, Keycloak gestiona identidades y roles como equivalente abierto de Entra ID, mientras que los logs de Airflow cubren la observabilidad. Git y una integración continua simple automatizan el despliegue, y Docker Compose empaqueta los servicios sin recurrir a Kubernetes, que el equipo no podría administrar y que el volumen tampoco exige.

3.4 Seguridad, gobierno, calidad y operación

Concretamos, asimismo, los controles que en Azure resuelve el proveedor y que en esta alternativa deben configurarse de forma explícita (DAMA International, 2017):

- Accesos: roles en Keycloak bajo mínimo privilegio, replicados como seguridad a nivel de fila en PostgreSQL y en Superset.
- Protección: cifrado en tránsito (TLS) y en reposo, seudonimización de clientes y enmascarado del texto libre.
- Calidad: tests de dbt que impiden pedidos sin identificador de cliente o stock negativo, verifican la correspondencia entre facturas y ventas y exigen una frescura máxima de 12 horas.
- Gobierno: dueño de negocio por dominio, certificación de indicadores por el Data Steward y linaje documentado en dbt.
- Operación: copias diarias con pruebas de restauración trimestrales, entornos separados y formación por fases antes de ampliar el uso a los 45 asesores.

3.5 Costo total, riesgos y condición de revisión

Al no existir licencias por usuario ni por conector, el gasto recurrente se concentra en infraestructura y en horas de ingeniería.

Componente	Estimación anual (USD)	Supuesto
Cómputo (Airflow, PostgreSQL y Superset en un servidor de 4 vCPU y 16 GB)	≈ 700	Instancia de nube de gama media o servidor propio virtualizado
Almacenamiento de objetos (MinIO, menos de 100 GB)	≈ 60	Disco o almacenamiento en la nube de bajo costo
Copias de seguridad y pruebas de restauración	≈ 120	Snapshots diarios, retención de 30 días
Implementación inicial y formación (solo año 1)	≈ 2.000	Apoyo externo puntual para cerrar la brecha 1/5
Total año 1	≈ 2.880	—
Total años siguientes	≈ 880	Sin costo de implementación inicial

Tabla 37. *Costo total estimado de la plataforma open source por componente*

Frente a los aproximadamente USD 4.600 anuales que estimamos para Azure con el plan E5, el gasto directo recurrente resulta menor. Sin embargo, la comparación queda incompleta si se detiene ahí, porque el ahorro se traslada a horas internas de administración (parches, alta disponibilidad, seguridad a nivel de fila manual) que Azure incluye como servicio gestionado (Mell & Grance, 2011).

Riesgo	Probabilidad	Impacto	Responsable	Mitigación
Falta de perfiles en Airflow, dbt y PostgreSQL	Alta	Alto	Unidad BI	Formación por fases y soporte externo en la implementación
Ausencia de SLA comercial al autoalojar	Media	Alto	TI	Soporte empresarial de un partner o runbooks internos
Carga operativa de infraestructura propia	Alta	Medio	TI	PostgreSQL gestionado en la nube en vez de autoalojar toda la pila
Seguridad por fila y gobierno menos maduros que en Power BI	Media	Alto	Unidad BI	Vistas SQL y roles de PostgreSQL; pruebas de acceso por perfil
Dependencia de una persona clave	Media	Alto	Unidad BI	Documentación obligatoria y control de versiones en Git

Tabla 38. *Matriz de riesgos de la arquitectura open source y su mitigación*

La recomendación final fue un stack minimalista compuesto por PostgreSQL, Python con dbt, Apache Airflow, Apache Superset y Keycloak, empaquetado con Docker Compose. Esta opción complementa la arquitectura Azure seleccionada como ruta de salida sin dependencia de un único fabricante. Dejamos explícita, además, la condición que obligaría a revisarla: un crecimiento significativo de los canales digitales o la necesidad real de detectar quiebres de inventario en segundos y no en horas justificaría evaluar Kafka y un motor de streaming. El Capítulo 5 retomó esa condición al tratar el procesamiento en tiempo real.

CAPITULO 4 Ciencia de Datos y Machine Learning

En esta etapa del proyecto incorporamos una capacidad distinta de las anteriores. Hasta la arquitectura alternativa open source, el sistema explicaba qué había ocurrido y por qué; a partir de aquí, además, anticipa qué va a ocurrir y recomienda una acción concreta, sin sustituir nada de lo construido. Dado que el cálculo de valor exigía parámetros que ninguna etapa previa del diseño había fijado, declaramos supuestos nuevos, sujetos a validación por Finanzas. Son los siguientes: 12 tiendas físicas, coherentes con los 45 asesores y 70 usuarios considerados en la Unidad de Inteligencia Comercial y Datos y en la arquitectura cloud seleccionada; margen bruto del 45 %; facturación referencial de USD 8 millones, tomada del business case de la Unidad de Inteligencia Comercial y Datos, e inventario medio a costo de USD 1,3 millones, con una rotación aproximada de 3,4 veces al año.

4.1 Priorización de la decisión y objetivo del modelo

Evaluamos los cuatro casos de uso definidos en el análisis de datos no estructurados (Capítulo 1) con cinco criterios ponderados, cuyos pesos se fijaron antes de puntuar y no se modificaron después: impacto (30 %), disponibilidad y calidad de los datos (25 %), utilidad para el usuario (20 %), facilidad de implementación (15 %) y riesgo (10 %). En este último criterio, una puntuación alta indica menor exposición legal u operativa. La opción mejor valorada fue la reposición y transferencia semanal por referencia y tienda, con más de un punto de ventaja sobre la segunda. Este resultado converge con la matriz de casos de uso de datos no estructurados, que otorgó 14 sobre 15 a la detección de demanda no satisfecha.

Decisión candidata	Impacto / Datos / Usuario / Facilidad / Riesgo	Total	Prioridad
Reposición y transferencia semanal por SKU y tienda	5 / 4 / 5 / 4 / 5	4,60	1
Priorización de clientes potenciales	4 / 3 / 4 / 4 / 2	3,55	2
Alertas de reclamos críticos de servicio	2 / 3 / 4 / 3 / 3	2,90	3
Voz del cliente y evaluación de campañas	3 / 3 / 3 / 2 / 2	2,75	4

Tabla 39. *Matriz de priorización de decisiones candidatas, escala 1–5*

La priorización de clientes potenciales perdió sobre todo en riesgo, porque trata datos personales y constituye un perfilado individual que exigiría una evaluación de impacto bajo la Ley Orgánica de Protección de Datos Personales (Asamblea Nacional del Ecuador, 2021). La reposición, en cambio, opera exclusivamente sobre datos de producto y existencias. Con ello, el modelo extiende el caso 2 del análisis de datos no estructurados desde el plano descriptivo (detectar la demanda no atendida que ya se perdió) hacia el predictivo, es decir, hacia anticiparla antes de que el quiebre se produzca (Provost & Fawcett, 2013).

Elemento	Definición
Decisión que mejora	Cuánto reponer de cada referencia en cada tienda y qué transferir entre tiendas, cada semana; hoy la toma el jefe de tienda por intuición y cuando el quiebre ya ocurrió
Problema de ML	Regresión: predicción de demanda
Unidad de análisis	SKU-talla-color × tienda × semana
Variable objetivo	Unidades demandadas en la semana t+1; las semanas con existencia cero (demanda censurada) se imputan o se excluyen
Horizonte y frecuencia	Siete días; inferencia semanal (domingo por la noche) y reentrenamiento mensual
Línea base sin IA	Media móvil de las últimas cuatro semanas por SKU y tienda
Técnica elegida	Gradient boosting sobre datos tabulares
Criterios técnicos	WMAPE ≤ 30 %, mejora mínima del 20 % sobre la línea base y recall ≥ 0,70 en quiebres
Criterios de negocio	−25 % de ventas perdidas por rotura, −8 % de inventario medio y −8 horas semanales en pedidos

Tabla 40. *Ficha del modelo: definición, línea base, técnica y criterios de éxito*

Optamos por una solución técnica proporcional a la escala de RM. El gradient boosting maneja de forma nativa la estacionalidad, las variables categóricas y las series intermitentes. Descartamos las redes neuronales y el aprendizaje profundo, porque 12 tiendas y unas 18.000 combinaciones activas no justifican su costo de cómputo ni la especialización que exigirían a un equipo con brecha 1 sobre 5. Tampoco elegimos un ARIMA por referencia, que supondría modelar miles de series intermitentes (Hyndman & Athanasopoulos, 2021). Por último, no se usa la exactitud como métrica, ya que resulta inadecuada para un suceso tan minoritario como la rotura.

4.2 Fuentes de datos, variables y validación

El modelo requiere siete fuentes, cinco internas y dos externas públicas. Ninguna contiene datos personales, puesto que el modelo opera sobre producto, existencias y contexto, no sobre clientes. La fuente más débil es el registro de promociones, de calidad media-baja.

Fuente	Qué aporta al modelo	Histórico	Frecuencia	Calidad	Propietario
POS	Unidades vendidas por referencia, tienda y día	24 meses	Diaria	Alta	Ger. Comercial
Inventario (ERP)	Existencias por referencia y tienda; identifica la censura por rotura	24 meses	Cada 2–4 h	Media	Logística
Maestro de producto	Categoría, talla, color, precio, temporada y antigüedad	Vigente	Mensual	Media	Compras
Promociones	Descuento y campaña por referencia y fecha	18 meses	Semanal	Media-baja	Marketing
Transferencias	Movimientos entre tiendas ya ejecutados	24 meses	Diaria	Media	Logística
Calendario del Ecuador	Feriados, quincenas y fechas comerciales	Completo	Anual	Alta	Externa pública
Previsión meteorológica	Temperatura y precipitación por ciudad	24 meses	Diaria	Alta	Externa pública

Tabla 41. *Fuentes requeridas por el modelo, calidad y propiedad*

Agrupamos las variables en cinco bloques: histórico de demanda (rezagos de 1, 4 y 52 semanas, media móvil y tendencia), producto, comercial, disponibilidad y contexto. Para evitar la fuga de información del futuro, fijamos varias reglas: datos cerrados hasta el domingo anterior, calendario promocional congelado el viernes previo, exclusión de la existencia de la semana objetivo y uso del clima como previsión, nunca como valor observado. La validación es temporal, en cortes sucesivos, porque mezclar fechas al azar inflaría el desempeño (Hyndman & Athanasopoulos, 2021).

El conjunto de entrenamiento ronda 1,9 millones de filas (unas 18.000 combinaciones activas por 104 semanas), un volumen que no justifica un motor distribuido. El modelo tampoco altera ninguna de las capas existentes. Añade un entrenamiento mensual en Python sobre Azure SQL Database (o PostgreSQL) y una canalización de Azure Data Factory (o un DAG de Airflow) que se ejecuta

los domingos a las 22:00. Incorpora, asimismo, una tabla de hechos de predicción con la versión del modelo, que permite auditar cada pedido sugerido, y una vista en Power BI (o Apache Superset) con seguridad a nivel de fila por tienda (Kimball & Ross, 2013).

4.3 Área de aplicación y responsabilidades

El alcance comprende la reposición semanal y la transferencia entre tiendas de producto de la temporada en curso, en las 12 tiendas físicas. Quedan fuera la compra inicial de temporada, los precios y rebajas, el comercio electrónico y las referencias nuevas sin histórico; para estas últimas se aplica una regla por categoría hasta que acumulen cuatro semanas de ventas.

Departamento	Papel en el proceso	Diseño	Operación semanal	Ejecución del pedido	Medición
Gerencia Comercial	Patrocinador	A	I	I	A
Jefes de tienda	Usuario que decide	C	R	C	C
Logística e Inventario	Ejecutor y proveedor de datos	C	C	R	C
Compras	Proveedor de datos	C	I	C	I
Marketing	Proveedor del calendario promocional	C	C	I	I
Unidad BI y Datos	Responsable técnico del modelo	R	R	I	R
TI	Infraestructura	C	C	I	I

Tabla 42. *Departamentos implicados y matriz RACI (R: responsable; A: aprueba; C: consultado; I: informado)*

Dentro de la Unidad de Inteligencia Comercial y Datos, creada en el Capítulo 5 de la Parte I, el Data Engineer mantiene la canalización y el Analista de Negocio vigila el error y la aceptación, por lo que no se crea ningún puesto. Ninguna transferencia se ejecuta de forma automática. El jefe de tienda valida o ajusta la sugerencia y registra el motivo; así se preserva la responsabilidad humana y se genera información para corregir el modelo (Davenport & Harris, 2017).

4.4 Mejora esperada y piloto con grupo de control

Identificamos tres palancas de valor. La primera es reducir las ventas perdidas por rotura, sobre una línea base de margen perdido cercana a USD 108.000 anuales, con un objetivo de −25 %. Esta línea base no contradice la pérdida de USD 15.000 por semestre del diagnóstico inicial, porque miden cosas distintas: aquella cifra recoge solo la venta perdida por reacción tardía ante quiebres ya

producidos (unos USD 30.000 al año), mientras que esta recoge toda la venta perdida por rotura (8.000.000 × 3 % ≈ USD 240.000 anuales), expresada en margen (× 45 %). La primera está imputada a la Unidad de Inteligencia Comercial y Datos (Capítulo 5 de la Parte I); el modelo de reposición actúa sobre el remanente. La segunda es disminuir el inventario inmovilizado, que asciende a USD 1.300.000 a costo, con un objetivo de −8 %. La tercera es recortar las horas dedicadas a pedidos, de 18 a 10 horas semanales en las 12 tiendas. Estas metas son mejoras técnicas y no se trasladan íntegras al resultado. Por ello aplicamos la conversión mejora real = mejora técnica × adopción × ejecución correcta, con una adopción progresiva del 60 %, 85 % y 95 % en los tres años y un factor de ejecución correcta del 80 %.

La atribución se demostrará mediante un piloto de 12 semanas, iniciado al comienzo de una temporada. Participarán seis tiendas de tratamiento, que reciben el pedido sugerido, y seis de control, que mantienen el proceso actual, emparejadas por facturación mensual y mezcla de categorías. La métrica primaria es la diferencia en diferencias de la tasa de rotura y del margen perdido. Para continuar, exigimos una reducción relativa de la tasa de rotura de al menos 15 puntos porcentuales frente al control y una aceptación superior al 60 %. Al cierre se decidirá entre extender el modelo a las 12 tiendas, prorrogar el piloto ajustando el modelo o detener la iniciativa.

4.5 Business case y riesgos

La inversión inicial asciende a USD 29.000, distribuidos en datos (6.000), desarrollo (12.000), integración (5.000), formación (2.500) y piloto (3.500). El costo recurrente es de USD 9.300 anuales, de modo que el primer año exige USD 38.300. Calculamos el beneficio sobre margen y descontamos lo que en el business case de la Unidad de Inteligencia Comercial y Datos ya se imputó al BI descriptivo (10.000, 20.000 y 23.000 USD), para no contabilizarlo dos veces. Como esas cifras se estimaron sobre venta, descontarlas íntegras de una base expresada en margen es una decisión conservadora: reduce, y no infla, el beneficio atribuido al modelo.

Palanca	Fórmula	Año 1	Año 2	Año 3
Ventas perdidas por rotura	(8.000.000 × 3 % × 45 % − ya imputado al BI) × 25 % × adopción × 80 %	11.760	14.960	16.150
Coste de mantener inventario	1.300.000 × 8 % × 18 % × adopción × 80 %	8.990	12.730	14.230
Horas liberadas	8 h × 52 semanas × USD 15/h × adopción × 80 %	3.000	4.240	4.740
Total beneficios	—	23.750	31.930	35.120

Tabla 43. *Beneficio anual por palanca, con fórmula explícita, en USD*

En el escenario base, el resultado neto acumulado pasa de −14.550 USD en el año 1 a +33.900 USD en el año 3, con un VAN al 8 % de +26.400 USD y la

recuperación de la inversión hacia el mes 20. El escenario pesimista (adopción del 40–70 %, mejora del 15 % y margen del 35 %) no recupera la inversión, con un VAN de −8.300 USD, mientras que el optimista alcanza un VAN de +61.200 USD. El punto de equilibrio se sitúa en el 66 % del beneficio estimado. Fuera del retorno quedan, además, cerca de 79.000 USD de inventario liberado que mejoran la caja (Chopra & Meindl, 2019). Los principales riesgos identificados y su mitigación son los siguientes:

- Promociones mal registradas y censura de la demanda (probabilidad alta, impacto alto): certificar y congelar el calendario promocional; imputar o excluir las semanas con rotura.
- Baja aceptación del pedido sugerido (media, alto): validación humana obligatoria, motivo de rechazo e indicador semanal de aceptación.
- Deriva por cambio de colección y referencias nuevas (alta, medio): reentrenamiento mensual, umbral de WMAPE por categoría y regla por atributos.
- Doble contabilización con el business case de la Unidad de BI (media, alto): cálculo sobre el remanente y medición del piloto antes de consolidar cifras.

CAPITULO 5 Caso de Uso en Tiempo Real

5.1 Priorización del caso y análisis de latencia

En este capítulo incorporamos la tercera capacidad analítica del proyecto: el procesamiento de eventos a medida que ocurren. Evaluamos cinco casos candidatos con criterios fijados antes de conocer los resultados: sensibilidad al tiempo (30 %), impacto (25 %), capacidad de actuar (20 %), continuidad de los datos (15 %) y riesgo (10 %). La matriz seleccionó el hueco en sala y la rotura de curva de tallas, con 4,30 sobre 5, por un margen mínimo sobre la respuesta a consultas entrantes (4,25).

Caso candidato	Tiempo 30 %	Impacto 25 %	Actuar 20 %	Datos 15 %	Riesgo 10 %	Total
Hueco en sala y rotura de curva de tallas	4	4	5	4	5	4,30
Consulta entrante sin respuesta dentro del SLA	5	4	4	5	2	4,25
Desviación del ritmo de venta frente a la meta diaria	2	3	3	5	5	3,20
Abandono de carrito en comercio electrónico	3	3	3	4	3	3,15

Caso candidato	Tiempo 30 %	Impacto 25 %	Actuar 20 %	Datos 15 %	Riesgo 10 %	Total
Reclamo crítico detectado en conversación	4	2	4	3	2	3,15

Tabla 44. *Matriz de priorización de casos de uso en tiempo real, escala 1–5*

La puntuación no fue el único argumento. El asesor de sala ya está físicamente frente al producto, mientras que la consulta entrante involucra datos personales y depende de un asesor con disponibilidad irregular. Tampoco la latencia objetivo respondió a una preferencia tecnológica: la dedujimos del tiempo que exige la acción humana. Reponer desde la bodega de la tienda toma unos quince minutos, de modo que una detección en quince minutos, y no en segundos, ya captura casi todo el valor disponible.

Nivel	Latencia	Qué aporta en RM	Decisión
Lotes	Horas o días	No llega a tiempo: el hueco dura toda la jornada	Insuficiente
Micro-lotes	1–15 minutos	Detección en 15 min frente a una acción que tarda otros 15: captura casi todo el valor	Seleccionado
Flujo continuo	Segundos	Adelanta la detección unos minutos sobre una acción que no se acelera	Descartado
Latencia muy baja	Milisegundos	Sin ningún caso de uso que lo requiera	Descartado

Tabla 45. *Descomposición de la latencia y elección del nivel tecnológico*

Definimos el evento de negocio como la ausencia de ventas de una referencia de alta rotación durante una ventana de dos horas en la que debería haberse vendido. Dos reglas lo detectan. La Regla 1 (hueco en sala) contrasta, sobre la previsión horaria del modelo de reposición, las unidades esperadas con las vendidas mediante una distribución de Poisson, y emite alerta cuando la probabilidad de vender cero baja del 5 % (a partir de $\lambda \geq 3$). Exige, además, existencia positiva en el sistema: si la existencia es cero, se trata de una rotura real, que ya cubre el modelo de reposición. La Regla 2 (rotura de curva) emite una alerta de transferencia cuando las tallas centrales de un modelo tienen existencia cero en la tienda y existencia positiva en otra tienda de la misma zona.

5.2 Fuentes de datos y arquitectura

Fuente	Naturaleza	Qué aporta	Emisión en continuo	Propietario
POS	Flujo de eventos	Línea de ticket con referencia, talla, color, unidades e importe	Alta	Gerencia Comercial

Fuente	Naturaleza	Qué aporta	Emisión en continuo	Propietario
ERP de inventario	Flujo de eventos	Movimiento de existencias por referencia y tienda	Media (requiere CDC)	Logística
App de reposición	Flujo de eventos	Cierre de la alerta por parte del asesor	Alta	Unidad BI
Previsión horaria (II.4)	Dato de referencia	Unidades esperadas por ventana, tienda y referencia	No aplica (lote nocturno)	Unidad BI
Maestro y curva de tallas	Dato de referencia	Pesos de cada talla por categoría y ubicación en sala	No aplica (por temporada)	Compras
Turnos del personal	Dato de referencia	Asesor asignado a cada zona y franja	No aplica (semanal)	Recursos Humanos

Tabla 46. *Fuentes del caso de uso: flujos de eventos frente a datos de referencia*

El volumen descarta cualquier tentación de sobredimensionar. Los 2,2 millones de líneas de venta anuales estimadas en la selección de la arquitectura cloud, repartidos en doce horas de apertura, equivalen a 0,14 eventos por segundo en promedio y a un máximo cercano a 1,7 eventos por segundo en el pico del sábado. Son valores varios órdenes de magnitud inferiores a la carga para la que está diseñado un broker de mensajería distribuida. Esta cifra confirma, por una vía de cálculo independiente, la decisión que tomamos en la arquitectura alternativa open source de prescindir de Kafka (Kleppmann, 2017).

Optamos por una arquitectura Lambda ligera. Se conservan intactos el proceso nocturno y el almacén de datos (vía lenta) y se añade una vía rápida dedicada exclusivamente a este caso de uso, con micro-lotes de 15 minutos. Los eventos de la vía rápida y los cierres de alerta llegan también al almacén, con lo que se evitan un repositorio paralelo y una segunda versión de la verdad (Kleppmann, 2017).

Captura	Tabla caliente	Motor de reglas	Alerta	Cierre
Líneas de ticket del POS y movimientos del ERP, por CDC cada 15 min	Ventana móvil de 2 h por SKU-talla-color y tienda	Anomalía de Poisson ($\lambda \geq 3$) y rotura de curva de tallas	Al asesor de la zona: referencia, talla y ubicación en sala	Repuesto, no había stock o falsa alarma, en un toque

Figura 1. *Arquitectura Lambda ligera para el caso de uso en tiempo real de RM: tramos de la vía rápida.*

El alcance comprende las 12 tiendas físicas y solo las referencias de alta rotación de la temporada en curso. Quedan fuera el comercio electrónico, las referencias de baja rotación, el horario de cierre y la reposición planificada semanal, que sigue a cargo del modelo de reposición. El diseño incluye un control de fatiga de alertas: un máximo de seis alertas por tienda y turno, con elevación automática

de umbrales si la precisión cae por debajo del 60 %. Lo consideramos la condición de supervivencia del sistema, porque si más de la mitad de las alertas resultan falsas, el personal deja de atenderlas en cuestión de semanas.

Departamento	Papel	Lógica y umbrales	Atención de la alerta	Transferencia zonal	Medición
Asesores de sala	Actúan	I	R	I	I
Jefes de tienda	Supervisan y escalan	C	A	C	C
Logística e Inventario	Ejecutan la transferencia	C	I	R	C
Unidad BI y Datos	Diseña y calibra la lógica	R	C	I	R
Gerencia Comercial	Patrocina	A	I	I	A

Tabla 47. *Departamentos y matriz RACI del caso de uso en tiempo real (R: responsable; A: aprueba; C: consultado; I: informado)*

5.3 Mejora esperada y business case

Por construcción, esta palanca de valor queda separada de las de la Unidad de Inteligencia Comercial y Datos y del modelo de reposición. Aquellas exigen existencia cero en el sistema; esta, en cambio, actúa sobre producto que figura como disponible pero no es vendible en sala por extravío, prendas en probador, errores de escaneo o curva de tallas rota. Ton y Raman (2010) documentaron este fenómeno en el comercio de moda mediante un estudio longitudinal de cuatro años sobre 333 tiendas. Declaramos un supuesto nuevo, aún no medido: esa pérdida equivale al 2 % de la facturación, USD 160.000 de venta y USD 72.000 anuales de margen (al 45 %), cifra que el piloto deberá sustituir por una medición real.

Proponemos un piloto con cuatro tiendas de tratamiento y cuatro de control durante doce semanas, con auditoría física independiente por muestreo. Al igual que en el modelo de reposición, este diseño es la condición para convertir una mejora técnica en un beneficio de negocio defendible.

Concepto	Micro-lotes (seleccionado)	Flujo continuo (broker y streaming)
Inversión inicial	USD 16.000	USD 42.000
Costo recurrente anual	USD 7.700	USD 24.000
Beneficio captado frente al micro-lote	Referencia (100 %)	105 %, sin acelerar la acción humana de reposición

Tabla 48. *Costes de las dos velocidades comparadas: micro-lotes frente a flujo continuo*

Concepto	Año 1	Año 2	Año 3	Indicadores
Costos totales	−23.700	−7.700	−7.700	Acumulado: −39.100
Beneficios totales	+9.130	+17.900	+22.460	Acumulado: +49.490
Resultado neto anual	−14.570	+10.200	+14.760	ROI a 3 años: 26,6 %
Resultado neto acumulado	−14.570	−4.370	+10.390	VAN al 8 %: +6.970; TIR: 41,6 %
Escenario pesimista (acum.)	−19.250	−18.530	−15.340	VAN: −14.670; no recupera
Escenario optimista (acum.)	−9.530	+10.090	+36.410	VAN: +28.890

Tabla 49. *Business case a tres años de la alternativa seleccionada y escenarios, en USD*

El payback se sitúa en torno al mes 28, de modo que el caso recupera la inversión dentro del horizonte de tres años, aunque de forma ajustada: el punto de equilibrio exige el 83 % del beneficio estimado. El flujo continuo, por el contrario, destruye valor de forma inequívoca. Requiere USD 42.000 de inversión y USD 24.000 anuales para capturar apenas un 5 % más de beneficio, con un VAN incremental de −64.000 USD frente a los micro-lotes, porque adelantar la detección no acelera la reposición física. Por tanto, la variable crítica es la disciplina operativa con que el equipo de sala atienda las alertas, más que una tecnología modesta y ya resuelta. Entre los riesgos del modelo, el primero es la fatiga de alertas, con probabilidad alta e impacto muy alto; le siguen la exactitud del inventario de partida y el uso indebido de las alertas para evaluar el desempeño individual del personal. Este último se mitiga mediante la seudonimización de los identificadores, conforme a la Ley Orgánica de Protección de Datos Personales (Asamblea Nacional del Ecuador, 2021).

En consecuencia, recomendamos aprobar únicamente el piloto acotado de doce semanas y no el despliegue completo. Si la auditoría independiente confirma la reducción del hueco y una atención superior al 70 %, el sistema se extenderá a las 12 tiendas entre los meses 5 y 6; de lo contrario, se detendrá con una pérdida pequeña y documentada.

CONCLUSIONES Y APLICACIONES

En esta sección cerramos el trabajo con una síntesis de lo que aprendimos y de su aplicación al caso RM. Reunimos la síntesis financiera de los tres business case cuantificados, el mapa de riesgos que reapareció bajo distintas formas a lo largo de los diez capítulos del trabajo, la hoja de ruta integrada de implementación y las validaciones que consideramos pendientes antes de comprometer presupuesto.

Síntesis financiera consolidada

Solo tres componentes del sistema incluyen un business case cuantificado: la Unidad de Inteligencia Comercial y Datos, el modelo de reposición y la detección de huecos en sala. Los tres parten de la misma facturación referencial de RM, USD 8.000.000 anuales, y actúan sobre palancas de valor distintas y disjuntas por construcción, de modo que ninguno contabiliza dos veces el mismo beneficio. Para la Unidad de Inteligencia Comercial y Datos cuantificamos el ahorro de horas de consolidación manual y la pérdida evitada por reacción tardía ante quiebres ya producidos (existencia cero en sistema). Para el modelo de reposición estimamos el valor de evitar que esos quiebres lleguen a producirse. En el caso de huecos en sala, por su parte, medimos una pérdida diferente: el producto que figura disponible en el sistema pero no es vendible en sala, que por definición exige existencia positiva y, por tanto, nunca se superpone con las dos anteriores.

Capacidad	Inversión inicial	Costo recurrente anual	Beneficio anual en régimen	Resultado a 3 años
Unidad de BI (I.5)	USD 56.000 (implantación 45.000 y formación inicial 11.000)	USD 117.000–120.000	Ahorro de horas, errores y quiebres ya ocurridos	ROI −27 %; payback 4–4,5 años (escenario base)
Reposición con ML (II.4)	USD 29.000	USD 9.300	USD 35.120	ROI +59,6 %; VAN +USD 26.400; payback ≈ 20 meses
Huecos en sala en tiempo real (II.5)	USD 16.000	USD 7.700	USD 22.460	ROI +26,6 %; VAN +USD 6.970; TIR 41,6 %; payback ≈ 28 meses

Tabla 50. *Comparación de los tres business case cuantificados del proyecto*

Leídos en conjunto, los tres casos dicen más que cualquiera de ellos por separado. Los dos casos predictivos de la Parte II recuperan la inversión dentro

del horizonte de tres años, aunque con holguras muy distintas. La reposición de inventario ofrece el mejor retorno del proyecto porque actúa cada semana sobre el cien por ciento del flujo de inventario. El caso de huecos en sala, en cambio, es positivo pero ajustado (su punto de equilibrio exige el 83 % del beneficio estimado), ya que actúa sobre una fracción más acotada de la pérdida comercial. Solo la Unidad de BI no se recupera en tres años en el escenario base. A nuestro juicio, este resultado es coherente con su naturaleza de capacidad habilitante: se trata de la plataforma de la que dependen todos los demás componentes del sistema y no de una aplicación analítica puntual (Davenport & Harris, 2017).

Las tres cifras dependen, además, del mismo supuesto no confirmado: la facturación de USD 8.000.000 anuales. Las del modelo de reposición y del caso de huecos en sala descansan también en supuestos propios de volumen y en tasas de referencia que declaramos en cada capítulo. Por ello, antes de comprometer presupuesto sobre cualquiera de los tres business case, RM debería validar con Finanzas su facturación real.

Riesgos transversales y gobierno del dato

Identificamos cuatro riesgos que aparecen, con distintas formas, en varios capítulos del trabajo. Conviene leerlos como una sola familia de riesgos del proyecto y no como hallazgos aislados de cada capítulo.

Riesgo transversal	Dónde aparece	Mitigación común
Dependencia de un único proveedor tecnológico	I.3, II.2, II.3	Cláusulas contractuales de portabilidad; arquitectura open source documentada como alternativa (II.3)
Datos personales y cumplimiento de la LOPDP	I.4, II.1, II.5	Minimización, seudonimización, control de acceso por rol, evaluación de impacto antes de perfilar a escala (Asamblea Nacional del Ecuador, 2021)
Capacidad operativa de un equipo de 4,5 FTE	I.5, II.3, II.4, II.5	Ningún modelo nuevo crea un puesto adicional; cada capítulo dimensiona la fracción de tiempo que exige
Cifras preliminares sin validar con TI o Finanzas	I.2, I.3, I.5, II.2, II.4, II.5	Cada capítulo marca de forma explícita qué cifra es un supuesto declarado y qué verificación queda pendiente

Tabla 51. *Riesgos transversales del proyecto y su mitigación común*

El gobierno del dato es el hilo que conecta todo el trabajo. El esquema de control de acceso basado en roles que definimos junto con el modelo dimensional es el mismo que aplicamos a los cuatro tableros, el que extendimos a los datos no estructurados y el que reutilizamos sin modificaciones en el modelo de reposición y en el caso de huecos en sala para las tablas de predicciones y alertas. Más que cualquier cifra individual, esta reutilización evidencia que el proyecto construyó una plataforma y no una colección de iniciativas independientes (DAMA International, 2017; International Organization for Standardization, 2022).

Hoja de ruta integrada

Cada capítulo propuso un cronograma de implementación acorde con su alcance. Al integrarlos, los diez describen una secuencia razonable para llevar el sistema desde la propuesta académica hasta la operación. Las dos primeras fases reproducen el plan de implantación de aproximadamente 30 semanas del Capítulo 2 de la Parte I: la primera agrupa el diagnóstico, el gobierno y el diseño del modelo de datos, y la segunda, la construcción del Data Warehouse, los tableros, las pruebas, la capacitación y la puesta en marcha.

Fase	Capítulos que la componen	Duración referencial	Hito de salida
1. Gobierno y diseño	I.1, I.2, I.3	8 a 10 semanas	Fuentes certificadas, modelo dimensional validado y responsable de datos designado
2. Construcción de la plataforma	I.4, I.5	22 a 24 semanas	Data Warehouse, Data Marts y cuatro tableros de Power BI en producción, bajo la Unidad de Inteligencia Comercial y Datos
3. Extensión del alcance	II.1, II.2, II.3	8 a 12 semanas	Datos no estructurados integrados; plataforma cloud validada formalmente y alternativa open source documentada
4. Capacidades predictivas y en tiempo real	II.4, II.5	10 a 14 semanas	Modelo de reposición en operación; piloto de detección de huecos en sala (4 tiendas de tratamiento y 4 de control, 12 semanas) en curso
5. Mejora continua	Todos	Permanente	Revisión trimestral de indicadores, reentrenamiento de modelos y actualización del business case

Tabla 52. *Hoja de ruta integrada del proyecto, por fases*

Dos secuencias merecen especial atención. En primer lugar, el modelo de reposición y la detección de huecos en sala, que contienen las capacidades más avanzadas del proyecto, dependen por completo del Data Warehouse y de la Unidad de Inteligencia Comercial y Datos diseñados en la Parte I; ninguno es viable mientras esas dos piezas no operen de forma estable. En segundo lugar, los pilotos con grupo de control que exigimos tanto para el modelo de reposición como para la detección de huecos en sala son los que convierten una mejora técnica en un beneficio de negocio defendible. Sin ellos, no podríamos atribuir con confianza ninguna de las cifras anteriores al modelo antes que a la variación natural del negocio.

Validaciones pendientes antes de ejecutar

En cada capítulo declaramos qué cifra era un supuesto del equipo y qué verificación quedaba pendiente antes de comprometer presupuesto. Reunidas en un solo lugar, estas validaciones constituyen la verdadera condición de entrada del proyecto, por sólido que sea el diseño técnico de cada capacidad.

Responsable de validar	Qué debe confirmar	Capítulo de origen
Finanzas	Facturación anual real de RM, hoy asumida como USD 8.000.000 en los tres business case	I.5, II.4, II.5
TI	Plan de Microsoft 365 contratado (si incluye o no Power BI Pro vía plan E5), que mueve el costo anual de la plataforma entre USD 4.600 y USD 16.400	II.2
TI	Si la Landing Zone actual tiene, o no, capacidades de gobierno y catalogación propias de un data lake corporativo completo	I.3
Gerencia Comercial	Tasas reales de conversión de clientes potenciales y rotación de inventario	I.1, II.4
Comercial y Talento Humano	Costos de mercado de los cinco roles de la Unidad de Inteligencia Comercial y Datos, hoy estimados en USD 93.000 anuales	I.5
Logística e Inventario	Disponibilidad y calidad del histórico de quiebres de stock necesario para entrenar y evaluar el modelo de reposición	II.4
Comercial / Auditoría	El supuesto de que el producto no vendible en sala equivale al 2 % de la facturación (USD 160.000 de venta; USD 72.000 de margen), que el piloto de doce semanas debe sustituir por una medición real	II.5
Legal / Recursos Humanos	Política de uso que prohíba emplear los identificadores seudonimizados de cajero o asesor para evaluar el desempeño individual, comunicada a la plantilla antes de la puesta en marcha	II.1, II.5

Tabla 53. *Validaciones pendientes antes de ejecutar, por responsable y capítulo de origen*

Ninguna de estas validaciones invalida el diseño propuesto en los diez capítulos; todas las anticipamos como condición para pasar de la propuesta académica a una decisión de inversión real.

Conclusiones generales

- Consideramos que el Sistema de Inteligencia Comercial de RM resuelve, de forma demostrable en el plano del diseño, el problema que diagnosticamos al inicio del trabajo: información dispersa, procesos manuales y ausencia de una única versión de la verdad. La arquitectura en tres capas y el modelo dimensional, los cuatro tableros de Power BI y la Unidad de Inteligencia Comercial y Datos forman un conjunto coherente. Su elección tecnológica quedó validada formalmente con la selección de la arquitectura cloud y

cuenta con una alternativa portable documentada en la arquitectura open source.

- Las dos capacidades que incorporamos al cierre del proyecto, los datos no estructurados y el procesamiento predictivo o en tiempo real, no sustituyen ninguna capa existente. Ambas se integran sobre la misma arquitectura, el mismo esquema de gobierno y el mismo equipo de 4,5 FTE, sin crear un solo puesto adicional.
- El caso de negocio del proyecto es heterogéneo, y esa heterogeneidad resulta informativa para RM. La reposición de inventario y la detección de huecos en sala recuperan la inversión dentro de tres años, con un margen mucho más amplio en el primer caso. La Unidad de BI, por su parte, exige un horizonte más largo porque es la capacidad habilitante de la que dependen todas las demás.
- Ningún business case del proyecto contabiliza dos veces el mismo beneficio. En los tres business case declaramos de forma explícita sobre qué palanca actúa cada uno y cuál deja fuera: existencia cero reaccionada tarde, existencia cero evitada y disponibilidad en sistema sin venta posible en sala, respectivamente. Esta disciplina permite sumar sus resultados sin inflar el caso de negocio consolidado.
- Todas las cifras del proyecto (la facturación referencial de USD 8.000.000, los volúmenes de datos y las tasas de conversión y de pérdida) son supuestos que declaramos como equipo y no mediciones confirmadas de RM. Antes de cualquier decisión de inversión, corresponde validarlas con TI, Finanzas y la Gerencia Comercial.
- El gobierno del dato, que abarca el esquema de control de acceso basado en roles, la clasificación por sensibilidad y el cumplimiento de la Ley Orgánica de Protección de Datos Personales, constituye la columna vertebral técnica del proyecto. Lo definimos una sola vez, junto con el modelo dimensional, y lo reutilizamos sin cambios en los capítulos posteriores que lo requieren.
- Recomendamos a RM avanzar primero con las Fases 1 y 2 de la hoja de ruta integrada (gobierno, modelo dimensional, Data Warehouse, tableros y la Unidad de Inteligencia Comercial y Datos), porque constituyen la base de la que dependen todas las capacidades posteriores. Las Fases 3 y 4, en cambio, deberían quedar condicionadas a la validación de los supuestos financieros y de volumen declarados en cada capítulo.

REFERENCIAS

Asamblea Nacional del Ecuador. (2021). *Ley Orgánica de Protección de Datos Personales*. Registro Oficial, Quinto Suplemento N.º 459, 26 de mayo de 2021. https://www.gob.ec/regulaciones/ley-organica-proteccion-datos-personales

Chen, H., Chiang, R. H. L., & Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. *MIS Quarterly*, *36*(4), 1165–1188. https://doi.org/10.2307/41703503

Chopra, S., & Meindl, P. (2019). *Supply chain management: Strategy, planning, and operation* (7.ª ed.). Pearson.

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

Davenport, T. H. (2014). *Big data at work: Dispelling the myths, uncovering the opportunities*. Harvard Business Review Press.

Davenport, T. H., & Harris, J. G. (2017). *Competing on analytics: The new science of winning* (ed. actualizada). Harvard Business Review Press.

FinOps Foundation. (2026). *FinOps framework*. https://www.finops.org/framework/

Han, J., Kamber, M., & Pei, J. (2012). *Data mining: Concepts and techniques* (3.ª ed.). Morgan Kaufmann.

Hyndman, R. J., & Athanasopoulos, G. (2021). *Forecasting: Principles and practice* (3.ª ed.). OTexts. https://otexts.com/fpp3/

International Organization for Standardization. (2022). *Information security, cybersecurity and privacy protection — Information security management systems — Requirements* (ISO/IEC Standard No. 27001:2022). https://www.iso.org/standard/27001.html

Kimball, R., & Ross, M. (2013). *The data warehouse toolkit: The definitive guide to dimensional modeling* (3.ª ed.). Wiley.

Kleppmann, M. (2017). *Designing data-intensive applications: The big ideas behind reliable, scalable, and maintainable systems*. O'Reilly Media.

Kotler, P., & Keller, K. L. (2016). *Dirección de marketing* (15.ª ed.). Pearson Educación.

Laudon, K. C., & Laudon, J. P. (2022). *Management information systems: Managing the digital firm* (17.ª ed.). Pearson.

Mell, P., & Grance, T. (2011). *The NIST definition of cloud computing* (NIST Special Publication 800-145). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.SP.800-145

Microsoft. (2025a). *Introducción a Azure Data Factory*. Microsoft Learn. https://learn.microsoft.com/es-es/azure/data-factory/introduction

Microsoft. (2025b). *Documentación de Power BI*. Microsoft Learn. https://learn.microsoft.com/es-es/power-bi/

Microsoft. (2025c). *Seguridad a nivel de fila con Power BI*. Microsoft Learn. https://learn.microsoft.com/es-es/fabric/security/service-admin-row-level-security

Parmenter, D. (2020). *Key performance indicators: Developing, implementing, and using winning KPIs* (4.ª ed.). Wiley.

Pérez Galán, J. L. (2025a). *Fuentes de datos* [Material de clase]. Westfield Business School.

Pérez Galán, J. L. (2025b). *¿Qué es cloud, cloud computing y cloud storage?* [Material de clase]. Westfield Business School.

Pérez Galán, J. L. (2026a). *Business case y creación del departamento* [Material de clase]. Westfield Business School.

Pérez Galán, J. L. (2026b). *Obtención de datos no estructurados* [Material de clase]. Westfield Business School.

Project Management Institute. (2021). *A guide to the project management body of knowledge (PMBOK guide)* (7.ª ed.).

Provost, F., & Fawcett, T. (2013). *Data science for business: What you need to know about data mining and data-analytic thinking*. O'Reilly Media.

Sharda, R., Delen, D., & Turban, E. (2023). *Business intelligence, analytics, and data science: A managerial perspective* (6.ª ed.). Pearson.

TIENDEC S.A. (2026). *Informe interno de gestión comercial* [Documento interno no publicado].

Ton, Z., & Raman, A. (2010). The effect of product variety and inventory levels on retail store sales: A longitudinal study. *Production and Operations Management*, *19*(5), 546–560. https://doi.org/10.1111/j.1937-5956.2010.01120.x