Maestría en

ADMINISTRACIÓN DE EMPRESAS

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

AUTORES:

Msc. Mayra Gabriela Llambo

Msc. Erika Josselyn Llerena

Msc. Geovanna Paola Melo

Msc. Estefania Michelle Pérez

Msc. Johanna Aracely Rosas

TUTOR: Msc. José Antonio García Fernández

Implementación de una solución de Business Intelligence en CVP Centro Constructor Vásquez Pacheco para maximizar las ventas y toma de decisiones.

REINVENTEMOS EL FUTURO

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

Implementación de una solución de Business Intelligence en CVP Centro Constructor Vásquez Pacheco para maximizar las ventas y toma de decisiones.

Por

Mayra Gabriela Llambo
Erika Josselyn Llerena
Geovanna Paola Melo
Estefania Michelle Pérez.
Johanna Aracely Rosas

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: . Mayra Gabriela Llambo, . Erika Josselyn Llerena Guerra, . Geovanna Paola Melo Medina, . Estefania Michelle Pérez, . Johanna Aracely Rosas Guillen 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, . Mayra Gabriela Llambo, Erika Josselyn Llerena Guerra, Geovanna Paola Melo Medina, Estefania Michelle Pérez, Johanna Aracely Rosas Guillen, declaramos bajo juramento que el trabajo de titulación titulado **Propuesta "Implementación de una solución de Business Intelligence en CVP Centro Constructor Vásquez Pacheco para maximizar las ventas y toma de decisiones"** 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.



Mayra Gabriela Llambo

Correo electrónico:

mayrallambo@hotmail.com



Erika Josselyn Llerena

Correo electrónico:

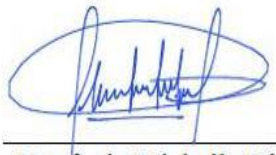
erikajosselyn2894@gmail.com



Geovanna Paola Melo Medina

Correo electrónico:

paomelo43@gmail.com



Estefania Michelle Pérez

Correo electrónico:

estefaniapperezz@hotmail.com



Johanna Aracely Rosas Guillen

Correo electrónico:

johannarosas9807@gmail.com

Autorización de Derechos de Propiedad Intelectual

Nosotros, . Mayra Gabriela Llambo, . Erika Josselyn Llerena Guerra, . Geovanna Paola Melo Medina, . Estefania Michelle Pérez, . Johanna Aracely Rosas Guillen, en calidad de autores del trabajo de investigación titulado Titulo del trabajo de **investigación "Implementación de una solución de Business Intelligence en CVP Centro Constructor Vásquez Pacheco para maximizar las ventas y toma de decisiones"**, 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, 23 de octubre de 2026.



Mayra Gabriela Llambo

Correo electrónico:

mayrallambo@hotmail.com



Erika Josselyn Llerena

Correo electrónico:

erikajosselyn2894@gmail.com



Geovanna Paola Melo Medina

Correo electrónico:

paomelo43@gmail.com



Estefania Michelle Pérez

Correo electrónico:

estefaniapperezz@hotmail.com



Johanna Aracely Rosas Guillen

Correo electrónico:

johannarosas9807@gmail.com

Dedicatorias y Agradecimientos

En primer lugar, agradezco a Dios por ser el motor de mi vida y ser la luz en mi camino que me guía cuando voy a tomar decisiones.

A mi esposo, que es mi complemento y mi mejor amigo que ha estado todos estos anos a mi lado siendo mi fan número uno y empujándome a hacer mejor.

A mis padres por haberme educado bajo valores y principios y que han sido parte de lo que yo soy y como es mi carácter.

A mis compañeras por haber compartido juntas este proyecto de colaboración ya que sin su ayuda no lo hubieras hecho posible.

ÍNDICE DEL DOCUMENTO

CAPÍTULO 1 12
Presentación de la empresa 12
Definición inicial del problema 16
Áreas, objetivos y necesidades detectadas 17
Indicadores clave de desempeño (KPI) 18
CAPÍTULO 2 25
Origen de la información y recurrencia de carga: fuentes internas y externas 25
Sistemas y destino de la carga 27
Almacenamiento en data warehouse y datamarts 29
Proceso ETL (extracción, transformación y carga) 37
Carencias y problemas en la carga de datos 39
Informes generados actualmente, necesidades detectadas y necesidades no cubiertas 41
CAPÍTULO 3 47
Procesos de carga de datos (ETL, Data Lake) 47
Gobernanza y calidad de datos y accesibilidad 52
Horarios de carga y accesibilidad 53
Variables disponibles y acceso al Data Warehouse 54
Sistemas implicados en los procesos y recursos IT o externos empleados 56
Sistemas implicados en los procesos. 56
CAPÍTULO 4 61
Análisis de herramientas ETL y selección de la alternativa más adecuada 61
Usuarios y perfiles de uso de cada dashboard 62
Necesidades de información y acceso por departamento y perfil 68
Revisión y cumplimiento de los objetivos iniciales 71
CAPÍTULO 5 75
Business case y creación del departamento 75
Caracterización del departamento comercial 75
Principales problemas y limitaciones 77
Definición de las carencias en formación y recursos 78
Coordinación de los flujos entre otras áreas 79
Definición del organigrama, dependencias de otras áreas y solución de problemas 80
Business case de costos y beneficios 83
CAPÍTULO 6 86
Selección e identificación de fuentes de datos externas e internas, estructuradas o semiestructuradas 86

Definición de la relación entre los datos 88
Casos de uso y modelo de consumo de bi 90
Big data: impacto en el proyecto, áreas, departamentos y procesos de negocio 94
CAPÍTULO 7 98
Análisis de las alternativas de diferentes proveedores propuestos 98
Selección de la alternativa de proveedor tecnológico 99
Definir nuestro modelo de cloud computing como SaaS, PaaS o IaaS 102
Flujo de la integración de nuestra arquitectura big data con el datawarehouse o estructura bi alternativa 103
CAPÍTULO 8 114
Complementación del ecosistema de big data y analítica avanzada 114
Complementación del ecosistema de big data 115
CAPÍTULO 9 121
Identificar y definir el objetivo de aplicación del modelo de aplicación de Data Science y Machine Learning en el proyecto 121
Fuentes de datos utilizadas 122
Fuentes externas de datos 123
Área de aplicación y departamentos de negocios implicados 124
Identificación de la mejora esperada 124
Mejora esperada en ventas 125
Cuantificación de la mejora en el área de aplicación a través de un Business Case 126
CAPÍTULO 10 130
Identificar y definir el objetivo de aplicación del modelo de aplicación de Real Time 130
Fuentes de datos utilizadas 130
Área de aplicación y departamentos de negocio implicados 131
Identificación de la mejora esperada 134
Cuantificar la mejora en el área de aplicación a través de un Business Case 135
CONCLUSIONES Y APLICACIONES 138
REFERENCIAS 140

Índice de tablas

Tabla 1. Datos de identificación de la empresa 12
Tabla 2. Indicadores operativos y financieros de la empresa 13
Tabla 3. Indicadores clave de desempeño del área Comercial 19
Tabla 4. Indicadores clave de desempeño del área de Operaciones 20
Tabla 5. Indicadores clave de desempeño financieros y transversales 22

Tabla 6. Dimensiones corporativas del data warehouse 31
Tabla 7. Dimensiones del datamart comercial 32
Tabla 8. Tablas de hechos del datamart comercial 33
Tabla 9. Indicadores de marketing soportados por el datamart comercial 34
Tabla 10. Dimensiones del datamart de bodega 35
Tabla 11. Tablas de hechos del datamart de bodega 35
Tabla 12. Indicadores de inventario soportados por el datamart de bodega 36
Tabla 13. Fuentes de datos del proceso ETL 37
Tabla 14. Fases del proceso ETL 38
Tabla 15. Informes generados actualmente por la empresa 42
Tabla 16. Dimensiones corporativas del Data Warehouse 48
Tabla 17. Fuentes de datos del proceso ETL 50
Tabla 18. Dimensiones de calidad de datos y metas 52
Tabla 19. Perfil de uso del dashboard ejecutivo o gerencial 62
Tabla 20. Perfil de uso del dashboard comercial o de ventas 63
Tabla 21. Perfil de uso del dashboard de marketing 64
Tabla 22. Perfil de uso del dashboard de bodega e inventarios 64
Tabla 23. Perfil de uso del dashboard de operaciones y logística 65
Tabla 24. Perfil de uso del dashboard financiero y transversal 66
Tabla 25. Perfil de uso del dashboard de administración del sistema de BI 67
Tabla 26. Recursos externos y servicios subcontratados 82
Tabla 27. Costos del equipo de Business Intelligence y de implementación (USD) 83
Tabla 28. Datos estructurados internos considerados por reporte 86
Tabla 29. Principales relaciones de datos propuestas para CVP 89
Tabla 30. Matriz de priorización de KPI por nivel de impacto 91
Tabla 31. Especificación funcional de los dashboards por perfil de usuario 92
Tabla 32. Impacto esperado del modelo analítico por área de negocio 95
Tabla 33. Criterios de evaluación 107
Tabla 34. Matriz de evaluación ponderada 108
Tabla 35. Complementación del ecosistema de Big Data 115
Tabla 36. Lenguajes de programación del ecosistema de datos 117
Tabla 37. Ecosistemas de bases de datos y búsqueda 119
Tabla 38. Fuentes internas de datos del modelo 122

Tabla 39. Fuentes externas de datos complementarias 123
Tabla 40. Áreas de negocio implicadas en la aplicación del modelo 124
Tabla 41. Mejora esperada e indicadores de validación por ámbito de aplicación 125
Tabla 42. Línea base financiera y operativa de CVP en 2025 126
Tabla 43. Resultado económico de los escenarios y su tratamiento financiero 128
Tabla 44. Departamentos y responsabilidades en el análisis en tiempo real 131
Tabla 45. Indicadores vinculados con el análisis en tiempo real 133
Tabla 46. Flujo propuesto del Business Case 136

Índice de figuras

Figura 1. Flujo de la información operativa y contable a través del sistema SAGA 27
Figura 2. Flujo de datos de la arquitectura propuesta 30
Figura 3. Flujos de trabajo diario y semanal para la extracción, validación y disponibilidad de la información 53
Figura 4. Organigrama propuesto con la incorporación del Departamento de Business Intelligence 81
Figura 5. Arquitectura propuesta del Data Warehouse corporativo y de los Data Marts 95

INTRODUCCIÓN Y OBJETO DEL PROYECTO

Implementar una solución de Business Intelligence en CVP Centro Constructor Vásquez Pacheco que integre y analice la información comercial de la empresa para fortalecer la toma de decisiones estratégicas y maximizar las ventas a nivel nacional mediante el uso de indicadores y análisis de datos en tiempo real.

CAPÍTULO 1

Presentación de la empresa.

CVP Centro Constructor Vásquez Pacheco (en adelante, CVP) es una empresa ecuatoriana constituida como persona natural, cuya razón social corresponde a María Elena Vásquez Pacheco. Su matriz se ubica en la avenida 24 de Mayo y El Vergel, cantón Cañar, y su cobertura geográfica comprende la provincia de Cañar, con presencia en los cantones Cañar y El Tambo. La Tabla 1 resume los datos de identificación de la empresa.

Tabla 1

Datos de identificación de la empresa

Concepto	Descripción
Nombre comercial	CVP Centro Constructor Vásquez Pacheco
Razón social	María Elena Vásquez Pacheco
Registro Único de Contribuyentes (RUC)	0300746740001
Tipo de empresa	Persona natural
Fecha de inicio de actividades	20 de diciembre de 1995
País de origen	Ecuador
Cobertura geográfica	Provincia de Cañar, con presencia en los cantones Cañar y El Tambo
Matriz	Avenida 24 de Mayo y El Vergel, cantón Cañar, Ecuador

Nota. Elaboración propia a partir de la información proporcionada por CVP Centro Constructor Vásquez Pacheco.

En cuanto a su dimensión operativa y financiera, la Tabla 2 presenta los principales indicadores reportados por la empresa: 16 colaboradores, ventas totales por $3,077,453.94, compras totales por $2,679,367.38 y una cartera de 9,000 clientes activos. Se observa que el valor de las compras supera al de las ventas en $126,076.33. Esta diferencia puede asociarse a la reposición de existencias, circunstancia que refuerza la pertinencia de monitorear la rotación de inventarios y el margen bruto, indicadores que se desarrollan más adelante en este capítulo.

Tabla 2

Indicadores operativos y financieros de la empresa

Indicador	Valor
Colaboradores	16
Ventas totales	$3,077,453.94
Compras totales	$2,679,367.38
Clientes activos	9,000

Nota. Elaboración propia a partir de la información proporcionada por CVP Centro Constructor Vásquez Pacheco.

Actividad económica

CVP desarrolla actividades orientadas a la comercialización de materiales para la construcción y a la prestación de servicios relacionados con el sector. Sus principales actividades económicas son tres. La primera es la venta al por menor de artículos de ferretería, herramientas manuales y eléctricas, materiales para bricolaje, alambres y cables

eléctricos, cerraduras, accesorios, extintores, equipos para jardinería y demás productos especializados para la construcción. La segunda es la comercialización de materiales de construcción, acabados, maquinaria y equipos para proyectos de obra civil. La tercera es el alquiler de grúas con operador y de otros equipos especializados para la construcción.

Estructura organizacional

La empresa cuenta con una estructura organizacional dividida en siete áreas estratégicas: Dirección, Talento Humano, Comercial, Tesorería, Operaciones, Calidad y Expansión. El presente estudio se centra en las áreas Comercial y de Operaciones, debido a su relación directa con la generación y el uso de la información de ventas, inventarios y abastecimiento.

Reseña histórica

La empresa nació en 1995 bajo la denominación Ferretería Hermanos Vásquez Pacheco, por iniciativa del señor Carlos Vásquez. Tras su fallecimiento, la dirección quedó a cargo de la señora María Elena Vásquez Pacheco, quien consolidó el negocio y promovió su crecimiento.

En 2015, Carlos Zambrano Vásquez, hijo de la señora Vásquez Pacheco, se incorporó a la empresa para liderar su modernización. Bajo su gestión se amplió el portafolio de productos y la tradicional ferretería se transformó en un centro constructor. Como parte de este proceso, en 2020 se inauguraron nuevas instalaciones que fortalecieron la imagen y la capacidad de la empresa. En 2025, CVP continuó su expansión con la apertura de una sucursal en el cantón El Tambo.

Actualmente, con más de 30 años de trayectoria, CVP es una de las empresas referentes del sector de la construcción y la ferretería en la provincia de Cañar, reconocida por la calidad de sus productos y por su compromiso con el servicio al cliente.

Principales proveedores

CVP mantiene relaciones comerciales con importantes proveedores nacionales e internacionales del sector de la construcción, entre los que destacan Adelca, Unión Cementera Nacional (UCEM – Cemento Guapán), Gala Importaciones, Intaco, Kubiec e Italpisos. Estas alianzas estratégicas contribuyen a garantizar el abastecimiento de productos de alta calidad, lo que fortalece la competitividad de la empresa y la satisfacción de sus clientes.

Misión, visión y valores

Misión

La misión de CVP consiste en satisfacer las necesidades de sus clientes mediante la comercialización oportuna de materiales de construcción, productos de ferretería y acabados de alta calidad, ofreciendo un servicio personalizado y soluciones integrales que contribuyan al desarrollo seguro y exitoso de sus proyectos.

Visión

La visión de CVP es ser líder en el mercado ecuatoriano de materiales para la construcción, ferretería y acabados, destacándose por la innovación, la excelencia en el servicio y la mejora continua, y consolidándose como un referente de confianza y desarrollo para sus clientes y para el sector de la construcción.

Valores

Los valores que orientan la actuación de la empresa son la innovación, la empatía, la responsabilidad y el orden.

Definición inicial del problema

CVP ha experimentado un crecimiento en sus operaciones y en el volumen de información generada por las ventas, los clientes, los inventarios y las compras. Sin embargo, la empresa no cuenta con una herramienta que integre y analice estos datos de manera eficiente, lo que dificulta la toma de decisiones estratégicas y la identificación de oportunidades para incrementar las ventas.

La inteligencia de negocios (*Business Intelligence*, BI) comprende el conjunto de técnicas, sistemas y prácticas que permiten recopilar, integrar y analizar datos con el fin de comprender el negocio y apoyar decisiones oportunas (Chen et al., 2012). Por ello, se propone implementar una solución de BI que transforme los datos de CVP en información útil para optimizar la gestión comercial y apoyar el crecimiento de la empresa, con proyección a nivel nacional.

Objetivo general inicial

Implementar una solución de BI en CVP que integre y analice la información comercial de la empresa, con el fin de fortalecer la toma de decisiones estratégicas y maximizar las ventas a nivel nacional mediante el uso de indicadores y de análisis de datos en tiempo real.

Objetivos específicos iniciales

Para el cumplimiento del objetivo general se plantean los siguientes objetivos específicos:

1. Diagnosticar el estado actual de la gestión de la información comercial de la empresa, identificando las principales fuentes de datos, los procesos y las necesidades de información para la toma de decisiones.
2. Diseñar e implementar una plataforma de BI que consolide la información de ventas, clientes, inventarios y compras en un único sistema de análisis.
3. Desarrollar tableros de control (*dashboards*) e indicadores clave de desempeño (KPI) que permitan monitorear las ventas, la rentabilidad, el comportamiento de los clientes y el desempeño comercial.
4. Identificar oportunidades de crecimiento mediante el análisis de las tendencias de consumo, los productos de mayor rotación, los segmentos de clientes y el comportamiento del mercado.
5. Mejorar la calidad y la velocidad de la toma de decisiones comerciales, proporcionando a la gerencia información confiable, actualizada y disponible en tiempo real.

Áreas, objetivos y necesidades detectadas

Como parte del proceso de fortalecimiento organizacional y de crecimiento estratégico de CVP, resulta indispensable establecer objetivos claros para cada una de las áreas funcionales que participan directamente en el cumplimiento de la misión empresarial.

La definición de estos objetivos permite orientar los esfuerzos del personal, optimizar los recursos disponibles y mejorar la coordinación entre departamentos, con el propósito de incrementar la competitividad de la empresa. En este sentido, el presente informe desarrolla los objetivos de las áreas Comercial y de Operaciones, considerando los subprocesos que las integran y su importancia dentro de la cadena de valor de la organización.

Área Comercial

El objetivo del área Comercial es incrementar la participación de CVP en el mercado de materiales para la construcción mediante estrategias comerciales enfocadas en el crecimiento de las ventas, el fortalecimiento del posicionamiento de la marca, la fidelización de los clientes y una adecuada gestión del inventario, garantizando altos niveles de satisfacción y rentabilidad a nivel nacional.

Área de Operaciones

El objetivo del área de Operaciones es optimizar la gestión operativa mediante procesos eficientes de almacenamiento, recepción, conservación y despacho de materiales, garantizando el abastecimiento oportuno y la continuidad de las operaciones comerciales con altos estándares de calidad y seguridad.

Indicadores clave de desempeño (KPI)

Los indicadores clave de desempeño (*Key Performance Indicators*, KPI) son medidas que reflejan el grado de cumplimiento de los factores críticos de éxito de una organización

(Parmenter, 2015). Los KPI de CVP responden a la necesidad estratégica de implementar una solución de BI que permita integrar y analizar la información comercial de la empresa, fortalecer la toma de decisiones y maximizar las ventas a nivel nacional. Como referencia del contexto operativo y financiero en el que se aplican, se consideran los valores presentados en la Tabla 2. Los indicadores se organizan en tres grupos: los del área Comercial, los del área de Operaciones y los de carácter financiero y transversal. Cada uno se acompaña de su fórmula de cálculo, su frecuencia de medición y una meta sugerida.

Indicadores del área Comercial

La Tabla 3 presenta los nueve indicadores definidos para el área Comercial, orientados a medir el crecimiento de las ventas, el cumplimiento del presupuesto, la captación y la fidelización de los clientes, y la calidad del servicio.

Tabla 3

Indicadores clave de desempeño del área Comercial

KPI	Fórmula de cálculo	Frecuencia	Meta sugerida
Crecimiento porcentual de ventas	((Ventas del mes actual − Ventas del mes anterior) / Ventas del mes anterior) × 100	Mensual	≥ 5 %
Cumplimiento de presupuesto comercial	(Ventas reales / Presupuesto de ventas) × 100	Mensual	≥ 95 %
Número de clientes nuevos	Conteo de clientes nuevos registrados en el período	Mensual	≥ 50 por mes

Tasa de fidelización	(Clientes recurrentes / Total de clientes activos) × 100	Trimestral	≥ 60 %
Ticket promedio	Total de ventas / Número de transacciones	Mensual	≥ $280
Nivel de satisfacción del cliente	Promedio de las encuestas de satisfacción (escala de 1 a 5)	Mensual	≥ 4.2 sobre 5
Ventas corporativas (%)	(Ventas a constructoras y contratistas / Ventas totales) × 100	Mensual	≥ 25 %
Tasa de venta cruzada	(Transacciones con dos o más categorías / Total de transacciones) × 100	Mensual	≥ 30 %
Tasa de conversión de prospectos	(Clientes nuevos / Prospectos contactados) × 100	Mensual	≥ 15 %

Nota. Elaboración propia. Las metas corresponden a valores sugeridos por el equipo de investigación.

Indicadores del área de Operaciones

La Tabla 4 recoge los diez indicadores del área de Operaciones, que permiten evaluar la eficiencia de la recepción, el almacenamiento y el despacho de materiales, así como la seguridad en la bodega y la trazabilidad de los productos.

Tabla 4

Indicadores clave de desempeño del área de Operaciones

KPI	Fórmula de cálculo	Frecuencia	Meta sugerida

Tiempo promedio de recepción	Horas totales de recepción / Número de recepciones realizadas	Semanal	≤ 2 horas
Tiempo promedio de despacho	Horas entre el pedido confirmado y la entrega al cliente	Semanal	≤ 4 horas
Nivel de cumplimiento de pedidos	(Pedidos completos y a tiempo / Total de pedidos) × 100	Semanal	≥ 98 %
Productos dañados	(Unidades dañadas / Unidades manejadas) × 100	Mensual	≤ 0.5 %
Índice de accidentes en bodega	(Número de accidentes / Horas trabajadas) × 1,000,000	Mensual	0
Trazabilidad de productos	(Productos con registro completo / Total de productos) × 100	Mensual	≥ 99 %
Eficiencia de *picking*	(Líneas correctas / Total de líneas de *picking*) × 100	Semanal	≥ 99.5 %
Tasa de devoluciones por error	(Devoluciones por error operativo / Total de entregas) × 100	Mensual	≤ 1 %
Ocupación de espacio en bodega	(Espacio utilizado / Espacio total disponible) × 100	Mensual	75 % – 85 %
Cumplimiento de normas de seguridad	(Controles de seguridad aprobados / Total de controles) × 100	Mensual	≥ 95 %

Nota. Elaboración propia. Las metas corresponden a valores sugeridos por el equipo de investigación.

Indicadores financieros y transversales

Los indicadores financieros y transversales, presentados en la Tabla 5, permiten evaluar la salud financiera general de la empresa y el desempeño de áreas transversales como Talento Humano y Tesorería.

Tabla 5

Indicadores clave de desempeño financieros y transversales

KPI	Fórmula de cálculo	Frecuencia	Meta sugerida
Margen bruto	((Ventas − Costo de ventas) / Ventas) × 100	Mensual	≥ 20 %
Margen operativo	(Utilidad operativa / Ventas) × 100	Mensual	≥ 8 %
Rotación de cartera de clientes	Ventas a crédito / Cartera promedio de clientes	Mensual	≥ 4 veces por año
Días de inventario (DIO)	Inventario promedio / Costo de ventas diario	Mensual	≤ 60 días
Productividad por colaborador	Ventas totales / Número de colaboradores	Anual	≥ $159,580
Tasa de retención de empleados	((Empleados al final − Empleados nuevos) / Empleados al inicio) × 100	Anual	≥ 85 %

Rotación de personal	(Bajas en el período / Promedio de empleados) × 100	Anual	≤ 15 %
Costo laboral sobre ventas	(Costo total de nómina / Ventas totales) × 100	Mensual	≤ 12 %
Ciclo de caja	DIO + Días de cartera − Días de proveedores	Mensual	≤ 45 días
Rentabilidad sobre activos (ROA)	(Utilidad neta / Activos totales) × 100	Trimestral	≥ 10 %

Nota. Elaboración propia. Las metas corresponden a valores sugeridos. La meta de productividad por colaborador se obtuvo dividiendo las ventas totales entre el número de colaboradores (Tabla 2), por lo que se evalúa con frecuencia anual.

En síntesis, este capítulo caracteriza a CVP como una empresa con más de tres décadas de trayectoria, en crecimiento y con un volumen creciente de información, y establece el marco de objetivos e indicadores que justifica la implementación de una solución de BI. El capítulo siguiente describe cómo se origina, se carga y se almacena actualmente la información de la empresa, y cómo se propone organizarla en un almacén de datos.

CAPÍTULO 2

Origen de la información y recurrencia de carga: fuentes internas y externas

Fuentes internas

Las fuentes internas representan el origen principal de los datos utilizados por la empresa. La información es generada por las diferentes áreas funcionales, entre ellas Comercial, compras, inventarios, contabilidad y tesorería, mediante el registro de las operaciones diarias en el sistema de gestión SAGA. Cada proceso ejecutado dentro de la organización produce información que posteriormente se utiliza para el control de las actividades y para la generación de informes.

En el área Comercial se registra la información correspondiente a las ventas realizadas, incluidas la emisión de facturas, los datos de los clientes, los productos comercializados, las cantidades vendidas y los valores de cada transacción. De forma simultánea, el sistema actualiza el inventario disponible, de modo que las existencias reflejan de manera inmediata los movimientos generados por las ventas.

El área de compras constituye otra fuente importante de información interna. En ella se registran diariamente los ingresos de las mercancías adquiridas a los proveedores, con datos sobre productos, cantidades, costos y documentos de compra. Estos registros mantienen actualizado el inventario y facilitan el control del abastecimiento.

Por su parte, el área contable utiliza la información generada por las operaciones comerciales para registrar las transacciones financieras, las cuentas por cobrar, las cuentas por pagar y los demás movimientos económicos necesarios para el control administrativo.

Dado que todos estos procesos están integrados en SAGA, la información operativa y contable permanece centralizada y disponible para su consulta.

Fuentes externas

Las fuentes externas corresponden principalmente a la información proporcionada por los clientes y los proveedores. Los clientes generan datos durante los procesos de cotización, venta y facturación, mientras que los proveedores suministran información relacionada con órdenes de compra, facturas, listas de precios y documentos asociados al abastecimiento de materiales. Estos datos se incorporan a SAGA mediante los registros de cada transacción. Adicionalmente, la propuesta de BI contempla incorporar métricas de marketing digital provenientes de redes sociales, que actualmente no forman parte de la información gestionada en SAGA.

Recurrencia de la carga

En cuanto a la recurrencia de la carga de datos, CVP mantiene un esquema de actualización en tiempo real. Cada operación realizada dentro de la empresa se registra de inmediato en SAGA, en el momento en que ocurre, sin necesidad de procesos posteriores de consolidación manual. Este mecanismo permite que la información permanezca actualizada de forma continua y facilita el control de las operaciones, el seguimiento del inventario, la consulta de la información financiera y la generación de informes con datos oportunos y confiables.

La actualización inmediata constituye un elemento importante para la gestión de la empresa, ya que permite disponer de datos consistentes durante toda la jornada laboral, mejorar el seguimiento de las operaciones comerciales y proporcionar información confiable

para la elaboración de informes y para el apoyo a la toma de decisiones. Cabe precisar que la actualización en tiempo real se refiere al registro transaccional en SAGA; la extracción de los datos hacia el almacén de datos propuesto se programa con una frecuencia diaria, como se detalla en el proceso ETL.

Sistemas y destino de la carga

En CVP, el sistema utilizado para gestionar la información operativa es SAGA, herramienta principal para el registro y la administración de las actividades diarias de la empresa. A través de este sistema se procesan las operaciones de ventas, compras, inventarios y contabilidad, lo que permite que la información operativa permanezca centralizada en una sola plataforma. Esta integración facilita el control de los procesos y evita que los datos transaccionales se encuentren dispersos en diferentes sistemas.

El flujo de la información se inicia cuando cada área registra sus operaciones directamente en SAGA, como se ilustra en la Figura 1. El área Comercial ingresa las ventas realizadas a los clientes, genera la factura electrónica correspondiente y actualiza automáticamente el inventario. De igual manera, el área de compras registra el ingreso de la mercancía adquirida a los proveedores, lo que permite mantener un control permanente sobre las existencias y sobre los movimientos de abastecimiento. El área contable, por su lado, utiliza el mismo sistema para registrar y consultar la información financiera derivada de las operaciones comerciales.

Figura 1

Flujo de la información operativa y contable a través del sistema SAGA



Nota. Elaboración propia a partir de la información proporcionada por CVP Centro Constructor Vásquez Pacheco.

Una vez registrada, la información queda almacenada en la base de datos de SAGA, que se convierte en la fuente principal para la consulta de la información operativa y administrativa. De esta manera, los registros de ventas, compras, inventarios y contabilidad se concentran en un único sistema, lo que favorece la disponibilidad y la consistencia de los datos para las diferentes áreas de la organización.

El destino de la carga de datos corresponde principalmente al propio sistema SAGA, desde el cual se generan los informes que la empresa requiere. Estos informes pueden visualizarse directamente en el sistema o exportarse a Microsoft Excel cuando se necesita realizar análisis adicionales, elaborar indicadores específicos o presentar información

consolidada a la gerencia. La exportación a Excel permite complementar los reportes de SAGA con análisis personalizados, tablas dinámicas, gráficos y cálculos adicionales que apoyan el seguimiento de las ventas, las compras, los inventarios, las cuentas por cobrar, las cuentas por pagar y los resultados financieros. De este modo, Excel actúa como herramienta de apoyo para el análisis, mientras que SAGA continúa siendo la fuente oficial de los datos de la empresa.

En términos generales, el flujo de la información en CVP sigue una secuencia estructurada: las áreas registran sus operaciones en SAGA, el sistema almacena la información en su base de datos central y, posteriormente, los usuarios consultan los informes directamente en el sistema o los exportan a Excel para efectuar análisis más detallados y apoyar la toma de decisiones administrativas y estratégicas.

Almacenamiento en data warehouse y datamarts

Un almacén de datos es una colección de datos orientada a temas, integrada, no volátil y variante en el tiempo, que se utiliza para apoyar las decisiones de la gerencia (Inmon, 2005). Su diseño suele apoyarse en el modelado dimensional, en el cual la información se organiza en tablas de hechos, que registran los eventos medibles del negocio, y en tablas de dimensión, que describen el contexto de dichos eventos (Kimball y Ross, 2013). El diseño propuesto para CVP se basa en el análisis de las áreas Comercial y de Operaciones, en sus fuentes de datos, principalmente el sistema SAGA, y en los KPI definidos en el Capítulo 1 para el monitoreo del desempeño organizacional.

Arquitectura general del data warehouse

La arquitectura propuesta sigue un modelo centralizado, con un almacén de datos corporativo que alimenta datamarts especializados por área funcional. El sistema fuente principal es SAGA, que se complementa con datos de Excel para análisis adicionales y con métricas de marketing digital provenientes de redes sociales. La Figura 2 muestra el flujo de datos, desde los sistemas fuente hasta los tableros de control y los reportes.

Figura 2

Flujo de datos de la arquitectura propuesta



Nota. Elaboración propia. Los sistemas fuente comprenden SAGA (ventas, compras, inventarios, contabilidad, clientes y proveedores), archivos de Excel (análisis adicionales, indicadores personalizados y encuestas de satisfacción) y redes sociales (métricas de marketing digital).

Data warehouse corporativo

El almacén de datos corporativo centraliza la información histórica de toda la empresa con la mínima granularidad posible, lo que permite realizar análisis transversales y alimentar los datamarts especializados. La Tabla 6 presenta las dimensiones corporativas, junto con la frecuencia con la que se actualiza cada una.

Tabla 6

Dimensiones corporativas del data warehouse

Dimensión	Descripción	Frecuencia de actualización
DIM_TIEMPO	Fechas, períodos, temporadas y días hábiles	Diaria
DIM_CLIENTE	Clientes, segmentos, tipo (corporativo o minorista) y ubicación	Diaria
DIM_PRODUCTO	Productos, categorías, subcategorías, proveedor y código SKU	Diaria
DIM_PROVEEDOR	Proveedores, tipo, ubicación y condiciones comerciales	Diaria
DIM_EMPLEADO	Colaboradores, área, cargo y fecha de ingreso	Diaria
DIM_SUCURSAL	Sucursales (Cañar y El Tambo) y ubicación geográfica	Diaria
DIM_CANAL_VENTA	Canal de venta (físico, digital o telefónico)	Diaria
DIM_CATEGORIA_PRODUCTO	Categorías: construcción, ferretería, acabados y maquinaria	Diaria
DIM_METODO_PAGO	Efectivo, crédito, transferencia y tarjeta	Diaria
DIM_CAMPANA_MARKETING	Campañas publicitarias, tipo, canal y presupuesto	Por campaña

DIM_UBICACION_BODEGA	Zona, pasillo, estante y nivel de almacenamiento	Diaria

Nota. Elaboración propia. SKU = unidad de mantenimiento de existencias (stock keeping unit). La frecuencia indica el período máximo de refresco; cuando la fuente se actualiza con menor periodicidad, la carga incorpora únicamente los cambios disponibles.

Datamarts especializados

Los datamarts son subconjuntos del almacén de datos orientados a áreas específicas del negocio. Cada uno contiene las dimensiones y los hechos relevantes para sus KPI, organizados para responder con rapidez a las consultas analíticas (Kimball y Ross, 2013). Se proponen dos datamarts: uno comercial y otro de bodega.

Datamart comercial (DM_COMERCIAL)

El datamart comercial está orientado al análisis de las ventas, los clientes, el comportamiento de compra y el desempeño comercial, y soporta los KPI de la subárea de Ventas y Marketing. Las dimensiones incluidas se presentan en la Tabla 7, y las tablas de hechos, en la Tabla 8.

Tabla 7

Dimensiones del datamart comercial

Dimensión	Descripción en el contexto comercial
DIM_TIEMPO	Tiempo de análisis (día, semana, mes, trimestre y año)
DIM_CLIENTE	Perfil de los clientes, segmentación y fidelización

DIM_PRODUCTO	Catálogo de productos, categorías y rotación
DIM_EMPLEADO	Vendedores y asesores comerciales
DIM_SUCURSAL	Ubicación de la venta
DIM_CANAL_VENTA	Canal por el que se realizó la venta
DIM_CAMPANA_MARKETING	Campañas asociadas a las ventas

Nota. Elaboración propia con base en el análisis de las áreas Comercial y de Operaciones.

Tabla 8

Tablas de hechos del datamart comercial

Tabla de hechos	Descripción
HECHOS_VENTAS_RESUMEN	Resumen diario de ventas por cliente, producto y vendedor
HECHOS_CLIENTES	Comportamiento de los clientes: frecuencia, recencia y ticket
HECHOS_MARKETING	Impacto de las campañas en las ventas y en la captación de clientes
HECHOS_VENTAS_CRUZADAS	Análisis de ventas con múltiples categorías
HECHOS_VENTAS_CORPORATIVAS	Ventas a constructoras y contratistas

Nota. Elaboración propia con base en el análisis de las áreas Comercial y de Operaciones.

El datamart comercial soporta diecisiete KPI. Nueve de ellos corresponden a los indicadores del área Comercial definidos en la Tabla 3. Los ocho restantes, relacionados con el marketing digital, se detallan en la Tabla 9.

Tabla 9

Indicadores de marketing soportados por el datamart comercial

KPI	Fórmula o definición
Alcance en redes sociales	Total de personas que visualizan el contenido publicado
Tráfico hacia medios digitales	Visitas al sitio web y a los perfiles sociales
Conversión de campañas	(Clientes captados por la campaña / Alcance de la campaña) × 100
Reconocimiento de marca	Porcentaje de encuestados que reconocen la marca en la provincia
Retorno de la inversión en marketing (ROI)	((Ingresos generados − Costo de la campaña) / Costo de la campaña) × 100
Costo de adquisición de cliente (CAC)	Gasto en marketing / Clientes nuevos obtenidos
Tasa de interacción (*engagement rate*)	(Interacciones totales / Alcance) × 100
Tasa de crecimiento de seguidores	((Seguidores actuales − Seguidores anteriores) / Seguidores anteriores) × 100

Nota. Elaboración propia con base en el análisis de las áreas Comercial y de Operaciones.

Datamart de bodega (DM_BODEGA)

El datamart de bodega está orientado al control de inventarios, a los movimientos de bodega, al abastecimiento y a la gestión operativa del almacén, y soporta los KPI de la bodega comercial y del área de Operaciones. Sus dimensiones se presentan en la Tabla 10, y sus tablas de hechos, en la Tabla 11.

Tabla 10

Dimensiones del datamart de bodega

Dimensión	Descripción en el contexto de bodega
DIM_TIEMPO	Períodos de análisis del inventario
DIM_PRODUCTO	Productos en inventario, categorías y rotación
DIM_PROVEEDOR	Proveedores de la mercadería
DIM_SUCURSAL	Bodega o sucursal donde se almacena la mercadería
DIM_EMPLEADO	Personal de bodega responsable
DIM_UBICACION_BODEGA	Zona, pasillo, estante y nivel de almacenamiento

Nota. Elaboración propia con base en el análisis de las áreas Comercial y de Operaciones.

Tabla 11

Tablas de hechos del datamart de bodega

Tabla de hechos	Descripción
HECHOS_INVENTARIO_DIARIO	Registro diario de los niveles de inventario por producto y ubicación
HECHOS_MOVIMIENTOS	Entradas, salidas, transferencias y ajustes de inventario

HECHOS_RECEPCIONES	Recepciones de mercadería de los proveedores
HECHOS_DESPACHOS	Despachos de pedidos a los clientes
HECHOS_PICKING	Operaciones de *picking* y preparación de pedidos
HECHOS_PERDIDAS	Pérdidas por deterioro, daño, robo o vencimiento
HECHOS_CONTROLES_SEGURIDAD	Controles de seguridad realizados en la bodega
HECHOS_ACCIDENTES	Registro de accidentes laborales en la bodega

Nota. Elaboración propia con base en el análisis de las áreas Comercial y de Operaciones.

El datamart de bodega soporta dieciocho KPI. Diez de ellos corresponden a los indicadores del área de Operaciones definidos en la Tabla 4. Los ocho restantes, relativos al control de inventarios, se detallan en la Tabla 12.

Tabla 12

Indicadores de inventario soportados por el datamart de bodega

KPI	Fórmula o definición
Nivel de disponibilidad de inventario	(SKU disponibles / Total de SKU en el catálogo) × 100
Rotación de inventarios	Costo de ventas del período / Inventario promedio del período
Exactitud del inventario	(1 − (Σ|Inventario físico − Inventario en sistema| / Σ Inventario en sistema)) × 100
Tiempo promedio de reposición	Días entre el pedido de compra y la recepción de la mercadería
Productos agotados	(Número de SKU sin existencias / Total de SKU) × 100

Pérdidas por inventario	(Valor de las pérdidas / Valor total del inventario) × 100
Índice de cobertura de existencias	Días de inventario disponible según la demanda promedio
Tasa de productos próximos a vencer	(Productos con vencimiento en menos de 30 días / Total del inventario) × 100

Nota. Elaboración propia con base en el análisis de las áreas Comercial y de Operaciones.

Proceso ETL (extracción, transformación y carga)

El proceso ETL garantiza la calidad, la consistencia y la disponibilidad de los datos en el almacén de datos y en los datamarts. Consiste en extraer los datos de los sistemas fuente, transformarlos en un área de preparación (*staging area*) y cargarlos en las estructuras dimensionales de destino (Kimball y Ross, 2013). Las fuentes de datos consideradas, los datos que se extraen de cada una, su frecuencia y su prioridad se presentan en la Tabla 13.

Tabla 13

Fuentes de datos del proceso ETL

Fuente	Datos extraídos	Frecuencia	Prioridad
SAGA – Módulo Ventas	Facturas electrónicas, notas de venta y devoluciones	Diaria	Alta
SAGA – Módulo Compras	Órdenes de compra, recepciones y facturas de proveedores	Diaria	Alta
SAGA – Módulo Inventarios	Movimientos de existencias, ajustes y transferencias	Diaria	Alta

SAGA – Módulo Contabilidad	Asientos contables y estados financieros	Diaria	Alta
SAGA – Módulo Clientes	Maestro de clientes, crédito y cobranza	Diaria	Alta
Excel – Marketing	Métricas de redes sociales, campañas y encuestas	Semanal	Media
Excel – Talento Humano	Nómina, asistencia, accidentes y evaluaciones	Mensual	Media
Sistema BI (registros de uso)	Uso del sistema, tiempos de respuesta y errores	Diaria	Baja

Nota. Elaboración propia con base en el análisis de las áreas Comercial y de Operaciones.

El proceso se desarrolla en tres fases, que se describen en la Tabla 14. La conexión con SAGA se realizará directamente a su base de datos, ya sea SQL Server o PostgreSQL, según el motor que utilice la instalación de la empresa. Las métricas de redes sociales se obtendrán mediante interfaces de programación (API) y se complementarán con los archivos de Excel del área de Marketing. Para preservar el historial de los cambios en las dimensiones se aplicará la técnica de dimensiones de cambio lento de tipo 2, que registra cada modificación como una nueva fila (Kimball y Ross, 2013).

Tabla 14

Fases del proceso ETL

Fase	Actividades

1. Extracción	Conexión directa con la base de datos de SAGA; exportación programada de archivos de Excel; consulta de las API de redes sociales para las métricas de marketing.
2. Transformación	Limpieza (eliminación de duplicados y corrección de formatos); normalización (estandarización de códigos, nombres y unidades); enriquecimiento (cálculo de métricas derivadas, como margen y rotación); validación (reglas de calidad de datos: completitud, consistencia y exactitud); historización (dimensiones de cambio lento de tipo 2).
3. Carga	Carga inicial con la población histórica del almacén de datos (de tres a cinco años de información); carga incremental diaria de hechos y dimensiones, con excepción de la dimensión de campañas, que se actualiza por campaña; refresco de los datamarts según la prioridad de la fuente; validación posterior a la carga (integridad referencial y totales de control).

Nota. Elaboración propia con base en el análisis de las áreas Comercial y de Operaciones.

Carencias y problemas en la carga de datos

El análisis del esquema actual de carga y uso de la información permite identificar nueve carencias, que se agrupan en cuatro ejes.

Calidad de los datos y dependencia del ingreso manual

La primera carencia es la dependencia del ingreso manual de la información. Cada área registra sus operaciones directamente en SAGA, por lo que cualquier error humano al digitar los datos puede afectar la calidad de la información y de los análisis posteriores.

Orientación operativa de SAGA y dependencia de Excel

La segunda carencia es que la información está disponible principalmente para la operación y no para el análisis. SAGA permite administrar las operaciones diarias, pero no está orientado al análisis avanzado del comportamiento de las ventas, de las tendencias de los clientes ni de las proyecciones comerciales. La tercera carencia deriva de la anterior: la empresa depende de Microsoft Excel para los análisis adicionales, ya que debe exportar la información desde SAGA para crear indicadores, tablas dinámicas y gráficos. Este procedimiento incrementa el tiempo de procesamiento y el riesgo de errores. A ello se suma una cuarta carencia, el riesgo de duplicidad y de versiones diferentes de los archivos, pues el trabajo con múltiples archivos de Excel puede generar varias versiones del mismo reporte y, en consecuencia, diferencias en la información utilizada por las distintas áreas.

Integración limitada y visualización insuficiente

La quinta carencia es la falta de integración con otras fuentes de información. SAGA concentra los datos internos de ventas, compras, inventarios y contabilidad, pero no integra datos de otras fuentes estratégicas, como las campañas de marketing, las redes sociales, el comportamiento del mercado o los indicadores comerciales nacionales. La sexta carencia es la demora en la generación de información para la gerencia: al requerir exportaciones y análisis manuales, la obtención de reportes estratégicos puede retrasarse, lo que dificulta una respuesta rápida ante los cambios del mercado. La séptima es la limitada visualización de los KPI, dado que no se dispone de tableros de control en tiempo real que permitan monitorear indicadores como las ventas por sucursal, el margen de utilidad, la rotación de inventarios, el desempeño de los vendedores o los productos más vendidos.

Capacidad analítica y expansión nacional

La octava carencia es la escasa capacidad para realizar análisis históricos y predictivos. Aunque la información se almacena en SAGA, no se aprovecha mediante herramientas de BI que permitan identificar patrones de compra, estacionalidad o proyecciones de ventas futuras. Finalmente, la novena carencia es la dificultad para apoyar la expansión nacional. El proyecto busca maximizar las ventas a nivel nacional; sin embargo, la forma actual de gestionar la información limita la obtención de una visión integral y oportuna para planificar el crecimiento comercial y tomar decisiones estratégicas basadas en datos.

Informes generados actualmente, necesidades detectadas y necesidades no cubiertas

La información constituye uno de los activos más importantes para la gestión empresarial, pues permite controlar las operaciones, evaluar el desempeño y apoyar la toma de decisiones estratégicas. De acuerdo con el documento de referencia del programa, las organizaciones generan diferentes tipos de reportes (operativos, corporativos, comerciales, legales y analíticos), según las necesidades de cada área.

CVP es una empresa dedicada a la comercialización de materiales para la construcción, ferretería y acabados, con 9,000 clientes activos y una estructura organizacional conformada por las áreas de Dirección, Talento Humano, Comercial, Tesorería, Operaciones, Calidad y Expansión. En este contexto, la disponibilidad de información confiable resulta indispensable para garantizar la eficiencia operativa y el crecimiento del negocio. Si bien la empresa genera información para el desarrollo de sus actividades diarias, existen oportunidades para fortalecer el análisis de datos mediante reportes gerenciales e indicadores que faciliten una administración basada en información oportuna.

Informes que se generan actualmente

Por la naturaleza operativa y comercial de CVP, los informes que la empresa genera actualmente son principalmente de tipo operativo y comercial, y permiten controlar las actividades diarias. La Tabla 15 resume su contenido y su finalidad.

Tabla 15

Informes generados actualmente por la empresa

Informe	Contenido	Finalidad
Reporte diario de ventas	Facturación diaria por vendedor, sucursal o punto de atención; número de clientes atendidos; productos vendidos; formas de pago; valor promedio por compra; cumplimiento de metas comerciales	Evaluar el comportamiento comercial y controlar el desempeño del personal de ventas
Reporte de inventarios	Existencias disponibles; productos agotados; productos con baja rotación; ingresos de mercadería; salidas de bodega; ajustes de inventario; inventarios físicos	Garantizar la disponibilidad de materiales para atender la demanda comercial
Reporte de compras	Órdenes de compra; proveedores; recepción de mercadería; costos de adquisición; cumplimiento de entregas; productos pendientes de recibir	Mantener abastecida a la empresa y controlar la relación con los proveedores estratégicos
Reporte financiero	Ventas mensuales; compras; flujo de caja; ingresos; egresos; cuentas por cobrar; cuentas por pagar	Controlar la liquidez y apoyar la planificación financiera

Reporte de despacho y bodega	Pedidos preparados y entregados; tiempos de despacho; productos dañados; devoluciones; incidentes de almacenamiento	Evaluar la eficiencia logística y el cumplimiento de las entregas
Reportes comerciales básicos	Clientes nuevos y frecuentes; ventas por categoría y por vendedor; productos más vendidos y de menor rotación	Dar seguimiento a la actividad comercial y apoyar las decisiones del área de ventas

Nota. Elaboración propia a partir de la información proporcionada por CVP Centro Constructor Vásquez Pacheco.

Necesidades detectadas

A partir del análisis del proyecto y de los objetivos planteados para las áreas Comercial y de Operaciones, se identifican tres necesidades de información que actualmente requieren fortalecerse.

Implementación de indicadores de desempeño. Aunque la empresa dispone de información operativa, es necesario consolidarla en indicadores que permitan evaluar el cumplimiento de los objetivos estratégicos. Entre los principales se encuentran el cumplimiento de las metas comerciales, la rotación de inventario, el nivel de servicio al cliente, el tiempo promedio de despacho, la exactitud del inventario y la rentabilidad por línea de productos. Estos indicadores se vinculan con los KPI definidos en las Tablas 3, 4 y 5 y están alineados con los objetivos de las áreas Comercial y de Operaciones. Si bien algunos informes actuales incorporan datos relacionados con ellos, como el cumplimiento de metas en el reporte diario de ventas, estos no se encuentran consolidados en un sistema formal de KPI con fórmulas, frecuencias y metas definidas.

Integración de la información entre áreas. La información generada por ventas, bodega, compras y operaciones se procesa de forma independiente en cada área y, además, los análisis complementarios que se realizan fuera de SAGA, como los de Excel, no se consolidan en un repositorio común. Por ello, es necesario integrar todos los procesos para que la Dirección disponga de información consolidada y en tiempo real. Esta integración permitirá una mejor planificación de las compras, una reducción de los sobreinventarios, una mejor coordinación entre ventas y bodega, y una mayor rapidez en la atención al cliente.

Reportes gerenciales. La alta dirección requiere reportes ejecutivos que resuman el comportamiento general del negocio. Estos reportes deberían incluir las ventas por sucursal, la rentabilidad por línea de negocio, los comparativos mensuales, la evolución de los clientes, el cumplimiento de los presupuestos y los indicadores estratégicos. Este tipo de informes corresponde a la categoría de reportes corporativos descrita en el documento de referencia del programa.

Necesidades no cubiertas

A pesar de que CVP dispone de información operativa para el desarrollo de sus actividades, aún existen seis necesidades estratégicas que no se encuentran completamente cubiertas.

Tablero de control gerencial. Actualmente no se evidencia la existencia de un tablero de control integral que concentre los principales indicadores del negocio. La implementación de un tablero permitiría visualizar en tiempo real las ventas, los inventarios, las compras, la rentabilidad, los despachos, el cumplimiento de los objetivos y los indicadores comerciales.

Inteligencia de negocios. La empresa aún no dispone de herramientas de BI que transformen grandes volúmenes de datos en conocimiento útil para la toma de decisiones. Su implementación permitiría detectar tendencias de ventas, identificar los productos de mayor crecimiento, pronosticar la demanda, optimizar los niveles de inventario y analizar la rentabilidad por cliente.

Análisis predictivo. Actualmente, las decisiones relacionadas con las compras y los inventarios se toman principalmente con base en la experiencia y en la demanda histórica. La empresa podría incorporar modelos predictivos para estimar las ventas futuras, anticipar los productos de alta demanda, reducir los quiebres de inventario, planificar las campañas comerciales y optimizar las compras. Esta necesidad coincide con el enfoque de analítica y minería de datos descrito en el documento de referencia del programa.

Automatización de reportes. La consolidación de gran parte de la información requiere procesos manuales. La automatización de los reportes permitiría reducir los tiempos de elaboración, evitar los errores humanos, disponer de información actualizada y mejorar la confiabilidad de los datos.

Sistema integrado de información. Se identifica la necesidad de contar con una plataforma que integre todas las áreas estratégicas de la empresa (Dirección, Talento Humano, Comercial, Tesorería, Operaciones, Calidad y Expansión), incluidos los procesos de compras y de marketing. La integración facilitaría la trazabilidad de la información, eliminaría la duplicidad de registros y mejoraría la coordinación entre los departamentos.

Indicadores de satisfacción del cliente. Actualmente no se evidencia un sistema estructurado para medir de forma permanente la satisfacción del cliente. Se recomienda

implementar reportes sobre el nivel de satisfacción, el tiempo de respuesta, los reclamos, las devoluciones, la fidelización y la recomendación del servicio, esta última medida mediante el *Net Promoter Score* (NPS).

En síntesis, CVP registra sus operaciones en tiempo real en SAGA, que constituye la fuente oficial y centralizada de sus datos transaccionales. No obstante, el análisis de esa información depende de exportaciones manuales a Excel y no existe una capa analítica que integre las distintas fuentes ni que ofrezca indicadores y tableros actualizados. La arquitectura de almacén de datos y datamarts, junto con el proceso ETL descrito, constituye la base técnica para cubrir las necesidades identificadas y alcanzar los objetivos planteados en el Capítulo 1.

CAPÍTULO 3

Procesos de carga de datos (ETL, Data Lake)

Este capítulo detalla los componentes del proceso ETL (Extract, Transform, Load), la arquitectura del Data Warehouse, los Data Marts especializados, las fuentes de datos, el flujo de información y las políticas de gobernanza y calidad de datos.

Arquitectura general del Data Warehouse

La arquitectura propuesta sigue un modelo centralizado: un Data Warehouse corporativo alimenta Data Marts especializados por área funcional. El sistema fuente principal es SAGA, complementado con archivos de Excel para análisis adicionales y con métricas de redes sociales.

Flujo de datos. La información circula en la secuencia sistemas fuente, staging area, Data Warehouse, Data Marts y dashboards o reportes.

Sistemas fuente. SAGA, el sistema de planificación de recursos empresariales (ERP, por sus siglas en inglés) principal, aporta datos de ventas, compras, inventarios, contabilidad, clientes y proveedores. Excel aporta análisis adicionales, indicadores personalizados y encuestas de satisfacción, y las redes sociales, métricas de marketing digital.

Data Warehouse comercial

El Data Warehouse corporativo centraliza la información histórica de la empresa con el mínimo nivel de granularidad, lo que permite análisis transversales y alimenta los Data Marts especializados.

Dimensiones corporativas (tablas de dimensión). La Tabla 16 presenta las dimensiones corporativas, su descripción y su frecuencia de actualización.

Tabla 16

Dimensiones corporativas del Data Warehouse

Dimensión	Descripción	Frecuencia de actualización
DIM_TIEMPO	Fechas, períodos, temporadas y días hábiles	Diaria
DIM_CLIENTE	Clientes, segmentos, tipo (corporativo o retail) y ubicación	Diaria
DIM_PRODUCTO	Productos, categorías, subcategorías, proveedor y SKU	Diaria
DIM_PROVEEDOR	Proveedores, tipo, ubicación y condiciones comerciales	Diaria
DIM_EMPLEADO	Colaboradores, área, cargo y fecha de ingreso	Diaria
DIM_SUCURSAL	Sucursales (Cañar y El Tambo) y ubicación geográfica	Diaria
DIM_CANAL_VENTA	Canal de venta (físico, digital o telefónico)	Diaria
DIM_CATEGORIA_PRODUCTO	Categorías: construcción, ferretería, acabados y maquinaria	Diaria
DIM_METODO_PAGO	Efectivo, crédito, transferencia y tarjeta	Diaria
DIM_CAMPANA_MARKETING	Campañas publicitarias, tipo, canal y presupuesto	Por campana

Nota. SKU = unidad de mantenimiento de existencias (stock keeping unit). Elaboración propia.

Data Marts especializados

Los Data Marts son subconjuntos del Data Warehouse orientados a áreas específicas del negocio. Cada uno contiene las dimensiones y los hechos necesarios para sus indicadores clave de desempeño (KPI) y está optimizado para consultas analíticas rápidas.

Data Mart comercial (DM_COMERCIAL). Este Data Mart se orienta al análisis de ventas, clientes, comportamiento de compra y desempeño comercial, y soporta los KPI de la subárea de Ventas y Marketing.

Dimensiones incluidas. DIM_TIEMPO, para el análisis por día, semana, mes, trimestre y año; DIM_CLIENTE, con el perfil, la segmentación y la fidelización de los clientes; DIM_PRODUCTO, con el catálogo, las categorías y la rotación; DIM_EMPLEADO, que identifica a vendedores y asesores comerciales; DIM_SUCURSAL, con la ubicación de la venta; DIM_CANAL_VENTA, con el canal por el que se realizó la venta; y DIM_CAMPANA_MARKETING, con las campañas asociadas a las ventas.

Data Mart de bodega (DM_BODEGA). Este Data Mart se orienta al control de inventarios, los movimientos de bodega, el abastecimiento y la gestión operativa del almacén, y soporta los KPI de Bodega Comercial y de Operaciones.

Dimensiones incluidas. DIM_TIEMPO, con los períodos de análisis de inventario; DIM_PRODUCTO, con los productos en inventario, las categorías y la rotación; DIM_PROVEEDOR, con los proveedores de mercadería; DIM_SUCURSAL, con la bodega o sucursal donde se almacena; DIM_EMPLEADO, con el personal de bodega responsable; y

DIM_UBICACION_BODEGA, con la zona, el pasillo, el estante y el nivel de almacenamiento.

Proceso ETL (Extract, Transform, Load)

El proceso ETL garantiza la calidad, la consistencia y la disponibilidad de los datos en el Data Warehouse y en los Data Marts. Consta de tres fases: extracción, transformación y carga.

Fuentes de datos. La Tabla 17 resume las fuentes de datos identificadas, los datos que se extraen de cada una, la frecuencia de extracción y la prioridad asignada.

Tabla 17

Fuentes de datos del proceso ETL

Fuente	Datos extraídos	Frecuencia	Prioridad
SAGA, módulo de ventas	Facturas electrónicas, notas de venta y devoluciones	Diaria	Alta
SAGA, módulo de compras	Órdenes de compra, recepciones y facturas de proveedores	Diaria	Alta
SAGA, módulo de inventarios	Movimientos de stock, ajustes y transferencias	Diaria	Alta
SAGA, módulo de contabilidad	Asientos contables y estados financieros	Diaria	Alta
SAGA, módulo de clientes	Maestro de clientes, crédito y cobranza	Diaria	Alta

Excel, marketing	Métricas de redes sociales, campañas y encuestas	Semanal	Media
Excel, recursos humanos	Nómina, asistencia, accidentes y evaluaciones	Mensual	Media
Sistema de BI (logs)	Uso del sistema, tiempos de respuesta y errores	Diaria	Baja

Nota. BI = Business Intelligence. Elaboración propia.

Flujo ETL por fase. Cada fase del proceso comprende las actividades que se describen a continuación.

Fase 1: extracción. La extracción se realiza mediante conexión directa a la base de datos de SAGA (SQL Server o PostgreSQL), exportación programada de archivos de Excel y consulta de las interfaces de programación de aplicaciones (API) de redes sociales para obtener las métricas de marketing.

Fase 2: transformación. Comprende cinco procesos: limpieza (eliminación de duplicados y corrección de formatos), normalización (estandarización de códigos, nombres y unidades), enriquecimiento (cálculo de métricas derivadas, como margen y rotación), validación (reglas de calidad de datos: completitud, consistencia y exactitud) e historización (manejo de cambios lentos mediante dimensiones de cambio lento [SCD] de tipo 2).

Fase 3: carga. La carga inicial puebla el Data Warehouse con el histórico de los últimos 3 a 5 años. Luego, la carga incremental actualiza diariamente los hechos y las dimensiones (la dimensión de campañas se actualiza por campaña) y los Data Marts se

refrescan de forma programada según su prioridad. Por último, la validación posterior a la carga verifica la integridad referencial y los totales.

Gobernanza y calidad de datos y accesibilidad

Se establecen políticas y controles que garantizan la confiabilidad de la información durante todo su ciclo de vida en el ecosistema de Business Intelligence (BI). La Tabla 18 define las dimensiones de calidad y sus metas.

Tabla 18

Dimensiones de calidad de datos y metas

Dimensión de calidad	Definición	Meta
Completitud	Todos los campos obligatorios tienen valor	≥ 99 %
Exactitud	Los datos coinciden con la fuente oficial (SAGA)	≥ 98 %
Consistencia	Formatos y códigos estandarizados en todo el Data Warehouse	100 %
Actualidad	Datos disponibles con un retraso máximo de 24 horas	≤ 24 h
Trazabilidad	Registro del origen, las transformaciones y el responsable	100 %
Seguridad	Acceso controlado por roles y niveles de sensibilidad	100 %

Nota. Elaboración propia.

Horarios de carga y accesibilidad

La información de reportes y bases de datos se gestiona mediante una carga incremental diaria. Cada día a las 8:00 a. m., el asistente contable extrae y descarga los reportes de las operaciones del día anterior (ventas, cartera, compras, proveedores e inventarios). Este horario se establece porque, al finalizar la jornada laboral, no se generan nuevos movimientos contables.

Tras la extracción, el contador revisa y valida los reportes, y verifica que los datos coincidan y estén conciliados con el sistema contable SAGA. Como etapa final, el gerente revisa y aprueba la información para asegurar su exactitud, integridad y consistencia antes de ponerla a disposición de los departamentos. Una vez aprobada, la información queda disponible para los procesos operativos, administrativos y de análisis de la organización.

Además, cada semana se revisa y concilia la información con SAGA para garantizar su calidad, exactitud e integridad, y para asegurar que esté lista para generar y actualizar los KPI. La Figura 3 sintetiza los flujos de trabajo diario y semanal.

Flujo de trabajo semanal y mensual

Figura 3

Flujos de trabajo diario y semanal para la extracción, validación y disponibilidad de la información



Nota. Elaboración propia.

Variables disponibles y acceso al Data Warehouse

Las variables disponibles para el análisis en CVP Centro Constructor Vásquez Pacheco (en adelante, CVP) provienen de SAGA, el sistema de gestión empresarial que constituye la fuente principal de información operativa, administrativa y financiera de la empresa. Al integrar en una sola plataforma los procesos de ventas, compras, inventarios y contabilidad, el sistema ofrece variables que permiten dar seguimiento a las actividades comerciales y apoyar la toma de decisiones.

En el área comercial, en SAGA se registra el número de factura, la fecha de emisión, la identificación del cliente, el código y la descripción del producto, la cantidad vendida, el precio unitario, los descuentos, los impuestos y el valor total de la transacción. Estas

variables permiten analizar el comportamiento de las ventas, identificar los productos de mayor rotación, evaluar el desempeño comercial y comparar períodos.

En el módulo de inventarios, el sistema almacena el código, la descripción, la categoría y la ubicación del producto, así como las existencias, el costo, el precio de venta y los movimientos de ingreso y salida. Estas variables facilitan el control de inventarios, el seguimiento del abastecimiento y la planificación de compras, y permiten mantener existencias adecuadas para atender la demanda.

En compras, SAGA registra proveedores, órdenes de compra, facturas de adquisición, fechas de ingreso, cantidades, costos unitarios y valores totales. Estas variables permiten controlar el abastecimiento, evaluar las adquisiciones y dar seguimiento a la relación con los proveedores.

En el área contable y financiera, el sistema dispone de variables sobre cuentas por cobrar, cuentas por pagar, ingresos, egresos y movimientos contables. Su integración facilita la elaboración de informes financieros y el control de la situación económica de la organización.

Actualmente, la empresa no cuenta con un Data Warehouse para el almacenamiento de información analítica, por lo que el acceso a las variables se realiza directamente desde SAGA, fuente oficial de información de todas las áreas. Los usuarios autorizados consultan los datos mediante los informes del sistema, según los permisos de cada perfil. Cuando se requieren análisis más específicos o informes gerenciales, la información se exporta a Microsoft Excel, donde se organiza, filtra y consolida. Excel complementa el análisis, pero no sustituye a SAGA, que continúa como repositorio principal de los datos.

Desde la perspectiva de BI, un Data Warehouse permitiría consolidar la información histórica de SAGA en un entorno especializado, integrar los datos de las distintas áreas y facilitar el desarrollo de indicadores estratégicos, tableros de control y análisis multidimensionales. Si bien el esquema actual (SAGA y exportaciones a Excel) satisface las necesidades operativas inmediatas, limita el análisis histórico e integrado que requiere la dirección.

Sistemas implicados en los procesos y recursos IT o externos empleados

La implementación de una solución de BI en CVP requiere identificar los sistemas que intervienen en los procesos de negocio y los recursos tecnológicos que soportarán la solución. La empresa utiliza actualmente un sistema ERP para gestionar sus operaciones comerciales y administrativas; la estrategia de BI incorpora herramientas que permitirán integrar, transformar, almacenar y analizar grandes volúmenes de información para sustentar decisiones basadas en datos.

Sistemas implicados en los procesos.

Los procesos de CVP abarcan las áreas comercial, de compras, inventarios, operaciones, contabilidad, tesorería y dirección. Cada una genera información que utilizará la plataforma de BI, cuya integración permitirá disponer de información centralizada, confiable y actualizada diariamente para apoyar la gestión estratégica.

Sistema ERP SAGA. SAGA es el principal sistema implicado y la fuente oficial de información de la empresa. Administra las operaciones diarias y concentra la información de ventas y facturación electrónica, compras y abastecimiento, inventarios, clientes,

proveedores, contabilidad, cuentas por cobrar y cuentas por pagar. Registra cada transacción en tiempo real, lo que mantiene actualizados los inventarios y la información financiera, y sirve como sistema fuente para extraer los datos que alimentarán el Data Warehouse corporativo.

Microsoft Excel. Aunque SAGA concentra la información operativa, la empresa utiliza Excel como herramienta complementaria para elaborar reportes personalizados, tablas dinámicas, indicadores específicos, análisis financieros, consolidaciones manuales y gráficos. En el proyecto de BI, Excel seguirá siendo una fuente secundaria, principalmente para indicadores de marketing, encuestas de satisfacción y análisis específicos aún no integrados al ERP.

Data Warehouse corporativo. La propuesta incorpora un Data Warehouse como repositorio central de información histórica, que integrará datos de SAGA, archivos de Excel, redes sociales y sistemas de BI. Almacenará los datos organizados por dimensiones (tiempo, clientes, productos, proveedores, empleados, sucursales y campañas de marketing), lo que facilita los análisis multidimensionales y las consultas de alto rendimiento.

Data Marts. A partir del Data Warehouse se implementarán Data Marts especializados que responden a las necesidades analíticas de cada área.

Data Mart comercial. Permite analizar ventas, clientes, ticket promedio, fidelización, marketing, ventas corporativas y venta cruzada. Su finalidad es apoyar las decisiones comerciales mediante indicadores estratégicos y dashboards especializados.

Data Mart de bodega. Concentra información sobre inventarios, movimientos de bodega, recepciones, despachos, picking, productos dañados y seguridad, y permite optimizar la gestión operativa y el control de inventarios mediante indicadores de desempeño logístico.

Plataforma de Business Intelligence. La plataforma de BI es el sistema analítico principal del proyecto. Sus funciones incluyen la integración automática de datos, la construcción de dashboards ejecutivos, el seguimiento de KPI, los reportes gerenciales, los análisis históricos y predictivos, los pronósticos de ventas y la visualización de indicadores actualizados según la frecuencia de carga. Así, transforma los datos operativos en información estratégica para la dirección.

Redes sociales y herramientas de marketing. Como fuentes externas, el proyecto integrará mediante API indicadores de redes sociales, como el alcance de las publicaciones, las interacciones, la conversión de campañas, el tráfico digital y el reconocimiento de marca. Estos datos complementarán la información comercial almacenada en el Data Warehouse.

Recursos IT o externos empleados

La solución de BI requiere una infraestructura tecnológica compuesta por recursos internos de tecnología de la información (IT) y servicios externos especializados.

Recursos tecnológicos internos (IT). Comprenden la infraestructura de servidores, la base de datos, las herramientas ETL, la plataforma de BI, los equipos informáticos y los mecanismos de seguridad.

Infraestructura de servidores. Se requieren servidores para la base de datos de SAGA, el Data Warehouse, la plataforma de BI, el almacenamiento histórico y los procesos

ETL. Deben garantizar disponibilidad, capacidad de procesamiento y seguridad de la información.

Base de datos. La arquitectura contempla una base de datos relacional (por ejemplo, SQL Server o PostgreSQL) que almacenará la información histórica, las dimensiones, las tablas de hechos, los KPI y los datos consolidados. Esta base será el núcleo del sistema analítico.

Herramientas ETL. Los procesos ETL automatizarán la integración de datos: extraerán información de SAGA, importarán archivos de Excel, obtendrán datos desde API, eliminarán registros duplicados, normalizarán formatos, validarán la calidad de los datos y cargarán la información al Data Warehouse. Así se garantiza que los análisis se basen en información consistente y confiable.

Plataforma BI. La empresa requerirá una herramienta especializada de visualización y análisis que es Power BI para construir dashboards, generar reportes interactivos, configurar alertas automáticas y realizar análisis multidimensionales. La utilizarán la gerencia y los responsables de cada área funcional.

Equipos informáticos. El proyecto utilizará la infraestructura existente: los computadores de los usuarios, la gerencia, el área comercial y las áreas administrativas, la conectividad de red y el acceso seguro al sistema.

Seguridad informática. Se implementarán mecanismos de protección como el control de accesos por roles, la autenticación de usuarios, las copias de seguridad, el cifrado de información, las auditorías de acceso y las políticas de recuperación ante desastres, que aseguran la integridad y la confidencialidad de la información empresarial.

Recursos externos. Comprenden las API de redes sociales, el soporte del ERP, la consultoría especializada en BI y, de forma opcional, los servicios en la nube.

API de redes sociales. Se utilizarán API para capturar automáticamente el alcance de las campañas, las interacciones, los seguidores y las conversiones, lo que enriquecerá los análisis comerciales y de marketing.

Servicios de soporte del ERP. Se requiere soporte técnico especializado del proveedor de SAGA para la configuración del sistema, la integración con BI, el mantenimiento, las actualizaciones y la resolución de incidencias.

Consultoría especializada en BI. Durante la implementación se recomienda contar con consultores externos para el diseño del Data Warehouse, el desarrollo de los procesos ETL, la construcción de dashboards, la optimización del modelo de datos y la capacitación de los usuarios. Su experiencia reducirá los riesgos y acelerará la adopción de la solución.

Servicios en la nube (opcional). Como estrategia de escalabilidad, la empresa podrá considerar servicios en la nube para el respaldo de información, el almacenamiento histórico, la publicación de dashboards, el acceso remoto y la alta disponibilidad.

CAPÍTULO 4

Análisis de herramientas ETL y selección de la alternativa más adecuada

La implementación de una arquitectura de BI requiere incorporar una herramienta de integración de datos que automatice los procesos ETL y asegure que la información de los sistemas operacionales se consolide de forma confiable para el análisis estratégico. Su selección debe responder a criterios técnicos y al ecosistema tecnológico de la organización, en particular a la compatibilidad con los sistemas existentes, la capacidad de automatización, la escalabilidad y el soporte para procesos de calidad de datos.

Entre las soluciones disponibles destacan SQL Server Integration Services (SSIS), Talend Open Studio, Pentaho Data Integration (PDI) y Apache NiFi. SSIS es una plataforma empresarial que facilita la construcción de procesos ETL mediante flujos de trabajo automatizados y una integración eficiente con bases de datos relacionales. Talend y Pentaho son alternativas de código abierto con amplia conectividad hacia distintas fuentes y gran flexibilidad de diseño. Apache NiFi se orienta a la gestión de flujos de datos continuos y al procesamiento en tiempo real, por lo que resulta apropiado para arquitecturas con alto nivel de integración entre múltiples sistemas.

En la arquitectura propuesta para CVP, SAGA es la fuente principal y el proceso ETL alimenta un Data Warehouse corporativo del que se derivan los Data Marts analíticos. En este contexto, SSIS es la herramienta más compatible, por su capacidad para programar cargas automáticas, ejecutar transformaciones complejas, validar la integridad de la información y administrar cargas incrementales, funcionalidades acordes con la frecuencia de actualización y los controles de calidad definidos para el modelo de datos. Además, su integración con

Microsoft SQL Server favorece la administración centralizada de la información y reduce la complejidad operativa del proceso ETL.

Desde la perspectiva de la arquitectura empresarial, SSIS fortalece la gobernanza de datos porque reduce la dependencia de procesos manuales, mejora la trazabilidad de las transformaciones y facilita la disponibilidad de información consistente para los indicadores y los tableros de control. En consecuencia, es la alternativa que mejor se adapta al ecosistema tecnológico planteado para CVP y contribuye a consolidar una infraestructura de BI que respalda la toma de decisiones con información íntegra, oportuna y de alta calidad, sin modificar la estructura del Data Warehouse ni de los procesos ETL ya definidos.

Usuarios y perfiles de uso de cada dashboard

Las Tablas 19 a 25 describen, para cada dashboard, los usuarios, el perfil de acceso, la frecuencia de uso, los KPI visibles, el nivel de detalle y, según corresponda, las alertas, las funcionalidades y otras características.

Dashboard ejecutivo o gerencial

Tabla 19

Perfil de uso del dashboard ejecutivo o gerencial

Aspecto	Definición
Usuarios	Gerente general, director comercial, director de operaciones y presidente del tribunal (visor externo)
Perfil de acceso	Lectura (dashboard de solo visualización, sin capacidad de filtrado avanzado)
Frecuencia de uso	Diaria (inicio de jornada) y según necesidad estratégica

KPI visibles	Crecimiento porcentual de ventas, cumplimiento de presupuesto, margen bruto, margen operativo, ROA, ciclo de caja, productividad por colaborador y decisión basada en datos (%)
Nivel de detalle	Consolidado por área y sucursal (Cañar y El Tambo), sin detalle a nivel de transacción individual
Alertas	Desviaciones superiores al 10 % en el presupuesto, quiebres críticos de stock y caídas en las ventas

Nota. ROA = retorno sobre los activos. Elaboración propia.

Dashboard comercial o de ventas

Tabla 20

Perfil de uso del dashboard comercial o de ventas

Aspecto	Definición
Usuarios	Jefe comercial, coordinador de ventas, vendedores y asesores comerciales
Perfil de acceso	Lectura para vendedores (solo sus métricas); lectura y escritura para el jefe comercial (ajuste de metas, filtros y segmentaciones)
Frecuencia de uso	Varias veces al día (monitoreo del cumplimiento diario)
KPI visibles	Crecimiento de ventas, cumplimiento de presupuesto, ticket promedio, tasa de fidelización, clientes nuevos, ventas corporativas (%), tasa de venta cruzada, conversión de prospectos y nivel de satisfacción del cliente
Nivel de detalle	Por vendedor, sucursal, canal de venta, segmento de cliente (corporativo o retail) y día, semana o mes

Funcionalidades	Comparativos de desempeño entre vendedores, ranking de productos y proyección frente a la meta mensual

Nota. Elaboración propia.

Dashboard de marketing

Tabla 21

Perfil de uso del dashboard de marketing

Aspecto	Definición
Usuarios	Responsable de marketing, community manager, jefe comercial (visor) y gerente general (visor)
Perfil de acceso	Lectura y escritura para marketing (configuración de campañas y carga de métricas externas) para el jefe comercial; lectura para la gerencia
Frecuencia de uso	Semanal (revisión de campañas) y mensual (análisis de ROI)
KPI visibles	Alcance en redes sociales, tráfico digital, conversión de campañas, reconocimiento de marca, ROI de marketing, CAC, tasa de interacción (engagement rate) y tasa de crecimiento de seguidores
Nivel de detalle	Por campaña, canal (Facebook, Instagram y web), período y sucursal
Fuentes adicionales	Datos importados desde Excel (métricas de redes) y API de redes sociales

Nota. ROI = retorno de la inversión; CAC = costo de adquisición de clientes. Elaboración propia.

Dashboard de bodega e inventarios

Tabla 22

Perfil de uso del dashboard de bodega e inventarios

Aspecto	Definición
Usuarios	Jefe de bodega, responsable de inventarios, personal de bodega (recepción y despacho), jefe comercial (visor de disponibilidad) y compras
Perfil de acceso	Lectura y escritura para el jefe de bodega (registro de ajustes y controles); lectura para el personal operativo (consulta de ubicaciones y stock) y para compras (niveles de reposición)
Frecuencia de uso	Diaria (control de inventario) y semanal (análisis de rotación)
KPI visibles	Nivel de disponibilidad de inventario, rotación de inventarios, exactitud del inventario, tiempo de reposición, productos agotados, pérdidas por inventario, índice de cobertura y productos de baja rotación.
Nivel de detalle	Por SKU, categoría, ubicación en bodega (zona, pasillo y estante) y proveedor
Alertas	Stock mínimo, productos sin movimiento por más de 60 días y discrepancias entre el inventario físico y el sistema

Nota. SKU = unidad de mantenimiento de existencias (stock keeping unit). Elaboración propia.

Dashboard de operaciones y logística

Tabla 23

Perfil de uso del dashboard de operaciones y logística

Aspecto	Definición
Usuarios	Jefe de operaciones, supervisor de bodega, personal de picking y despacho, responsable de calidad y responsable de seguridad

Perfil de acceso	Lectura y escritura para el jefe de operaciones; lectura para supervisores y personal de línea
Frecuencia de uso	Diaria (seguimiento de despachos), semanal (eficiencia operativa) y mensual (seguridad y cumplimiento)
KPI visibles	Tiempo promedio de recepción, tiempo promedio de despacho, nivel de cumplimiento de pedidos, productos dañados, índice de accidentes, trazabilidad, eficiencia de picking, tasa de devoluciones por error, ocupación de espacio y cumplimiento de normas de seguridad
Nivel de detalle	Por pedido, empleado de bodega, turno y tipo de operación (recepción, picking o despacho)
Integración	Vinculado con el dashboard de bodega para la trazabilidad completa del producto

Nota. Elaboración propia.

Dashboard financiero y transversal

Tabla 24

Perfil de uso del dashboard financiero y transversal

Aspecto	Definición
Usuarios	Contador, tesorero, gerente general y jefe de talento humano
Perfil de acceso	Lectura y escritura para contador y tesorero (ajuste de proyecciones); lectura para la gerencia
Frecuencia de uso	Mensual (cierres) y trimestral (reportes a la dirección)

KPI visibles	Margen bruto, margen operativo, rotación de cartera, días de inventario (DIO), productividad por colaborador, tasa de retención de empleados, rotación de personal, costo laboral sobre ventas, ciclo de caja y ROA
Nivel de detalle	Consolidado mensual, por área de costos y por línea de producto
Restricciones	Acceso restringido por la sensibilidad de los datos salariales y financieros

Nota. DIO = days inventory outstanding; ROA = retorno sobre los activos. Elaboración propia.

Dashboard de administración del sistema de BI

Tabla 25

Perfil de uso del dashboard de administración del sistema de BI

Aspecto	Definición
Usuarios	Administrador del sistema de BI, asistente contable (carga de datos) y consultor externo de BI
Perfil de acceso	Administrador total (configuración de ETL, gestión de usuarios, auditoría y calidad de datos)
Frecuencia de uso	Diaria (monitoreo de cargas) y semanal (validación de calidad de datos)
KPI visibles	Tiempo de generación de reportes, usuarios activos del BI, precisión de pronósticos, tasa de adopción del sistema, disponibilidad del sistema, tiempo de respuesta de dashboards, calidad de datos, número de dashboards activos y satisfacción de usuarios
Funcionalidades	Monitoreo de procesos ETL, registros (logs) de acceso, gestión de roles y permisos y validación de integridad referencial

Nota. Elaboración propia.

Necesidades de información y acceso por departamento y perfil

Departamento de ventas

El Departamento de Ventas requiere información actualizada para analizar el comportamiento comercial y tomar decisiones orientadas a incrementar los ingresos. Es fundamental conocer los materiales de construcción de mayor demanda, identificar los productos de alta y baja rotación, evaluar el desempeño de los vendedores y comprender los patrones de compra de los clientes. Sus principales reportes son el Dashboard de Ventas, Ventas por Producto, Ventas por Categoría, Ventas por Cliente, Cumplimiento de Metas Comerciales y Tendencias de Venta, que permiten optimizar la estrategia comercial, mejorar la ubicación de los productos en la tienda y aumentar las oportunidades de venta.

Acceso del departamento. El departamento puede consultar los reportes comerciales, acceder al historial de ventas y de clientes y visualizar los indicadores de desempeño comercial. No tiene permisos para modificar información financiera ni inventarios maestros.

Departamento de marketing

El Departamento de Marketing necesita información que permita medir el impacto de las campañas publicitarias y las promociones y comprender el comportamiento de los clientes, lo que facilita la segmentación del mercado, el diseño de estrategias comerciales y la fidelización. Sus reportes más importantes son el Dashboard de Marketing, Análisis de Promociones, Segmentación de Clientes, Análisis de la Compra y del Medio por el que se Realiza, Captación de Clientes y Fidelización de Clientes, que identifican las campañas con mayor retorno, los productos más atractivos y las oportunidades para incrementar las ventas.

Acceso del departamento. El departamento puede consultar indicadores de ventas relacionados con campañas, información de clientes con fines analíticos y resultados de promociones y del comportamiento de compra. No tiene acceso a información financiera confidencial ni a costos internos.

Departamento de bodega

El Departamento de Bodega requiere información actualizada sobre el estado del inventario para garantizar la disponibilidad de los materiales de construcción y evitar pérdidas por desabastecimiento o sobreinventario, y necesita conocer la rotación de los productos para optimizar el almacenamiento. Sus reportes son el Dashboard de Inventario, Rotación de Inventario, Control de Stock, Productos Más Vendidos, Inventario Obsoleto, Kardex e Inventario Valorizado, que facilitan la planificación de compras, el control de existencias y la reposición oportuna.

Acceso del departamento. El departamento puede consultar y actualizar el inventario, acceder al Kardex y a los movimientos de bodega y visualizar los niveles de stock, las entradas y las salidas. No tiene acceso a información financiera estratégica ni presupuestaria.

Departamento financiero

El Departamento Financiero necesita información precisa y actualizada para controlar la situación económica de la empresa y apoyar las decisiones estratégicas, en particular sobre ingresos, gastos, flujo de efectivo, cuentas por cobrar, cuentas por pagar y rentabilidad. Sus reportes son el Dashboard Financiero, Flujo de Caja, Estado de Resultados, Cuentas por Cobrar, Cuentas por Pagar, Presupuesto frente a Ejecutado y Proyecciones Financieras, que

permiten evaluar la liquidez, controlar el cumplimiento del presupuesto y planificar las necesidades de financiamiento.

Acceso del departamento. El departamento tiene acceso completo a la información financiera, incluidos ingresos, costos, gastos, utilidades, presupuestos y proyecciones, y administra las cuentas por cobrar y por pagar. No puede modificar información comercial ni de inventario fuera de los procesos autorizados.

Departamento de Business Intelligence

El Departamento de Business Intelligence es una propuesta de implementación a corto plazo. Será responsable de integrar, analizar y transformar los datos de todas las áreas para generar información estratégica que facilite la toma de decisiones. Trabajará con reportes operacionales, tácticos y estratégicos: los operacionales monitorean las actividades diarias, como las ventas y los inventarios; los tácticos analizan el desempeño de cada departamento y detectan oportunidades de mejora; y los estratégicos apoyan a la alta dirección mediante indicadores corporativos, análisis de rentabilidad, pronósticos de ventas y evaluación del cumplimiento de los objetivos del negocio.

Acceso del departamento. El departamento tiene acceso completo a la información de todas las áreas, incluidos el histórico y los datos actualizados para análisis predictivos y estratégicos. Administra el Data Warehouse, los procesos ETL, los permisos y la seguridad de los reportes; crea, modifica y publica dashboards; y gestiona los KPI y los modelos analíticos.

Revisión y cumplimiento de los objetivos iniciales

La revisión del cumplimiento de los objetivos permite evaluar si las actividades desarrolladas responden a los propósitos planteados al inicio de la investigación. En este trabajo, centrado en el diseño de una solución de BI para CVP, la revisión determina el nivel de cumplimiento respecto de los objetivos establecidos para fortalecer la gestión de la información comercial y operativa de la empresa.

El proyecto surgió de la necesidad de mejorar la toma de decisiones mediante la integración de la información de las distintas áreas. Para ello, se analizaron los procesos actuales, se identificaron las limitaciones en el manejo de datos y se diseñó una arquitectura de BI compuesta por un Data Warehouse, Data Marts especializados, procesos ETL y KPI que proporcionan información confiable, integrada y disponible para la administración. A continuación, se analiza el grado de cumplimiento de cada objetivo según los resultados obtenidos.

Desarrollo

Revisión del objetivo general. El objetivo general consistió en diseñar una solución de BI que integrara la información comercial de CVP para fortalecer la toma de decisiones estratégicas y contribuir al incremento de las ventas a nivel nacional. Se cumplió mediante una propuesta integral que contempla la arquitectura del sistema de BI, el diseño del Data Warehouse, la definición de los Data Marts, la estructuración del proceso ETL y la identificación de los indicadores que medirán el desempeño organizacional.

La propuesta responde a las necesidades detectadas en el diagnóstico institucional, el cual evidenció que la empresa depende de SAGA y del uso complementario de Excel para

generar reportes, lo que dificulta la integración de la información y limita la obtención de análisis oportunos para la alta dirección.

Cumplimiento del diagnóstico organizacional. Uno de los primeros objetivos específicos fue analizar la situación actual de la gestión de la información. Se cumplió mediante la identificación de las fuentes internas y externas de datos, la descripción del funcionamiento de SAGA y el análisis del flujo de información entre las áreas comercial, de operaciones, de compras y de contabilidad. También se evaluó cómo se generan los reportes y se determinaron las debilidades en la disponibilidad y la calidad de la información.

El diagnóstico identificó la dependencia del ingreso manual de datos, el uso de múltiples archivos de Excel para el análisis, la ausencia de tableros de control, la limitada capacidad de análisis histórico y predictivo, y la dificultad para integrar información estratégica de redes sociales y campañas de marketing. Estos hallazgos justificaron una solución de BI adaptada a las necesidades de la empresa.

Cumplimiento del diseño de la solución de Business Intelligence. Otro objetivo específico fue diseñar una solución tecnológica que consolidara la información empresarial para el análisis estratégico. Se alcanzó mediante una arquitectura de datos basada en un Data Warehouse corporativo alimentado por procesos ETL desde SAGA, archivos de Excel y otras fuentes complementarias. Sobre esta infraestructura se diseñaron Data Marts especializados para las áreas comercial y de bodega, que organizan la información según las necesidades de cada proceso.

Además, se estableció una estructura de dimensiones y tablas de hechos que permitirá realizar consultas analíticas eficientes, conservar el histórico y elaborar reportes dinámicos y dashboards gerenciales orientados a la toma de decisiones.

Cumplimiento de la definición de indicadores de desempeño. Un componente esencial fue la definición de KPI para medir objetivamente el cumplimiento de los objetivos estratégicos. El proyecto estructuró 55 indicadores distribuidos entre las áreas de ventas, marketing, bodega comercial, bodega de operaciones y finanzas, además de los indicadores específicos para evaluar el desempeño de la solución de BI. Estos indicadores miden el crecimiento de las ventas, la fidelización de clientes, la rotación de inventarios, la rentabilidad, el cumplimiento de pedidos, la productividad del personal, el desempeño financiero y la calidad de la información, y ofrecen una visión integral del comportamiento organizacional.

Cumplimiento de la integración de la información para la toma de decisiones. El proyecto también buscó integrar la información de las distintas áreas para facilitar la toma de decisiones estratégicas. Se diseñó un modelo que centraliza los datos de ventas, compras, inventarios, marketing, operaciones y contabilidad, elimina la fragmentación existente y permite a la gerencia disponer de información consolidada, consistente y actualizada para evaluar el desempeño institucional.

Esta integración mejorará la coordinación entre departamentos, optimizará la planificación de compras, reducirá los sobreinventarios, incrementará la disponibilidad de productos y fortalecerá el control de los procesos comerciales y operativos.

Cumplimiento de la propuesta de análisis estratégico. Finalmente, el proyecto proporciona herramientas que apoyan la planificación estratégica y el crecimiento empresarial. La implementación propuesta permitirá analizar los históricos de ventas, identificar tendencias de consumo, evaluar el comportamiento de los clientes, analizar la rentabilidad por producto, medir el impacto de las campañas de marketing y generar proyecciones que apoyen la expansión comercial a nivel nacional. Así, la organización contará con información de mayor valor para reducir la incertidumbre y fundamentar sus decisiones en evidencia objetiva.

CAPÍTULO 5

Business case y creación del departamento

El Departamento Comercial de CVP Centro Constructor Vásquez Pacheco constituye una de las áreas fundamentales para el funcionamiento y el crecimiento de la organización, debido a que concentra las actividades relacionadas con la generación de ventas, la captación y fidelización de clientes, el posicionamiento de la marca y la gestión de las relaciones comerciales. El presente capítulo caracteriza dicho departamento, identifica las limitaciones que condicionan su gestión y las carencias de formación y recursos que deben superarse, define el organigrama y los recursos externos necesarios para incorporar un Departamento de Business Intelligence (BI) y expone el business case de costos y beneficios de la iniciativa.

Caracterización del departamento comercial

Dentro del área se identifican principalmente las subáreas de Ventas y Marketing. La subárea de Ventas tiene como propósito maximizar los ingresos mediante una gestión comercial eficiente, satisfacer las necesidades de los clientes e incrementar las ventas y las relaciones comerciales de largo plazo. Entre sus objetivos se encuentran incrementar las ventas mensuales, mejorar la experiencia del cliente, captar nuevos clientes, fortalecer la fidelización, desarrollar indicadores de desempeño, incrementar las ventas corporativas y promover la venta cruzada entre las diferentes categorías de productos.

Por su parte, Marketing tiene como función contribuir al posicionamiento de CVP mediante estrategias tradicionales y digitales. Esta subárea adquiere especial importancia

debido a que la empresa busca ampliar su presencia y fortalecer su reconocimiento más allá de sus mercados actuales, por lo que necesita relacionar las acciones de marketing con resultados comerciales concretos, como la captación de clientes, las conversiones, las ventas y el retorno de la inversión.

Situación actual de la gestión comercial

Actualmente, la gestión comercial se apoya principalmente en el sistema SAGA ERP, que concentra la información generada por las operaciones de la empresa. En el área Comercial se registran las ventas, las facturas electrónicas, los clientes, los productos, las cantidades comercializadas y los valores de cada transacción. Además, las operaciones comerciales actualizan la información de inventario, lo que permite mantener las existencias relacionadas con los movimientos de venta.

Esta centralización constituye una fortaleza importante, ya que evita que la información operativa se encuentre completamente dispersa entre diferentes sistemas. SAGA integra también información de compras, inventarios, contabilidad, cuentas por cobrar y cuentas por pagar, de manera que las distintas áreas trabajan sobre una fuente institucional de información.

Sin embargo, existe una diferencia importante entre disponer de información y utilizarla estratégicamente. Aunque SAGA permite registrar y consultar las operaciones, no está orientado al análisis avanzado del comportamiento de las ventas, de las tendencias de los clientes ni de las proyecciones comerciales. Por esta razón, cuando se requieren análisis más específicos, los usuarios deben exportar la información a Excel para realizar procesos adicionales de filtrado, consolidación, tablas dinámicas y elaboración de gráficos.

Principales problemas y limitaciones

El diagnóstico de la situación actual permite identificar cinco limitaciones principales en la gestión comercial de CVP.

Dependencia de procesos manuales

Una de las principales limitaciones es la dependencia del ingreso manual de información. Cada área registra sus operaciones en SAGA y cualquier error humano puede afectar la calidad de los datos y, posteriormente, de los análisis realizados con ellos. En consecuencia, la calidad del análisis comercial depende directamente de la calidad del registro inicial: las inconsistencias en los nombres de productos, clientes o códigos dificultan la segmentación y la comparación de la información.

Dependencia de Excel

Aunque Excel es una herramienta útil para los análisis complementarios, su utilización como mecanismo principal para consolidar determinados reportes genera riesgos, entre ellos la duplicidad de archivos, la coexistencia de distintas versiones de la información y un mayor tiempo de procesamiento. Esta situación puede provocar que diferentes usuarios trabajen con versiones distintas de un mismo reporte, lo que dificulta que la organización tenga una única visión de sus resultados comerciales.

Falta de integración de información externa

SAGA concentra principalmente la información interna de la empresa, pero no integra por completo fuentes estratégicas como las redes sociales, las campañas de marketing, el comportamiento del mercado o los indicadores comerciales externos. Esta situación limita la

capacidad del Departamento Comercial para relacionar, por ejemplo, una campaña de marketing con los clientes captados, las ventas generadas y el retorno de la inversión.

Ausencia de dashboards en tiempo real

En la actualidad no se dispone de tableros de control que permitan visualizar de manera integral indicadores como las ventas por sucursal, el margen, la rotación de inventarios, el desempeño de los vendedores o los productos más vendidos. Esto obliga a consultar diferentes reportes para obtener una visión completa del desempeño comercial.

Limitada capacidad predictiva

Otra debilidad es la limitada capacidad para desarrollar análisis históricos y predictivos. Aunque la empresa almacena información en SAGA, actualmente no la aprovecha plenamente mediante herramientas de Business Intelligence para identificar patrones de compra, estacionalidad o posibles comportamientos futuros de la demanda.

Definición de las carencias en formación y recursos

A partir del diagnóstico realizado en CVP, se identificaron varias carencias relacionadas con la formación del personal y con la disponibilidad de recursos tecnológicos. Estas limitaciones podrían dificultar la implementación y el funcionamiento adecuado de un área de Business Intelligence, debido a que la empresa todavía requiere fortalecer las capacidades necesarias para recopilar, analizar e interpretar la información de manera estratégica.

En el ámbito de la formación, es necesario capacitar al personal en análisis e interpretación de datos, elaboración y seguimiento de indicadores de gestión, manejo de bases de datos y utilización de herramientas de inteligencia de negocios, especialmente Power BI. El desarrollo de estas competencias permitirá que los colaboradores no se limiten al registro de

información, sino que puedan comprenderla, identificar tendencias y utilizarla como respaldo para la toma de decisiones en las diferentes áreas de la organización.

En cuanto a los recursos, la empresa requiere una infraestructura tecnológica que facilite el almacenamiento, la integración y el procesamiento de los datos provenientes del ERP SAGA. Esto comprende licencias de software, equipos con capacidad suficiente, herramientas para automatizar los procesos de extracción, transformación y carga de datos (ETL) y un entorno destinado al desarrollo de un Data Warehouse. Asimismo, se recomienda asignar a una persona responsable de Business Intelligence, encargada de supervisar la calidad de los datos, administrar las herramientas implementadas y garantizar que la información se encuentre disponible, actualizada y sea confiable para todas las áreas.

Coordinación de los flujos entre otras áreas

El área de Business Intelligence deberá funcionar de manera transversal y mantener una coordinación permanente con Ventas, Compras, Inventarios, Contabilidad y Gerencia. Cada una de estas áreas genera información relevante para el funcionamiento de CVP, por lo que será necesario establecer procedimientos claros para el registro, la validación, la actualización y el intercambio de datos. Esta coordinación permitirá consolidar la información proveniente del ERP SAGA en un repositorio centralizado y reducir inconsistencias, duplicidades y retrasos en la generación de reportes.

El flujo de información deberá iniciar con el registro oportuno de las operaciones en cada departamento. Posteriormente, los datos serán revisados, integrados y procesados por el área de Business Intelligence para generar indicadores, informes y tableros de control, cuyos resultados se compartirán con los responsables de cada área y con la gerencia. Con ello se busca

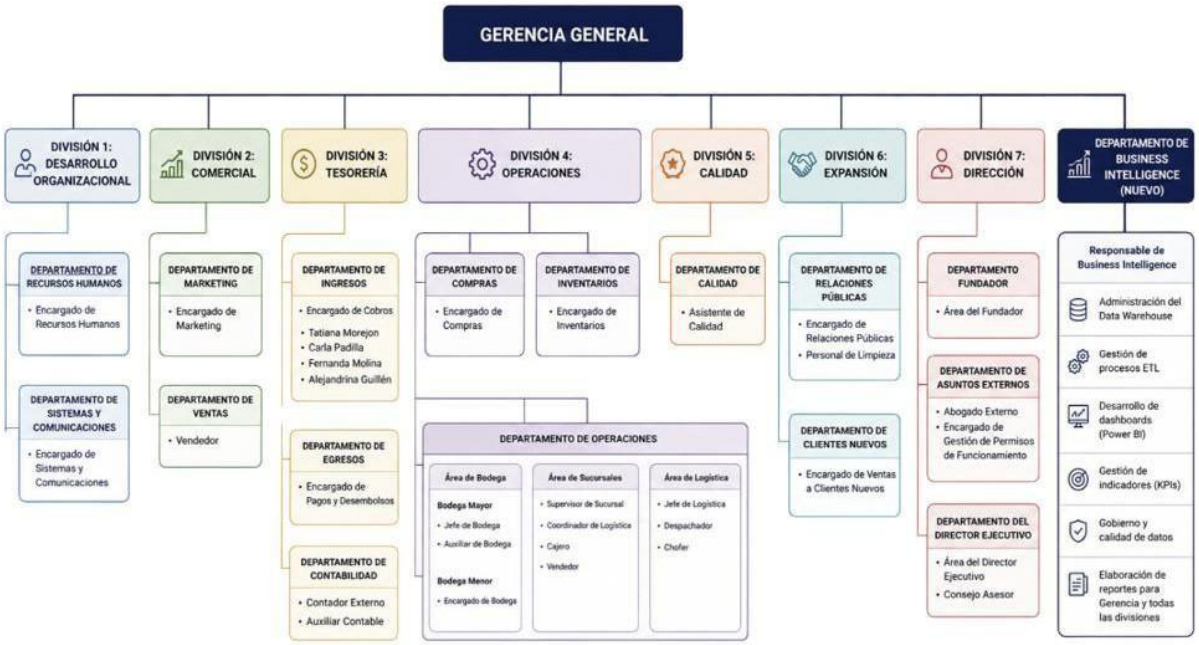
evaluar el desempeño de la empresa, detectar variaciones en las ventas, controlar los niveles de inventario, analizar las compras y apoyar la planificación financiera y comercial.

Una adecuada coordinación permitirá, además, definir responsabilidades sobre la calidad de la información y establecer canales de retroalimentación para corregir errores o registros incompletos. De esta manera, Business Intelligence no funcionará como un departamento aislado, sino como un apoyo estratégico que transforme los datos generados por toda la empresa en información útil para identificar tendencias, anticipar riesgos, reconocer oportunidades de mejora y fortalecer la toma de decisiones.

Definición del organigrama, dependencias de otras áreas y solución de problemas

Con base en la estructura organizacional actual de CVP, se propone incorporar el Departamento de Business Intelligence como una unidad de apoyo estratégico con dependencia directa de la Gerencia General, según se muestra en la Figura 4. Esta ubicación responde a la naturaleza transversal del área, cuya función será consolidar y analizar la información generada por todas las divisiones de la empresa sin sustituir las responsabilidades de cada una. De esta manera, el Departamento de BI trabajará de forma coordinada con las divisiones de Desarrollo Organizacional, Comercial, Tesorería, Operaciones, Calidad, Expansión y Dirección, y obtendrá la información necesaria para generar indicadores, informes ejecutivos y tableros de control que respalden la toma de decisiones. Según Laudon y Laudon (2022), la integración de los sistemas de información con la estructura organizacional permite mejorar la coordinación entre las áreas y fortalecer la gestión estratégica de la información.

Figura 4. Organigrama propuesto con la incorporación del Departamento de Business Intelligence



Fuente: Elaboración propia, adaptado del organigrama institucional de CVP Centro Constructor Vásquez Pacheco (2026).

La dependencia directa de la Gerencia General permitirá que el Departamento de Business Intelligence supervise la calidad de los datos, identifique inconsistencias en la información y proponga acciones de mejora para optimizar los procesos internos. Asimismo, facilitará la estandarización de los indicadores de desempeño y reducirá la duplicidad de información entre departamentos, lo que garantiza que todas las áreas trabajen con datos confiables y actualizados. Como señalan Turban et al. (2021), una adecuada gobernanza de datos y una estructura organizacional claramente definida son factores clave para que las

iniciativas de Business Intelligence generen valor y apoyen la toma de decisiones basada en evidencia.

Recursos externos destinados: subcontratación

La puesta en marcha del ecosistema de BI requiere experiencia especializada que la empresa no posee internamente. Por esta razón se contempla subcontratar los servicios que se resumen en la Tabla 1 y que se describen a continuación.

Tabla 26. Recursos externos y servicios subcontratados

Categoría	Servicio externo	Alcance
Consultoría especializada	Especialistas en Business Intelligence	Diseño del Data Warehouse y modelado dimensional; desarrollo e implementación de los procesos ETL; construcción de dashboards e indicadores (KPI); optimización del modelo de datos; capacitación a usuarios finales.
Soporte técnico del ERP	Proveedor de SAGA	Configuración del sistema para la integración con BI; soporte técnico especializado y mantenimiento; actualizaciones; resolución de incidencias.
Fuentes externas de marketing	APIs de redes sociales	Captura automática de alcance de campañas, interacciones, seguidores, conversiones y tráfico digital de plataformas como Facebook e Instagram.
Servicios en la nube (opcional)	Proveedores cloud	Respaldo de información histórica; almacenamiento del Data Warehouse; publicación de dashboards con acceso remoto; alta disponibilidad del sistema.

Fuente: Elaboración propia (2026).

La consultoría especializada se recomienda para reducir riesgos y acelerar la adopción de la solución. El soporte técnico del proveedor del ERP resulta necesario porque SAGA es el sistema fuente principal. Las APIs de redes sociales permiten enriquecer el Data Mart Comercial con métricas digitales, y los servicios en la nube se proponen como estrategia de escalabilidad y respaldo.

A estas líneas se suman dos recursos que no se plantean de forma explícita como subcontratación, pero que son necesarios: una plataforma de visualización de BI (se sugiere Power BI o una equivalente), lo que implica licenciamiento y posible soporte del proveedor; y una infraestructura de servidores, que podría incluir servicios de hosting o de centros de datos externos si la empresa no dispone de servidores propios adecuados.

Business case de costos y beneficios

Costos

La estimación de costos considera la remuneración del equipo de BI y la inversión de implementación. La Tabla 2 detalla los cargos, sus funciones y responsabilidades y los sueldos correspondientes, expresados en dólares estadounidenses.

Tabla 27. Costos del equipo de Business Intelligence y de implementación (USD)

Cargo	Funciones	Responsabilidades	Sueldo mensual	Sueldo anual

Analista de Business Intelligence	Transformar datos en información comprensible y útil que impulse la toma de decisiones.	Analizar datos; desarrollar paneles de control; generar reportes; traducir los requerimientos del negocio.	1,875.00	22,500.00
Arquitecto de datos	Crear un sistema de datos eficiente y ajustado a los objetivos de largo plazo del negocio.	Diseño del Data Warehouse; integridad y accesibilidad de los datos; alineación de los datos con los objetivos del negocio.	1,500.00	18,000.00
Desarrollador de BI	Desarrollar e implementar las herramientas y sistemas necesarios para el análisis y la generación de reportes.	Desarrollar herramientas de BI; crear los dashboards y reportes necesarios; mantener y actualizar las herramientas.	1,000.00	12,000.00
Costos de implementación				8,000.00
Costos totales				**60,500.00**

Fuente: Elaboración propia, adaptado del organigrama institucional de CVP Centro Constructor Vásquez Pacheco (2026).

El total de USD 60,500 resulta de sumar los sueldos anuales del equipo (USD 52,500) y los costos de implementación. A este monto se añade el licenciamiento de la plataforma de visualización: la licencia de Power BI Pro tiene un costo de mercado de USD 10 por usuario al mes. El costo de implementación para CVP asciende a USD 8,000 como inversión única.

Retorno de la inversión

El retorno de la inversión (ROI) de la implementación de BI en CVP se materializa principalmente en tres dimensiones: la reducción del tiempo destinado a la elaboración manual de reportes, la mejora en la toma de decisiones mediante información precisa y actualizada, y la detección temprana de problemas operativos y comerciales.

La automatización de los reportes relacionados con ventas, inventario, compras, cartera de clientes y desempeño comercial permitirá disminuir las tareas repetitivas, reducir los errores humanos y mejorar la disponibilidad de información para los responsables de cada área. Los equipos podrán así enfocarse en actividades de mayor valor estratégico, como la planificación comercial, la optimización de recursos y la mejora de la atención al cliente.

Asimismo, el análisis de datos históricos y en tiempo real permitirá identificar oportunamente situaciones como la disminución de las ventas, los productos con baja o alta rotación, el exceso o la falta de inventario, los retrasos en los pagos de los clientes y las desviaciones en los objetivos comerciales. La detección anticipada de estos factores facilitará la implementación de acciones correctivas, con ahorros operativos, una mejor gestión del capital de trabajo y una mayor rentabilidad para CVP (OKUNDATA, 2026).

CAPÍTULO 6

Selección e identificación de fuentes de datos externas e internas, estructuradas o semiestructuradas

Datos estructurados

Los datos estructurados corresponden a la información que ha sido transformada en un modelo de datos bien definido. Pueden ser generados por personas o por máquinas: en el primer caso, el ejemplo más habitual son las hojas de cálculo; en el segundo, los datos de puntos de venta, como las cantidades o los códigos de barras. Por la forma en que se organizan, son los datos más sencillos de analizar, por delante de los semiestructurados y los no estructurados (Bridge, 2026).

En el análisis se utilizarán los datos estructurados internos de la compañía que se obtienen de los reportes contables del sistema, los cuales se resumen en la Tabla 3.

Tabla 28. Datos estructurados internos considerados por reporte

Reporte	Campos considerados
Ventas	Fecha; número de factura; nombre del cliente; producto; cantidad vendida; precio unitario; descuento; subtotal; IVA; total.
Cuentas por cobrar	Cliente; número de factura; fecha de factura; fecha de vencimiento; valor original; abonos realizados; saldo pendiente; días de crédito; días de mora; estado (vigente o vencida); vendedor.
Inventario	Código del producto/SKU; descripción; unidad de medida; stock actual, mínimo y máximo; costo unitario; precio de venta; valor total del inventario; entradas;

	salidas; devoluciones; bodega; ubicación; fecha de última compra; fecha de última venta.
Cuentas por pagar	Código y nombre del proveedor; número de factura; fecha de factura; fecha de vencimiento; monto original; abonos; saldo pendiente; días para vencimiento; días de mora; estado; forma de pago.
Compras	Fecha de compra; número de factura; proveedor; código del producto; producto; cantidad comprada; costo unitario; descuento; IVA; total; forma de pago; condiciones de crédito.

Fuente: Elaboración propia (2026).

Datos no estructurados

Los datos no estructurados comprenden prácticamente cualquier información que no se encuentre en un formato específico, como el texto de un documento, una página web o los archivos de registro que no son fáciles de separar. También lo son los comentarios y publicaciones en redes sociales, que permiten conocer las futuras tendencias de compra o la efectividad de una campaña de marketing (Bridge, 2026). En el proyecto se trabajará con los siguientes datos no estructurados (IBM, 2026):

- Comentarios de los clientes y su nivel de satisfacción, incluidos los comentarios y reseñas sobre su experiencia en redes sociales y Google Business.
- Registros de quejas recibidas por WhatsApp o por llamadas telefónicas.
- Devoluciones por productos defectuosos.

Definición de la relación entre los datos

A partir de la identificación de las fuentes de datos de CVP, es necesario establecer las relaciones existentes entre los diferentes conjuntos de información antes de definir la arquitectura final de la solución de BI. La empresa dispone de información proveniente principalmente del sistema SAGA, donde se registran operaciones de ventas, compras, inventarios, clientes y procesos contables. Adicionalmente, se identifican datos provenientes de clientes y proveedores, así como información no estructurada relacionada con comentarios, quejas, devoluciones y experiencias de los clientes. Esta distinción resulta relevante porque, como señala IBM (2026), los datos estructurados se organizan en esquemas predefinidos que facilitan su análisis inmediato, mientras que los no estructurados requieren procesos previos de clasificación y transformación antes de integrarse a un modelo analítico.

Desde la perspectiva del modelado de datos, se recomienda que la propuesta de BI adopte un modelo dimensional basado en tablas de hechos y dimensiones. En el caso de CVP, las principales relaciones pueden estructurarse alrededor de dimensiones comunes como tiempo, producto, cliente, proveedor, vendedor y sucursal, que permitirán conectar los diferentes procesos empresariales.

Una relación fundamental corresponde a ventas e inventarios. El inventario contiene información sobre el código del producto, la descripción, las existencias, el stock mínimo y máximo, el costo unitario, el precio de venta, las entradas, las salidas, las devoluciones, la bodega y las fechas de última compra y venta. Al utilizar el código del producto como elemento común, será posible contrastar la demanda comercial con la disponibilidad física de los productos y desarrollar indicadores como rotación de inventario, cobertura de stock, productos

agotados y productos con baja rotación. La relación resulta especialmente relevante para CVP debido al volumen de productos comercializados y a la necesidad de garantizar su disponibilidad para atender la demanda.

La relación entre compras y cuentas por pagar permitirá complementar el análisis operativo con una perspectiva financiera. Las cuentas por pagar contienen información sobre el proveedor, la factura, las fechas de emisión y vencimiento, el monto original, los abonos, el saldo pendiente, los días para el vencimiento, los días de mora y el estado de la obligación. Al relacionarla con las compras efectuadas, la organización podrá determinar no solamente cuánto compra, sino también cuánto de esas adquisiciones permanece pendiente de pago y cuáles son los compromisos financieros próximos a vencer, lo que contribuirá al control de la liquidez y a una mejor planificación de los pagos.

Una relación adicional de importancia corresponde a ventas y cuentas por cobrar. Debido a que las cuentas por cobrar contienen el número de factura, el cliente, la fecha de factura, el vencimiento, el valor original, los abonos, el saldo pendiente y los días de mora, es posible utilizar la factura y el cliente como elementos de correspondencia entre ambas estructuras. Esta integración permitiría generar indicadores como cartera vencida, nivel de recuperación, días promedio de cobro y ventas a crédito por cliente. En consecuencia, la Gerencia podría diferenciar entre el volumen de ventas generado y el efectivo efectivamente recuperado, con una visión más completa del desempeño comercial y financiero. La Tabla 4 sintetiza las relaciones propuestas.

Tabla 29. Principales relaciones de datos propuestas para CVP

Relación	Elementos principales	Información que permite obtener
Ventas – Productos	Código de producto/SKU	Productos más vendidos, ingresos y rotación.
Ventas – Clientes	Cliente / código de cliente	Frecuencia de compra, valor del cliente y comportamiento.
Ventas – Vendedores	Vendedor	Cumplimiento de metas y desempeño comercial.
Ventas – Inventario	Producto + tiempo	Demanda, disponibilidad y riesgo de desabastecimiento.
Ventas – Cuentas por cobrar	Número de factura + cliente	Ventas a crédito, recuperación y cartera vencida.
Compras – Productos	Código de producto	Cantidades adquiridas y costos de adquisición.
Compras – Proveedores	Código de proveedor	Dependencia y comportamiento de proveedores.
Compras – Inventario	Producto + tiempo	Abastecimiento, entradas y reposición.
Compras – Cuentas por pagar	Factura + proveedor	Obligaciones pendientes y vencimientos.
Inventario – Bodega/Sucursal	Producto + ubicación	Existencias por ubicación y disponibilidad.
Todas las áreas – Tiempo	Fecha	Tendencias, comparaciones y estacionalidad.

Fuente: Elaboración propia (2026).

Casos de uso y modelo de consumo de bi

Matriz de priorización de KPI por nivel de impacto

Se definió un total de 55 KPI, organizados bajo una estrategia de despliegue por fases según su criticidad operativa y estratégica. La Tabla 5 presenta la matriz de priorización.

Tabla 30. Matriz de priorización de KPI por nivel de impacto

Nivel de prioridad	Áreas y métricas clave incluidas	Frecuencia de monitoreo	Fase de despliegue
Crítico	Crecimiento de ventas, cumplimiento de presupuesto, margen bruto, exactitud de inventario y cumplimiento de pedidos.	Diaria / semanal	Fase 1: implementación inmediata en el dashboard principal.
Importante	Ticket promedio, tasa de fidelización, rotación de inventarios, tiempo de despacho, ROI de marketing y productividad por colaborador.	Semanal / mensual	Fase 2: dashboards secundarios especializados.
Seguimiento	Alcance en redes, reconocimiento de marca, índice de accidentes, tasa de retención de personal y adopción del BI.	Mensual / trimestral	Fase 3: tableros transversales y de control de gestión.

Fuente: Elaboración propia (2026).

Especificación funcional de dashboards por perfil de usuario

El modelo de consumo contempla dashboards diferenciados según el perfil del usuario, de modo que cada área acceda a la información pertinente a su función. La Tabla 6 especifica el usuario objetivo, el caso de uso y las métricas clave de cada uno.

Tabla 31. Especificación funcional de los dashboards por perfil de usuario

Dashboard	Usuario objetivo	Caso de uso	Métricas clave
Ejecutivo / gerencial	Gerencia General y directores	Monitoreo transversal de la salud financiera y operativa de la empresa a alto nivel, sin detalle transaccional.	Crecimiento de ventas, márgenes brutos, ROA, ciclo de caja y productividad general.
Comercial y de ventas	Jefes comerciales y vendedores	Seguimiento operativo del rendimiento de ventas con filtros multidimensionales (vendedor, sucursal, canal y cliente).	Cumplimiento de metas de venta, ticket promedio (≥ USD 280) y desempeño de venta cruzada.
Bodega e inventarios	Jefes de bodega y Departamento de Compras	Gestión eficiente del stock, prevención de desabastecimiento y control de rotación de SKU en bodegas.	Rotación de inventario, disponibilidad (≥ 95%), productos agotados (≤ 5%) y alertas de stock mínimo.
Operaciones y logística	Supervisores de logística y despacho	Control de tiempos de atención, eficiencia en la preparación de pedidos y gestión de riesgos operativos.	Tiempo de recepción (≤ 2 h), tiempo de despacho (≤ 4 h), eficiencia de picking (≥ 99.5%) e índice de seguridad laboral.

Marketing digital y captación	Equipo de marketing y estrategia comercial	Evaluación del retorno de la inversión de campañas digitales e integración de métricas de redes sociales.	ROI de campañas (≥ 200%), costo de adquisición de clientes (CAC ≤ USD 15) y tasa de conversión digital.
Financiero y transversal	Contabilidad y tesorería	Control de la liquidez, seguimiento de la cartera de clientes y eficiencia del costo laboral.	Rotación de cartera, días de inventario (DIO ≤ 60 días) y relación costo laboral / ventas.
Administración del sistema BI	Administrador de TI / responsable de BI	Control de la calidad de los datos, la estabilidad técnica y el rendimiento de las plataformas de BI.	Completitud (≥ 99%), exactitud (≥ 98%), disponibilidad del sistema (≥ 99%) y tiempo de respuesta (≤ 3 s).

Fuente: Elaboración propia (2026).

Estructura organizacional y modelo operativo de BI

En cuanto a la ubicación organizacional, se crea el Departamento de Business Intelligence como unidad de apoyo estratégico con dependencia directa de la Gerencia General (Figura 1). Su operación es transversal y atiende a las áreas Comercial, Operaciones, Tesorería, Calidad, Expansión y Desarrollo Organizacional.

La estrategia de subcontratación y soporte técnico, ya descrita en la Tabla 1, se concreta en tres frentes: la consultoría externa, a cargo del diseño dimensional del Data Warehouse, el desarrollo de los procesos ETL en SQL Server Integration Services (SSIS), la maquetación de los tableros en Power BI y la capacitación del personal; el soporte del proveedor del ERP, para

la integración con la base de datos transaccional de SAGA; y las integraciones externas, mediante el consumo de APIs para la extracción automatizada de métricas de marketing digital.

Big data: impacto en el proyecto, áreas, departamentos y procesos de negocio

En el contexto de CVP, Big Data representa la capacidad de procesar y analizar de manera masiva datos estructurados, como las transacciones de ventas en SAGA, la facturación electrónica, el histórico de compras y los inventarios, y datos no estructurados o semiestructurados, como las métricas de redes sociales, el tráfico web, el comportamiento de navegación de los clientes, los registros (logs) del sistema y las encuestas de satisfacción.

Las V del Big Data aplicadas a CVP

Volumen. Manejo consolidado del historial transaccional de 9,000 clientes, de un catálogo con miles de SKU de construcción y ferretería y de los registros operativos de múltiples sucursales.

Velocidad. Procesamiento e integración diaria de la carga de datos (ETL), y en tiempo real en las operaciones, para alimentar los tableros de control ejecutivos y operativos a primera hora del día (8:00 a. m.).

Variedad. Unificación de los datos internos del ERP SAGA, de los archivos estructurados de Excel (recursos humanos y campañas) y de las métricas de entornos digitales y de marketing obtenidas mediante APIs.

Veracidad. Eliminación de errores de digitación y de duplicidades de registros de clientes (cédula frente a códigos antiguos) y estandarización de la nomenclatura de los

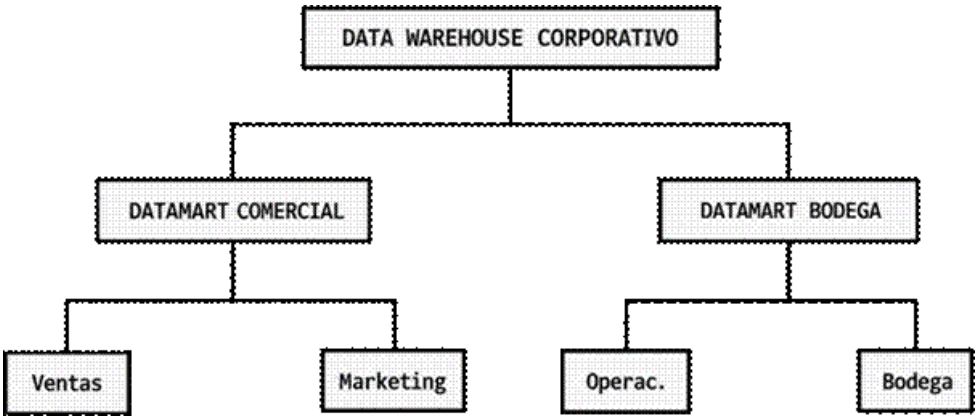
productos, mediante políticas estrictas de gobernanza y calidad de datos, con una meta de exactitud igual o superior al 98%.

Valor. Transformación de los datos crudos en conocimiento accionable para predecir la demanda, optimizar el capital de trabajo e incrementar el ticket promedio y las ventas corporativas.

Impacto por áreas y departamentos de negocio

El modelo analítico propuesto contempla la creación de un Data Warehouse corporativo central y de Data Marts especializados, que se muestran en la Figura 2 y que impactarán de forma directa en las áreas organizacionales que sintetiza la Tabla 7.

Figura 2. Arquitectura propuesta del Data Warehouse corporativo y de los Data Marts



Fuente: Elaboración propia (2026).

Tabla 32. Impacto esperado del modelo analítico por área de negocio

Área	Impacto transformativo	Data Mart asociado	Aporte clave
Comercial (Ventas y Marketing)	Pasa de una gestión reactiva de ventas a una estrategia analítica basada en segmentación de clientes, venta cruzada y analítica de conversión.	DM Comercial	Ventas: monitoreo continuo del cumplimiento de presupuestos, efectividad por asesor y captación de clientes corporativos (constructoras y contratistas), con la meta de representar al menos el 25% de las ventas totales. Marketing: medición del ROI de campañas (superior al 200%), costo de adquisición de clientes inferior a USD 15 y tasa de conversión de prospectos en canales digitales y redes sociales.
Operaciones y logística (Bodega)	Optimización de la cadena de suministro, trazabilidad total del inventario y erradicación de cuellos de botella en la recepción y el despacho.	DM Bodega	Tiempos promedio de recepción inferiores a 2 horas y de despacho inferiores a 4 horas; control de la ocupación de bodega (meta de 75% a 85%); eficiencia de picking de 99.5%; erradicación de pérdidas por deterioro o caducidad de materiales de construcción.

Financiera, Tesorería y Contabilidad	Visibilidad inmediata de la salud financiera, del margen de rentabilidad por línea de negocio y de la gestión del flujo de caja.	–	Ciclo de caja inferior a 45 días, rotación de cartera de clientes y días de inventario inferiores a 60; evaluación de la productividad por colaborador (referencia: USD 159,580) y del margen operativo (referencia: 8%).
Dirección General y Departamento de BI (propuesto)	Eliminación de la «ceguera estratégica» causada por reportes fragmentados en Excel.	–	Dashboard ejecutivo en tiempo real con alertas automáticas ante desviaciones superiores al 10% en presupuestos o quiebres críticos de stock; Departamento de BI responsable de la gobernanza de datos, la administración del ETL (SSIS) y el desarrollo de analítica predictiva.

Fuente: Elaboración propia (2026). El guion (–) indica que no se asocia un Data Mart específico en el diseño propuesto.

En conjunto, el análisis de las fuentes de datos, de las relaciones entre ellas y de los casos de uso muestra que CVP dispone de información suficiente y diversa para sustentar una solución de BI con enfoque de Big Data. La priorización de los KPI y la especificación de los dashboards permiten ordenar su implementación por fases, y los impactos esperados en las

áreas comercial, operativa y financiera constituyen la base para los capítulos posteriores del proyecto.

CAPÍTULO 7

Almacenamiento big data

Análisis de las alternativas de diferentes proveedores propuestos

- **Amazon Web Services**

Amazon Web Services (AWS) es el proveedor cloud con mayor cuota de mercado a nivel global, consolidándose como la opción líder en cuanto a infraestructura, servicios y alcance.

Su arquitectura altamente escalable y flexible permite abordar desde proyectos de startups hasta cargas empresariales complejas y críticas.

- **Microsoft Azure**

Azure se distingue por su profunda integración con el ecosistema Microsoft, incluyendo herramientas ampliamente adoptadas como Windows Server, Active Directory y Office 365. Esta compatibilidad nativa lo convierte en una opción especialmente atractiva para empresas que ya operan en entornos Microsoft.

- **Google Cloud Platform**

Su infraestructura global de alto rendimiento, combinada con una política de precios transparente y un fuerte compromiso con el open source, lo convierten en una plataforma especialmente atractiva para proyectos innovadores y entornos tecnológicos modernos que demandan flexibilidad, escalabilidad y potencia de cómputo.

- **IBM Cloud**

Su enfoque está claramente orientado al entorno empresarial, ofreciendo soluciones robustas con una fuerte especialización en entornos híbridos y multicloud. IBM Cloud destaca por su integración con tecnologías avanzadas como la inteligencia artificial a través de IBM Watson y el uso de blockchain para casos de uso corporativos.

- **Oracle Cloud Infrastructures**

OCI es especialmente atractivo para organizaciones que ya dependen del ecosistema Oracle, ya que proporciona compatibilidad total con Oracle Database y sus soluciones empresariales, facilitando así una transición fluida a la nube sin sacrificar eficiencia ni control.

Selección de la alternativa de proveedor tecnológico

A partir del análisis de las necesidades tecnológicas identificadas en CVP Centro Constructor Vásquez Pacheco, se considera necesario seleccionar un proveedor que permita soportar de manera integral la solución de Business Intelligence propuesta. La alternativa debe responder a los requerimientos actuales de la organización, pero también ofrecer capacidad de

crecimiento a medida que aumente el volumen de información, se incorporen nuevas fuentes de datos y se amplíen las necesidades analíticas de la empresa. En este sentido, la selección no debe limitarse únicamente al costo inicial de la herramienta, sino considerar aspectos como integración de datos, capacidad de almacenamiento, procesos ETL/ELT, seguridad, escalabilidad, administración y compatibilidad con las herramientas de visualización.

Actualmente, CVP utiliza SAGA como principal fuente de información para los procesos de ventas, compras, inventarios y contabilidad. La información generada por estas operaciones constituye la base para el desarrollo de la solución BI; sin embargo, cuando se requieren análisis específicos, los datos pueden ser exportados hacia Microsoft Excel para realizar procesos adicionales de organización y consolidación. Asimismo, el diagnóstico evidencia que la empresa no cuenta actualmente con un Data Warehouse orientado al análisis histórico y multidimensional de la información. Esta situación demuestra la necesidad de incorporar una infraestructura tecnológica que permita centralizar los datos y automatizar su integración, disminuyendo la dependencia de procesos manuales.

Después de considerar las alternativas disponibles, se selecciona Microsoft Azure como el proveedor tecnológico que mejor se adapta a las necesidades de CVP. La principal razón de esta elección es que Azure permite construir una arquitectura integral de datos en la nube, incorporando servicios para la extracción, transformación, almacenamiento, procesamiento y análisis de información. Dentro de esta arquitectura, Azure Data Factory puede desempeñar el papel de herramienta de integración y orquestación de datos, mientras que Azure Synapse Analytics puede utilizarse como plataforma para el almacenamiento y procesamiento analítico,

complementándose posteriormente con Power BI para la visualización de indicadores y dashboards.

Azure Data Factory resulta particularmente adecuado para la propuesta debido a que es un servicio administrado de integración de datos que permite automatizar procesos ETL y ELT, mover información entre diferentes fuentes y destinos y establecer flujos de trabajo programados. Microsoft señala que Data Factory permite integrar datos provenientes de diferentes almacenes y construir pipelines para mover y transformar información a escala.

Además, dispone de conectores para múltiples fuentes y puede trabajar tanto con información almacenada en la nube como con entornos locales. Esta característica resulta relevante para CVP, debido a que la solución propuesta debe partir de la información generada en SAGA y posteriormente incorporar otras fuentes, como archivos de Excel y datos externos relacionados con clientes, proveedores y comportamiento comercial.

Otro factor determinante es la posibilidad de mantener una arquitectura flexible respecto al origen de los datos. Azure Data Factory dispone de diferentes opciones de Integration Runtime, incluyendo alternativas para escenarios en la nube y entornos locales, lo que permite adaptar la arquitectura a la infraestructura existente de CVP. Esta capacidad resulta especialmente importante porque la empresa actualmente depende de SAGA como sistema transaccional principal y, por tanto, la implementación de BI no debería exigir una sustitución inmediata del sistema operativo existente. En lugar de reemplazar SAGA, Azure permitiría utilizarlo como fuente de datos y construir sobre él una capa analítica independiente.

Para la construcción del repositorio analítico, Azure Synapse Analytics constituye una alternativa adecuada para el Data Warehouse propuesto. Microsoft describe Synapse como un

servicio de análisis que integra capacidades de almacenamiento de datos empresariales, procesamiento de grandes volúmenes de información y herramientas de integración, además de presentar una integración directa con otros servicios del ecosistema Azure y Power BI. Esta característica se ajusta a la arquitectura planteada para CVP, en la cual los datos provenientes de ventas, compras, inventarios, cuentas por cobrar y cuentas por pagar deben consolidarse en un entorno centralizado para posteriormente ser utilizados en análisis multidimensionales.

La elección de Azure también presenta una ventaja importante frente a una solución compuesta por diferentes proveedores y herramientas independientes. La integración entre Azure Data Factory, Azure Synapse Analytics y Power BI permite construir un ecosistema tecnológico coherente, en el cual cada componente cumple una función específica dentro del ciclo de vida de los datos. Data Factory puede encargarse de la extracción, transformación y orquestación; Synapse puede proporcionar la infraestructura analítica; y Power BI puede utilizar la información consolidada para generar indicadores, reportes y tableros de control. Microsoft documenta específicamente la integración entre Azure y Power BI para escenarios de análisis empresarial, tanto para información almacenada en la nube como para determinados escenarios con datos locales.

Definir nuestro modelo de cloud computing como SaaS, PaaS o IaaS

Para la implementación de Business Intelligence en CVP Centro Constructor Vásquez Pacheco se propone adoptar PaaS (Platform as a Service) como modelo principal de Cloud Computing, utilizando Microsoft Azure como proveedor tecnológico y complementándolo con herramientas SaaS para el consumo y visualización de la información.

La elección de PaaS responde a que CVP requiere una plataforma capaz de integrar las diferentes fuentes de información de la organización, principalmente ERP SAGA, archivos de Excel y fuentes digitales, para posteriormente ejecutar procesos de extracción, transformación y carga (ETL), consolidar un Data Warehouse corporativo y generar Data Marts especializados para las diferentes áreas.

PaaS permite desarrollar esta arquitectura sin que CVP tenga que asumir directamente la administración de servidores y gran parte de la infraestructura tecnológica subyacente. Esto resulta especialmente conveniente para una empresa cuyo negocio principal es la comercialización de productos para la construcción y no la gestión de infraestructura informática.

Flujo de la integración de nuestra arquitectura big data con el datawarehouse o estructura bi alternativa

En la arquitectura propuesta para CVP Centro Constructor Vásquez Pacheco, el flujo de integración entre la arquitectura Big Data / sistemas fuentes y la estructura de Data Warehouse (DW) y Data Marts (DM) se organiza en una secuencia estructurada de cinco etapas principales:

Origen y Captura de Datos (Sistemas Fuente)

El flujo comienza en los sistemas de origen internos y externos que generan datos diariamente:

- **ERP SAGA (Fuente Interna Principal):** Registra en tiempo real las operaciones transaccionales de ventas, facturación electrónica, clientes, compras, movimientos de inventario y contabilidad.
- **Archivos Excel (Fuentes Secundarias Internas):** Almacenan métricas personalizadas, datos de Talento Humano y encuestas de satisfacción.
- **Redes Sociales y Marketing Digital (Fuentes Externas):** Capturan métricas de alcance, interacción y tráfico mediante APIs.

Extracción, Transformación y Carga (Proceso ETL)

El proceso ETL, gestionado a través de **SQL Server Integration Services (SSIS)**, actúa como el puente de integración entre las fuentes operativas y el almacenamiento analítico:

Extracción (Staging Area): Se conecta a la base de datos relacional de SAGA (SQL Server/PostgreSQL), lee archivos Excel programados y consulta APIs externas.

Transformación:

- **Limpieza y Estandarización:** Eliminación de duplicados y unificación de criterios (p. ej., corregir inconsistencias en cédulas/nombres de clientes y descripciones de SKU de materiales).
- **Normalización y Enriquecimiento:** Cálculo de métricas derivadas (márgenes, rotación) y manejo de dimensiones de cambio lento (SCD Tipo 2).
- **Validación de Calidad de Datos:** Verificación de reglas de completitud y consistencia.

Carga: Carga inicial histórica (últimos 3-5 años) y cargas incrementales automáticas.

Data Warehouse Corporativo (Modelo Centralizado)

La información consolidada y limpia se almacena en el Data Warehouse Comercial / Corporativo. Este repositorio centraliza el historial transaccional con la mínima granularidad y estructura los datos mediante Tablas de Dimensión compartidas:

- **Dimensiones Clave:** DIM_TIEMPO, DIM_CLIENTE, DIM PRODUCTO, DIM PROVEEDOR, DIM_EMPLEADO, DIM SUCURSAL (Cañar y El Tambo), DIM CANAL_VENTA, DIM CATEGORIA PRODUCTO, DIM_METODO_PAGO y DIM CAMPANA MARKETING.

Data Marts Especializados (Estructura BI Alternativa / Subconjuntos)

Para optimizar las consultas analíticas de áreas específicas sin impactar el DW central, el Data Warehouse alimenta Data Marts especializados estructurados alrededor de Tablas de Hechos y dimensiones relevantes:

- **Data Mart Comercial (DM_COMERCIAL):**
  - **Tablas de Hechos:** HECHOS VENTAS RESUMEN, HECHOS CLIENTES, HECHOS MARKETING, HECHOS VENTAS CRUZADAS, HECHOS VENTAS CORPORATIVAS.
  - **Soporte:** Análisis de ticket promedio, tasa de conversión, fidelización, cumplimiento de metas de ventas y ROI de marketing.
- **Data Mart Bodega / Logística (DM_BODEGA):**

- **Tablas de Hechos:** HECHOS INVENTARIO DIARIO, HECHOS MOVIMIENTOS, HECHOS RECEPCIONES, HECHOS DESPACHOS, HECHOS PICKING, HECHOS PÉRDIDAS, HECHOS CONTROLES SEGURIDAD, HECHOS ACCIDENTES.
- **Soporte:** Rotación de stock, tiempos de despacho/recepción, exactitud de inventario y eficiencia de picking.

Explotación de Datos y Capa de Presentación (Dashboards / BI)

Los Data Marts alimentan la plataforma de Business Intelligence (Power BI o equivalente) para la toma de decisiones mediante tableros e indicadores clave (KPIs) segmentados por perfil de usuario:

- **Dashboard Ejecutivo / Gerencial:** Visión consolidada de alto nivel (crecimiento de ventas, margen bruto/operativo, ROA, ciclo de caja).
- **Dashboard Comercial y Ventas:** Monitoreo diario por vendedor, sucursal, canal y segmento.
- **Dashboard de Bodega e Inventarios:** Monitoreo de rotación de SKU, stock mínimo y productos próximos a vencer.
- **Dashboard Operativo y Logístico:** Eficiencia de despachos, picking y tiempos de entrega.
- **Dashboard Financiero y Transversal:** Flujo de caja, días de inventario (DIO) y rotación de cartera.

Análisis crítico de Frameworks Big Data open source

La presente sección desarrolla la evaluación crítica de los principales frameworks y tecnologías Big Data de código abierto analizados, con el propósito de seleccionar aquellos componentes que resulten pertinentes para la arquitectura de Business Intelligence de CVP Centro Constructor Vásquez Pacheco.

La selección no se limita a identificar las herramientas más potentes del mercado, sino a determinar cuáles responden de manera justificada a las necesidades reales de la empresa, en coherencia con las decisiones tecnológicas adoptadas previamente (Microsoft Azure como proveedor cloud en modelo PaaS + SaaS, Azure Synapse Analytics como plataforma del Data Warehouse, SSIS/ADF como herramienta ETL y Power BI como capa de visualización)

Criterios de evaluación

La evaluación se estructuró mediante una matriz de decisión multicriterio ponderada. Los criterios y pesos se definieron a partir del diagnóstico organizacional y de la arquitectura propuesta:

Tabla 33

Criterios de Evaluación

Criterio	Peso	Justificación
A. Adecuación al perfil de datos de CVP (Volumen, Velocidad, Variedad)	25%	La empresa registra un volumen de información mediano (aprox. USD 2,5 millones en ventas anuales, 9.000 clientes activos, un catálogo de miles de SKUs), una velocidad de procesamiento predominantemente batch (carga diaria programada a las 8:00) y una

		variedad de datos en crecimiento (estructurados en SAGA, semiestructurados en Excel y APIs, y no estructurados proyectados: reseñas, comentarios y quejas).
B. Alineación con la arquitectura definida (Azure, Synapse, SSIS, Power BI)	20%	Los frameworks seleccionados deben integrarse con el ecosistema Microsoft ya adoptado, evitando la introducción de plataformas tecnológicas paralelas que incrementen la complejidad operativa.
C. Costo total de propiedad y perfiles de talento disponibles	20%	CVP cuenta con 16 colaboradores y una estructura de BI propuesta de tres perfiles. Las tecnologías que exigen administradores de clusters o especialistas exclusivos (ej. administradores Hadoop) resultan inviables.
D. Escalabilidad y evolución futura de la empresa	15%	La empresa posee un objetivo estratégico de expansión a nivel nacional; las tecnologías deben permitir crecer sin rediseñar la arquitectura.
E. Madurez tecnológica, soporte y comunidad	10%	Se priorizan tecnologías consolidadas, con trayectoria empresarial y comunidad activa, que reduzcan el riesgo de obsolescencia.
F. Simplicidad de implementación y mantenimiento	10%	La curva de aprendizaje y el esfuerzo de operación continua deben ser asumibles por el equipo interno.

Nota : Elaboración propia, adaptado del organigrama institucional de CVP Centro Constructor Vásquez Pacheco (2026).

Tabla 34.

Matriz de evaluación ponderada

Framework / Tecnología	A (25%)	B (20%)	C (20%)	D (15%)	E (10%)	F (10%)	PUNTAJE PONDERADO
Python (pandas, scikit-learn)	5	5	5	4	5	4	4.75
MongoDB / Cosmos DB	4	4	4	4	5	4	4.10
Elasticsearch	3	4	3	4	5	3	3.55
Apache Spark	2	5	3	5	5	2	3.55
R	4	3	3	3	4	2	3.25
Kafka / Kafka Streams	1	3	2	4	5	2	2.55
Hive	2	3	2	2	3	3	2.40
Redis	1	2	2	2	5	3	2.15
Keras / PyTorch	1	2	2	2	5	2	2.05
HBase / Cassandra	1	2	1	3	4	1	1.80
Apache Flink	1	2	1	3	2	1	1.60
Apache Hadoop / MapReduce	1	1	1	3	3	1	1.50
Apache Storm	1	1	1	2	3	1	1.35
Flume / Sqoop	1	1	1	1	1	1	1.00

Nota: Elaboración propia, adaptado del organigrama institucional de CVP Centro Constructor Vásquez Pacheco (2026).

Análisis crítico y justificación de la selección

Frameworks seleccionados para la Fase 1 (implementación inmediata):

Python (pandas, scikit-learn, statsmodels) — puntaje 4,75. Constituye la selección de mayor prioridad. Cubre la analítica predictiva requerida por los objetivos del proyecto (pronósticos de demanda, segmentación RFM de la cartera de 9.000 clientes, detección de productos de baja rotación), integra nativamente con Azure Machine Learning y permite la ejecución de scripts dentro de Power BI. Con un único Analista de Business Intelligence en la estructura propuesta, la consolidación de la preparación y el modelado de datos en un solo lenguaje de propósito general maximiza la mantenibilidad. Se descartó R (3,25) como alternativa exclusiva por su sintaxis de orientación académica y su menor versatilidad fuera del ámbito estadístico, aunque se reconoce su validez como complemento.

MongoDB (Azure Cosmos DB como servicio administrado equivalente) puntaje 4,10. Seleccionado en modalidad de reserva activa para la Fase 2. Responde a la única dimensión en la que el Data Warehouse relacional es estructuralmente débil: la variedad de datos no estructurados (reseñas de clientes en Google Business, comentarios en redes sociales, registros de quejas). Su modelo de documentos sin esquema tolera la ausencia de formato que, según el diagnóstico, caracteriza a estas fuentes (IBM, 2026).

Frameworks seleccionados en modalidad de reserva condicionada:

Elasticsearch puntaje 3,55. Reservado para la Fase 2, como capa de indexación y búsqueda de texto completo sobre el corpus de reseñas y comentarios. Adicionalmente, su capacidad de búsqueda semántica aporta un beneficio inmediato al problema de calidad de datos identificado en el diagnóstico: la inconsistencia en la nomenclatura de productos y clientes dificulta la localización de registros en SAGA, problema que la indexación flexible de Elasticsearch mitiga. Se complementa con MongoDB sin solapamiento funcional: MongoDB almacena, Elasticsearch indexa y busca.

Apache Spark puntaje 3,55. No se implementa en la Fase 1 por la inadecuación actual (procesamiento distribuido en memoria diseñado para volúmenes masivos), pero se constituye en la ruta tecnológica natural de evolución ante el crecimiento del volumen de datos (expansión nacional, incorporación de e-commerce o múltiples sucursales). Resulta especialmente coherente con la arquitectura elegida, dado que Azure Synapse Analytics incorpora un motor Spark integrado, lo que permitiría activar esta capacidad sin incorporar nuevos proveedores.

Frameworks rechazados con justificación explícita:

Apache Hadoop/MapReduce (1,50) y el ecosistema asociado Flume/Sqoop (1,00): tecnologías en declive desde mediados de la década de 2010 (Sqoop fue retirado del Apache Attic en 2021), que exigen administración de clusters especializada e infraestructura propia; su función de almacenamiento distribuido es asumida por Azure Synapse con menor costo total de propiedad. Apache Storm (1,35) y Apache Flink (1,60): orientados al procesamiento de flujos de millones de eventos por segundo en tiempo real; CVP opera con carga batch diaria, por lo que su incorporación carece de justificación de negocio (señal de alerta: la adopción de

tecnología sin necesidad que la respalde es un riesgo clásico de sobre ingeniería en proyectos BI de pymes). Kafka/Kafka Streams (2,55): bus de eventos cuya utilidad aparecería únicamente ante escenarios futuros de alta frecuencia (IoT, sensores, e-commerce de gran tráfico); se registra como opción de arquitectura condicionada. HBase/Cassandra (1,80) y Redis (2,15): bases de datos clave-valor diseñadas para tasas de escritura masivas y latencias de milisegundos en aplicaciones de alto tráfico; sin caso de uso en la arquitectura analítica de CVP. Keras/PyTorch (2,05): frameworks de aprendizaje profundo sin caso de uso justificado (ausencia de imágenes, procesamiento de lenguaje natural masivo o reconocimiento de patrones complejos); los modelos de pronóstico de series temporales de ventas se resuelven con técnicas estadísticas y de machine learning clásico en Python. Hive (2,40): su aporte conceptual —una interfaz SQL sobre grandes volúmenes de datos— ya es provista por Synapse SQL; se asimila como referencia conceptual, no como componente a incorporar.

Arquitectura resultante y hoja de ruta gobernada

La selección crítica conduce a una arquitectura de "Business Intelligence con componentes Big Data quirúrgicos", en la que el núcleo permanece en el Data Warehouse en la nube y los frameworks open source se incorporan de manera puntual, por fases y con criterios de activación medibles, en lugar de constituir una plataforma Big Data paralela:

- FASE 1 (implementación inmediata): SAGA / Excel / APIs → SSIS / Azure Data Factory (ETL batch) → Azure Synapse Analytics (Data Warehouse) → Power BI + Python (analítica predictiva).

- FASE 2 (activación condicionada al despliegue de la capa de satisfacción del cliente): incorporación de MongoDB/Cosmos DB (datos no estructurados) y Elasticsearch (indexación y minería de texto).
- FASE 3 (activación condicionada al crecimiento del volumen): motor Spark integrado en Synapse ante escenarios de múltiples sucursales, e-commerce o fuentes de alta frecuencia.

Los criterios de activación convertirán la selección en una hoja de ruta gobernada y no en una adquisición tecnológica anticipada. Se propone, como referencia, activar la Fase 3 cuando el volumen de datos procesado supere el umbral de gestión eficiente del entorno relacional (p. ej., ingesta superior a 50 GB diarios o más de 50 millones de registros anuales nuevos) o cuando se incorporen fuentes de streaming; y activar la Fase 2 cuando se formalice el proceso de captura sistemática de reseñas y comentarios de clientes.

CAPÍTULO 8

Complementación del ecosistema de big data y analítica avanzada

El proyecto de titulación para CVP Centro Constructor Vásquez Pacheco plantea la implementación de una solución de Business Intelligence (BI) orientada a integrar datos comerciales, operativos, de inventario y financieros provenientes de su ERP actual (SAGA) y fuentes complementarias (archivos Excel y redes sociales).

Actualmente, el flujo de datos dentro de CVP presenta las siguientes características y limitaciones:

- **Fuente Principal:** ERP SAGA sobre base de datos relacional (SQL Server / PostgreSQL).
- **Integración y ETL:** Procesos ETL propuestos con SQL Server Integration Services (SSIS).
- **Almacenamiento Analítico:** Data Warehouse Corporativo centralizado con Data Marts funcionales (*DM COMERCIAL* y *DM BODEGA*).
- **Procesamiento Actual:** Dependencia de exportaciones manuales a Microsoft Excel y carga de datos incremental diaria.
- **Limitaciones Detectadas:** Falta de análisis predictivo en tiempo real, imposibilidad de procesar datos no estructurados a gran escala (como imágenes de recepciones de obra o texto de interacciones con clientes), y riesgos de escalabilidad ante la expansión nacional planificada.

Para evolucionar desde una arquitectura tradicional de Business Intelligence hacia una Arquitectura Moderna de Big Data y Analítica Avanzada (Lakehouse), a continuación, se presenta una propuesta completa de complementación tecnológica.

Complementación del ecosistema de big data

Para responder a las demandas del crecimiento nacional, la variabilidad de la demanda en el sector ferretero y de la construcción, y la necesidad de analítica predictiva, se expande la arquitectura mediante la inclusión de herramientas de Big Data:

Tabla 35.-

Complementación del Ecosistema de Big Data

[Fuentes Heterogéneas]	→	**[Capa de Ingesta]**	→	**[Capa de Almacenamiento/ Procesamiento]**
- ERP SAGA (SQL)	→	- Apache Kafka (Events)	→	- Data Lake (AWS S3 / MinIO)
- Redes Sociales (APIs)	→	- Apache NiFi (Batch/APIs)	→	- Delta Lake / Apache Iceberg
- Excel / Documentos	→			Dara Warehuse

Nota: Elaboración propia, adaptado del organigrama institucional de CVP Centro Constructor Vásquez Pacheco (2026).

Arquitectura Data Lakehouse

Se propone transicionar de una arquitectura puramente de Data Warehouse a un Data Lakehouse, el cual combina la flexibilidad y bajo costo de almacenamiento de un Data Lake con las capacidades de gestión de transacciones ACID, esquema y rendimiento de un Data Warehouse.

- **Almacenamiento Cloud / On-Premise:** Utilización de Amazon S3 o MinIO (almacenamiento de objetos) para guardar datos crudos (estructurados, semiestructurados como JSON/XML, y no estructurados como imágenes de comprobantes o audios de servicio).

- **Capa de Formato Abierto (Delta Lake / Apache Iceberg):** Garantiza confiabilidad de los datos, viajes en el tiempo (time travel para auditar cambios de inventario) y procesamiento eficiente mediante archivos parquet optimizados.

Motores de Procesamiento Masivo y en Tiempo Real

- **Apache Spark / PySpark:** Sustituye el procesamiento rígido de scripts tradicionales para la transformación de grandes volúmenes de datos. Permite ejecutar transformaciones distribuidas sobre las tablas de ventas, movimientos de bodega e historial de clientes.

- **Apache Kafka:** Implementado para el procesamiento de eventos en tiempo real. Captura eventos críticos de manera instantánea, tales como:

    - Variaciones críticas de inventario (quiebres de stock en bodegas de Cañar o El Tambo).

- Eventos de facturación electrónica y cambios en la cartera de clientes a crédito.
- Clickstream y solicitudes de cotizaciones en la plataforma web o chatbot.

- **Apache NiFi:** Integrado en paralelo con SSIS para gestionar flujos de datos no estructurados o ingesta desde fuentes externas (APIs de proveedores como Adelca o UCEM, scraping de precios del mercado y métricas de redes sociales).

Tabla 36

SQL	PYTHON	R	SCALA / BASH
(Modelado y BI)	(ETL y ML)	(Estadística Avanzada)	(Orquestación / Spark)

Nota: Elaboración propia (2026).

Selección y justificación de lenguajes de programación

Para operativizar el ecosistema de datos, se definen lenguajes específicos según la capa del proceso analítico:**LENGUAJES DE PROGRAMACIÓN**

SQL (Structured Query Language)

- **Uso Principal:** Consultas analíticas, definición de vistas, agregaciones en los Data Marts y mantenimiento de la lógica de negocio en el ERP SAGA y Data Warehouse.

- **Justificación:** Es el estándar universal para manipulación de datos relacionales, garantizando compatibilidad con Power BI, SSIS y los sistemas existentes.

Python

- **Uso Principal:** Ingeniería de datos (Data Pipelines), automatización de transformaciones con PySpark y Pandas, y ciencia de datos (Machine Learning).
- **Librerías Específicas:**
    - Pandas / Polars: Para manipulación y limpieza rápida de archivos complementarios en Excel.
    - Scikit-Learn / XGBoost: Para modelado predictivo (proyección de demanda de materiales de construcción y segmentación RFM de clientes).
    - SQLAlchemy / PySpark: Para conectar las canalizaciones de datos entre SAGA, el Data Lake y la capa analítica.

R (Opcional / Analítica Financiera)

- **Uso Principal:** Análisis estadístico avanzado de series temporales para predicción de ventas mensuales y modelado econométrico frente a la inflación de insumos de construcción.

- **Justificación:** Complementario para la evaluación de riesgo crediticio de constructoras y clientes corporativos.

Bash / Shell Scripting

- **Uso Principal:** Automatización de tareas a nivel del sistema operativo, orquestación de respaldos, ejecución de scripts de validación diaria y gestión de contenedores Docker.

Ecosistema de bases de datos y búsqueda

Para evitar cuellos de botella y optimizar los tipos de datos manejados por CVP, se adopta un enfoque de Persistencia Políglota (Polyglot Persistence):

Tabla 37.- *Ecosistemas de Bases de Datos y Búsqueda*

Tipo de Base de Datos	Tecnología Propuesta	Rol dentro del Ecosistema de CVP	Justificación Técnica
Relacional / OLTP	PostgreSQL / SQL Server	Sistema ERP SAGA (Base operacional).	Garantiza transaccionalidad ACID para facturación, compras y contabilidad.
Relacional / OLAP	ClickHouse / Snowflake / PostgreSQL (Citus)	Almacenamiento del Data Warehouse Corporativo.	Consultas analíticas ultra rápidas con almacenamiento columnar para calcular KPIs agregados.

NoSQL Orientada a Documentos	MongoDB	Gestión de Catálogo extendido, cotizaciones complejas y Logs de eventos del sistema BI.	Permite esquemas flexibles para productos con múltiples atributos técnicos (cerámicas, accesorios, maquinaria).
NoSQL Clave-Valor / Cache	Redis	Memoria caché para dashboards en tiempo real e historial de sesiones de usuarios.	Acelera la respuesta del dashboard gerencial, logrando tiempos de carga menores a 3 segundos.
Motor de Búsqueda y Logs	Elasticsearch	Búsqueda avanzada de SKUs y análisis de logs de auditoría de datos.	Resuelve la falta de estandarización en nombres de productos mediante búsqueda difusa (fuzzy search).

Nota: Elaboración propia, adaptado del organigrama institucional de CVP Centro Constructor Vásquez Pacheco (2026).

CAPÍTULO 9

Identificar y definir el objetivo de aplicación del modelo de aplicación de Data Science y Machine Learning en el proyecto

Objetivo y alcance del modelo

El objetivo es aprovechar los registros históricos de CVP Centro Constructor Vásquez Pacheco para anticipar la demanda de productos y reconocer patrones de compra de sus clientes. A diferencia de los dashboards descriptivos ya propuestos, Data Science y Machine Learning permitirán generar pronósticos y segmentos que orienten decisiones de abastecimiento y acciones comerciales. El modelo se plantea como una fase posterior a la integración y depuración de los datos del ERP SAGA y no se presenta como una solución ya implementada.

Se priorizan dos casos de uso:

1. Pronóstico de demanda por producto o categoría, para apoyar la reposición y reducir excesos o faltantes de inventario.
2. Segmentación de clientes según recencia, frecuencia y valor monetario (RFM), con posibilidad de aplicar técnicas de clustering una vez comprobada la calidad de los registros.

Las recomendaciones automáticas de productos y el análisis de asociaciones entre ellos se consideran una evolución posterior y no un resultado inicial garantizado.

Fuentes de datos utilizadas

La Tabla 1 presenta las fuentes internas consideradas, sus variables prioritarias y su uso analítico.

Tabla 38

Fuentes internas de datos del modelo

Fuente	Variables prioritarias	Uso analítico
ERP SAGA – ventas	Fecha, factura, SKU, cantidad, precio, cliente, sucursal	Series históricas de demanda y análisis RFM
ERP SAGA – inventario y compras	Existencias, entradas y salidas, costos, fechas y plazos de reposición	Contrastar pronósticos con el stock y las decisiones de compra
Maestro de clientes	Identificador, historial de compras y canal	Segmentación y seguimiento de clientes
Excel / campañas (si están disponibles)	Promociones, campañas y fechas relevantes	Explicar variaciones de demanda y evaluar acciones comerciales

Nota. ERP = sistema de planificación de recursos empresariales; SKU = código de identificación del producto en inventario; RFM = recencia, frecuencia y valor monetario. Elaboración propia.

Antes de entrenar los modelos será necesario validar la disponibilidad histórica, la consistencia de SKU y clientes, las devoluciones, los períodos sin ventas y los posibles quiebres de stock. La tecnología de base de datos específica de SAGA y la cobertura histórica efectiva deberán confirmarse con el proveedor o el área de sistemas.

Fuentes externas de datos

Además de los registros internos, CVP puede incorporar fuentes externas que aporten contexto a los modelos predictivos (Tabla 2). Su integración se realizará únicamente cuando exista acceso autorizado, calidad suficiente y una relación comprobable con la demanda o el comportamiento de los clientes.

Tabla 39

Fuentes externas de datos complementarias

Fuente externa	Información relevante	Aplicación en Data Science
Redes sociales y canales digitales (APIs autorizadas)	Alcance, interacciones, tráfico y conversiones de campañas	Contrastar campañas con cambios en la demanda y en los segmentos de clientes
Información de proveedores	Listas de precios, disponibilidad, plazos de entrega y documentos de despacho	Ajustar decisiones de reposición y evaluar riesgos de abastecimiento
Solicitudes y comentarios de clientes	Cotizaciones, requerimientos de obra, preferencias y comentarios recibidos por canales habilitados	Identificar necesidades recurrentes y enriquecer perfiles de compra, previa depuración

Nota. API = interfaz de programación de aplicaciones. El uso de estas fuentes está sujeto a acceso autorizado y a la calidad de los datos. Elaboración propia.

Área de aplicación y departamentos de negocios implicados

La Tabla 3 distribuye la aplicación del modelo y las responsabilidades entre las áreas de la organización.

Tabla 40

Áreas de negocio implicadas en la aplicación del modelo

Área	Aplicación	Responsabilidad
Comercial / Marketing	Segmentación RFM y oportunidades de fidelización	Definir campañas, evaluar conversión y medir ventas incrementales
Bodega / Compras	Pronóstico y reposición de inventario	Revisar alertas y validar restricciones de abastecimiento
Finanzas	Evaluación de margen y capital de trabajo	Verificar el impacto económico y evitar el doble conteo
Gerencia / responsable BI	Seguimiento del desempeño del modelo	Aprobar pilotos, calidad de datos y decisiones

Nota. RFM = recencia, frecuencia y valor monetario; BI = Business Intelligence. Elaboración propia.

Identificación de la mejora esperada

El cambio propuesto consiste en pasar de decisiones apoyadas principalmente en reportes históricos a decisiones que incorporan estimaciones de demanda y perfiles de clientes. La mejora se comprobará mediante un piloto y la comparación de indicadores antes y después

de su aplicación, sin atribuir automáticamente al modelo todos los cambios del negocio (Tabla 41).

Tabla 41

Mejora esperada e indicadores de validación por ámbito de aplicación

Ámbito de aplicación	Mejora esperada	Indicadores para validar
Pronóstico de demanda	Menos exceso de inventario y menor riesgo de faltantes	Error del pronóstico (WAPE o MAE), días de inventario y quiebres de stock
Segmentación de clientes	Campañas comerciales más pertinentes	Ventas incrementales frente a un grupo comparable, ticket promedio y recompra
Gestión gerencial	Priorización de decisiones con evidencia	Uso de recomendaciones, frecuencia de revisión y acciones ejecutadas

Nota. WAPE = error porcentual absoluto ponderado; MAE = error absoluto medio. Elaboración propia.

Mejora esperada en ventas

En el área comercial, se espera que el modelo de Machine Learning permita identificar patrones de compra de los clientes y anticipar la demanda por categoría de producto. A partir del historial de facturación de SAGA se propone segmentar a los clientes mediante el análisis RFM para orientar las campañas comerciales. El objetivo es apoyar decisiones comerciales más precisas, no sustituir el criterio de los vendedores.

La cuantificación de este potencial comercial, a partir de la línea base de ventas reportada por CVP para 2025, se desarrolla en el Business Case de este capítulo, donde se presenta como un escenario y no como un efecto comprobado del modelo.

La mejora se validará con indicadores de crecimiento de ventas, ticket promedio y recompra, comparando un grupo piloto con su desempeño histórico o con un grupo comparable. Para atribuir resultados al modelo deberán considerarse las promociones, la estacionalidad, las variaciones de precios y la disponibilidad de inventario.

Cuantificación de la mejora en el área de aplicación a través de un Business Case

Para determinar la viabilidad de la implementación del modelo de Data Science y Machine Learning se desarrolla un Business Case financiero, a partir de los datos reales de CVP correspondientes a 2025.

Línea base financiera - operativa

La Tabla 5 resume los indicadores de partida y su clasificación según el origen del dato.

Tabla 42

Línea base financiera y operativa de CVP en 2025

Indicador	Valor	Clasificación
Ventas 2025	USD 3.077.453,94	Dato proporcionado por CVP
Compras 2025	USD 2.679.367,38	Dato proporcionado por CVP
Margen bruto actual	12,94 % (≈ 13 %)	Supuesto: punto medio del rango
Incremento supuesto de ventas	3%	Dato proporcionado por CVP
Ventas adicionales	USD 92.323,62	Dato proporcionado por CVP
Ventas Proyectadas	USD 3.169.777,56	Proyecciones
Margen comercial del escenario	18%	Proyecciones
Beneficio bruto sobre las ventas proyectadas	USD 570.559,96	Proyecciones

Inventario promedio	USD 349.859,08	Dato proporcionado por CVP
Cuentas por cobrar	USD 70.508,19	Dato proporcionado por CVP
Cartera vencida	USD 30.656,09	Dato proporcionado por CVP
Días de inventario	Aprox. 80 días	Estimación derivada; depende del margen supuesto

Nota. Valores en dólares estadounidenses (USD). Los días de inventario se calculan como inventario promedio / costo de ventas × 365, con el costo de ventas obtenido a partir del margen supuesto. Elaboración propia con datos proporcionados por CVP.

Escenario de ventas y segmentación

Se plantea, exclusivamente para evaluar la oportunidad, un incremento del 3 % en las ventas atribuible a acciones basadas en analítica. Sobre ventas de USD 3,077,453.94, ello representa USD 92,323.62 de facturación adicional. Con un margen bruto supuesto de 18%, el beneficio bruto incremental sería de USD 172,473.40 al año. Esta cifra no equivale a utilidad neta y requiere validación mediante un piloto comercial o la comparación con clientes similares no intervenidos.

Escenario de inventario y pronóstico

Utilizando el inventario promedio de USD 349.859,08 y el costo de ventas proyectado de USD 2.599.217,60, los días de inventario estimados son aproximadamente 49 días.

Si la analítica predictiva permitiera reducir el inventario a 40 días, manteniendo el nivel de servicio:

- Inventario objetivo: USD 284.845,76.
- Capital potencialmente liberado: USD 65.013,32.

Esta liberación corresponde a capital de trabajo, no a ingresos ni a utilidad neta.

Resultado económico y límites de la estimación

La Tabla 42 consolida los resultados de ambos escenarios y el tratamiento financiero que corresponde a cada uno.

Tabla 43

Resultado económico de los escenarios y su tratamiento financiero

Concepto	Impacto del escenario	Tratamiento financiero
Facturación incremental (+3 %)	USD 92.323,62 al año	Ingreso adicional supuesto
Margen bruto incremental	USD 172,473.40 al año	Beneficio bruto antes de costos adicionales
Inventario liberado (40a 49 días)	USD 65,013.32	Liberación potencial de capital, no utilidad
Cartera vencida	43,48 % de las cuentas por cobrar	Indicador contextual; no se monetiza como resultado del modelo priorizado

Nota. Los escenarios se calculan con las ventas de 2025 y el margen bruto supuesto de 18 % de la Tabla 5. El porcentaje de cartera vencida corresponde a la cartera vencida dividida entre las cuentas por cobrar. Elaboración propia.

Ruta de implementación y validación

La implementación se plantea en las siguientes etapas:

1. Depuración de datos históricos.
2. Definición de indicadores base.
3. Entrenamiento y comparación con un pronóstico sencillo.
4. Piloto de reposición y segmentación.
5. Medición de resultados frente a una referencia comparable.
6. Ajuste y ampliación gradual.

Se recomienda iniciar con los productos de mayor rotación y un conjunto delimitado de clientes. En el pronóstico se evaluará el error sobre períodos no utilizados para el entrenamiento; en las campañas se medirá el resultado frente a un grupo comparable. Esta secuencia permite separar el aporte real de los modelos de las variaciones estacionales, las promociones y los cambios de precios.

CAPÍTULO 10

Identificar y definir el objetivo de aplicación del modelo de aplicación de Real Time

Los consumidores actuales exigen respuestas inmediatas, personalización constante y procesos fluidos. El análisis de datos en tiempo real permite a las empresas ofrecer experiencias de cliente altamente adaptadas en el momento adecuado (Emilio, 2025).

Se aplicará el análisis de datos en tiempo real para conocer la satisfacción del cliente en el proceso posterior a la compra y verificar, cuando corresponda, si los productos llegaron a tiempo a su destino. De este modo se busca personalizar la experiencia del cliente en tiempo real.

Fuentes de datos utilizadas

Las fuentes de datos consideradas para este caso de uso son las siguientes; su disponibilidad se validará antes del piloto:

- Sistema de ventas ERP: detalle de la compra de productos y hora de la transacción.
- Sistema de logística: estado y ubicación del pedido, para saber si el producto está en camino, entregado o retrasado.
- Encuestas de satisfacción: opiniones y comentarios que el cliente decida aportar.
- CRM: historial de compras anteriores, reclamos y preferencias del cliente.

Área de aplicación y departamentos de negocio implicados

En CVP Centro Constructor Vásquez Pacheco, la aplicación prioritaria del análisis en tiempo real se concentra en el proceso posterior a la venta: preparación, despacho, transporte, confirmación de entrega y atención de incidencias. Esta delimitación responde al objetivo de conocer oportunamente el estado del pedido y mejorar la experiencia del cliente.

El caso de uso se plantea inicialmente para pedidos con entrega a domicilio en las operaciones de Cañar y El Tambo. La solución conectaría los eventos de venta del ERP SAGA con registros logísticos, actualizaciones de ubicación o estado del transporte y comunicaciones del cliente. Su alcance real dependerá de que CVP disponga de un sistema de despacho digital y de una fuente confiable de eventos durante el traslado; ambos componentes requieren validación antes de implementar el piloto.

Departamentos y responsabilidades

La Tabla 7 detalla la responsabilidad de cada área y la información o decisión que le corresponde.

Tabla 44

Departamentos y responsabilidades en el análisis en tiempo real

Área	Responsabilidad en el caso de uso	Información / decisión
Ventas y atención al cliente	Registrar correctamente el pedido, la dirección, el contacto y el compromiso de entrega; comunicar el enlace o código de seguimiento	Pedido confirmado, consulta de estado y gestión de reclamos

Bodega y despacho	Actualizar los hitos de preparación, salida y entrega al transportista; reportar faltantes o incidencias	Pedido preparado, despachado o retenido; hora de cada evento
Logística y transporte	Actualizar el estado y la ubicación cuando exista tecnología disponible; detectar demoras y coordinar acciones correctivas	En ruta, retraso, intento fallido, entregado y hora estimada
Servicio al cliente / CRM	Registrar consultas, quejas y satisfacción posterior a la entrega; cerrar incidentes	Tiempo de respuesta, motivo de reclamo y evaluación del servicio
Gerencia de operaciones	Supervisar la puntualidad, las incidencias y la calidad del servicio; priorizar medidas de mejora	Indicadores agregados y alertas relevantes
Responsable de BI / TI (función propuesta)	Integrar eventos, validar calidad y seguridad, y administrar alertas y visualizaciones	Integridad de eventos, disponibilidad, permisos y trazabilidad

Nota. BI = Business Intelligence; TI = tecnologías de la información; CRM = gestión de la relación con los clientes. Elaboración propia.

Flujo de información y respuesta operativa

El flujo de información y respuesta operativa comprende las siguientes etapas:

1. Pedido registrado en SAGA.
2. Preparación y despacho.
3. Generación de eventos de estado y ubicación.
4. Validación e integración de datos.
5. Regla de alerta por retraso o incidencia.

6. Aviso al responsable de logística.
7. Actualización al cliente.
8. Confirmación de entrega y encuesta de satisfacción.

La alerta debe originarse en el sistema operativo de seguimiento; Power BI puede apoyar el monitoreo y el análisis de los indicadores, pero no debe considerarse el único mecanismo de notificación inmediata. Para una primera fase resulta más realista implementar el seguimiento por estados e hitos; la geolocalización continua solo se incorporaría si existe conectividad, autorización y un dispositivo o aplicación GPS adecuados. Un código QR por sí solo permite consultar información, pero no genera ubicación en tiempo real.

Indicadores vinculados directamente con Real Time

La Tabla 8 presenta los indicadores propuestos, su forma de medición y su uso en la gestión.

Tabla 45

Indicadores vinculados con el análisis en tiempo real

Indicador	Cómo medirlo	Uso de gestión
Entregas puntuales (%)	Entregas dentro del plazo / total de entregas × 100	Detectar retrasos por ruta, sucursal o transportista
Tiempo de detección de incidencias	Minutos entre el evento y la alerta registrada	Comprobar la oportunidad del sistema
Tiempo de respuesta	Minutos entre la alerta y la primera acción documentada	Evaluar la capacidad de reacción

Pedidos con trazabilidad (%)	Pedidos con hitos completos / pedidos despachados × 100	Controlar la cobertura y la calidad de los datos
Consultas por estado del pedido	Número de consultas / número de entregas	Medir la reducción de la atención manual
Satisfacción posentrega	Promedio de encuestas o porcentaje de valoraciones positivas	Relacionar el servicio logístico con la experiencia

Nota. Los tiempos se expresan en minutos. Elaboración propia.

Condiciones para una implementación viable

Antes del piloto se requiere identificar quién registra cada hito, definir un identificador único de pedido, comprobar la disponibilidad de datos logísticos y de CRM, establecer reglas de alerta, controlar los accesos a la información de los clientes y obtener el consentimiento o la base legal aplicable para tratar datos de ubicación. Se pretende iniciar con una sucursal o una ruta de las operaciones de Cañar y El Tambo, medir los resultados y escalar únicamente cuando se compruebe la calidad de los datos y la utilidad operativa.

Identificación de la mejora esperada.

La implementación de un modelo de análisis de datos en tiempo real permitirá mejorar principalmente la experiencia y la satisfacción del cliente, ya que la empresa podrá identificar de manera inmediata el estado de los pedidos, los posibles retrasos y las opiniones de los consumidores. Esto facilitará que los responsables de las áreas comerciales y de logística actúen oportunamente ante cualquier inconveniente, en lugar de esperar a que el cliente presente un reclamo. La información actualizada también permitirá ofrecer una

atención más personalizada, considerando el historial de compras, las preferencias y las experiencias anteriores del cliente.

De acuerdo con Holmlund et al. (2020), el análisis de Big Data puede generar información útil para gestionar la experiencia del cliente y mejorar la toma de decisiones en los diferentes puntos de contacto con la empresa. Asimismo, Agag et al. (2024) encontraron que la analítica de marketing contribuye a la satisfacción del cliente mediante una mayor agilidad para responder a sus necesidades.

En este caso, se espera como resultado una reducción del tiempo de respuesta ante reclamos, una disminución de los problemas relacionados con entregas tardías, una mayor capacidad de seguimiento de los pedidos y un incremento de la satisfacción y la fidelización de los clientes. La información obtenida en tiempo real también permitirá detectar patrones de comportamiento y problemas recurrentes, lo que facilitará que la empresa adopte acciones preventivas y no solo correctivas.

Cuantificar la mejora en el área de aplicación a través de un Business Case

Para cuantificar el impacto de la implementación del análisis de datos en tiempo real en CVP, se propone desarrollar un Business Case enfocado principalmente en el proceso de despacho, el seguimiento de pedidos y la satisfacción del cliente.

La propuesta consiste en incorporar un sistema de rastreo para los pedidos despachados, asociado a un código que permita al cliente consultar desde cualquier dispositivo el estado de su pedido y, cuando la tecnología lo permita, su ubicación. La información generada se integraría en una plataforma de datos conectada con Power BI, lo que permitiría a la empresa

supervisar las entregas, identificar retrasos y responder oportunamente a posibles incidencias. Esta propuesta responde directamente al objetivo del proyecto de personalizar la experiencia del cliente y conocer oportunamente si los productos llegaron a su destino.

Flujo propuesto del Business Case

La Tabla 9 presenta, en orden, las etapas del flujo propuesto.

Tabla 46

Flujo propuesto del Business Case

N.º	Etapa	Descripción
1	Pedido del cliente	Compra registrada en el ERP
2	Despacho	Pedido preparado y asignado para entrega
3	Rastreo en tiempo real	Estados e hitos del pedido y, cuando exista tecnología, ubicación asociada al pedido
4	Código de seguimiento	El cliente consulta el estado de su pedido, y su ubicación si está disponible, desde su dispositivo
5	Plataforma de datos	Integra ERP, logística, CRM y rastreo
6	Power BI en tiempo real	Monitorea entregas, tiempos y retrasos, y apoya el análisis de las alertas
7	Respuesta inmediata	CVP identifica retrasos y gestiona incidencias
8	Experiencia del cliente	Encuesta de satisfacción posterior a la entrega
9	Medición de resultados	Puntualidad, reclamos, satisfacción y tiempos de respuesta

10	Impacto del Business Case	Menores costos, mayor eficiencia y mayor satisfacción y fidelización

Nota. ERP = sistema de planificación de recursos empresariales; CRM = gestión de la relación con los clientes. Elaboración propia.

El Business Case buscará demostrar que el valor del análisis en tiempo real no radica únicamente en conocer dónde está un pedido, sino en convertir esa información en acciones inmediatas que reduzcan incidencias, mejoren la eficiencia logística y eleven la satisfacción del cliente. La mejora se cuantificará mediante indicadores antes y después de la implementación (Tabla 8), lo que permitirá determinar tanto el impacto operativo como el retorno económico de la solución.

CONCLUSIONES Y APLICACIONES

1. Respuesta al problema y cumplimiento del objetivo general

El diagnóstico evidenció que CVP incrementó el volumen de información generada por sus ventas, clientes, inventarios y compras, pero depende del ERP SAGA y de hojas de cálculo de Excel, carece de un Data Warehouse, mantiene procesos de ingreso manual de datos y no dispone de dashboards ni de capacidad de análisis histórico y predictivo. Frente a esta situación, el proyecto planteó una solución integral que comprende un Data Warehouse con Data Marts especializados, un proceso de extracción, transformación y carga (ETL), 55 indicadores clave de desempeño, siete dashboards diferenciados por perfil de usuario y un Departamento de BI con dependencia directa de la Gerencia General. En coherencia con la concepción de la inteligencia de negocios como el conjunto de técnicas, sistemas y prácticas que permiten recopilar, integrar y analizar datos para apoyar decisiones oportunas (Chen et al., 2012), la propuesta cumple el objetivo general de integrar la información comercial de la empresa para fortalecer la toma de decisiones estratégicas y contribuir al incremento de las ventas.

2. Viabilidad técnica y económica

La implementación resulta viable siempre que se ejecute por fases y se sustente en una adecuada gobernanza de los datos. El equipo de BI representa un costo de USD 60,500, correspondiente a USD 52,500 de sueldos anuales y USD 8,000 de implementación, al que se suma el licenciamiento de la plataforma de visualización. La adopción de un modelo de computación en la nube de tipo PaaS sobre Microsoft Azure, complementado con Power BI,

permite que CVP no asuma la administración de la infraestructura tecnológica, ajena a su actividad principal. La priorización de los indicadores en tres fases ordena la inversión y, de acuerdo con Turban et al. (2021), la gobernanza de datos y una estructura organizacional claramente definida son factores clave para que estas iniciativas generen valor; en este sentido, la meta de exactitud de los datos igual o superior al 98 % constituye una condición de confiabilidad de los indicadores. Los beneficios cuantificados corresponden a escenarios supuestos: un incremento de ventas del 3 % (USD 76,598.73 de facturación adicional), un margen bruto incremental de USD 28,724.52 anuales y una liberación de capital de trabajo de USD 87,534.66 al reducir el inventario de 80 a 60 días. Por tanto, requieren validación mediante un piloto antes de considerarse resultados.

3. Evolución hacia la analítica predictiva y en tiempo real

El proyecto establece una ruta de evolución desde la analítica descriptiva hacia la predictiva y la de tiempo real. Una vez integrados y depurados los datos del ERP SAGA, se propone incorporar Big Data y Machine Learning como fase posterior, con dos casos de uso priorizados: el pronóstico de la demanda por producto o categoría y la segmentación de clientes según frecuencia y valor monetario. De manera complementaria, el análisis en tiempo real permitiría conocer la satisfacción del cliente al momento de la compra y el estado de las entregas. Con ello, CVP pasaría de reportar hechos ya ocurridos a anticipar la demanda, optimizar su inventario y personalizar la experiencia del cliente, siempre que cada etapa se valide de forma gradual y se mida su aporte frente a una referencia comparable.

REFERENCIAS

Armbrust, M., Ghodsi, A., Xin, R., & Zaharia, M. (2021). Lakehouse: A new generation of open platforms that unify data warehousing and advanced analytics. *Proceedings of the Conference on Innovative Data Systems Research (CIDR)*, 1-12.

Cámara de Comercio e Industria del Cañar. (2025). *Informe de desempeño del sector de la construcción y comercio al por mayor de materiales de ferretería*. Provincia del Cañar.

Centro Constructor Vásquez Pacheco. (2024). *Manual de procesos operativos y sistema ERP SAGA*. CVP.

Centro Constructor Vásquez Pacheco. (2025). *Información corporativa*. http://www.cvpcentroconstructor.com/

Centro Constructor Vásquez Pacheco. (2026a). *Documentación técnica, manual de procesos transaccionales del ERP SAGA e indicadores estratégicos de gestión*. Departamento de Sistemas y Gestión Operativa, CVP.

Centro Constructor Vásquez Pacheco. (2026b). *Proyecto de Business Intelligence: UIDE* [Documento académico no publicado].

Chambers, B., & Zaharia, M. (2018). *Spark: The definitive guide: Big data processing made simple*. O'Reilly Media.

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.

CloudAPPI. (2025). *Mejores proveedores cloud*. https://cloudappi.net/mejores-proveedores-cloud/

Conecta Software. (s. f.). Business Intelligence por departamentos. *Revista Conecta Software*. https://www.conectasoftware.com/magazine/bi-por-departamentos/

Contabilidad Finanzas. (s. f.). *Contabilidad por departamentos: Guía completa*. https://contabilidadfinanzas.com/contabilidad-por-departamentos/

Curto Díaz, J., & Conesa Caralt, J. (2010). *Introducción al Business Intelligence*. Editorial UOC.

Curto Díaz, J., & Conesa Caralt, J. (2015). *¿Cómo crear un Data Warehouse?* Editorial UOC.

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

Ehrenmueller-Jensen, M. (2024). *Data modeling with Microsoft Power BI: Self-service and enterprise data warehouse with Power BI*. Packt Publishing.

Espinosa Montiel, C. A. (2013). *Guía para implementar una solución Business Intelligence: Caso de estudio Empresa Espinosa & Espinoza* [Trabajo de titulación, Pontificia Universidad Católica del Ecuador].

Fuller, P. D. (2024). *Beginning Power BI for business users: Learning to turn data into insights*. John Wiley & Sons.

Garg, N. (2018). *Apache Kafka quick start guide: Leverage Apache Kafka to build real-time data pipeline models*. Packt Publishing.

Golfarelli, M., & Rizzi, S. (2009). *Data warehouse design: Modern principles and methodologies*. McGraw-Hill Education.

Google Cloud. (s. f.). *¿Qué es ETL?* https://cloud.google.com/learn/what-is-etl?hl=es-419

IBM. (s. f.). *¿Qué es ETL?* https://www.ibm.com/mx-es/think/topics/etl

Inmon, W. H. (2005). *Building the data warehouse* (4.ª ed.). Wiley Publishing.

Joyanes Aguilar, L. (2019). *Inteligencia de negocios y analítica de datos: Una visión global de Business Intelligence & Analytics*. Alfaomega Colombiana.

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

Laudon, K. C., & Laudon, J. P. (2022). *Sistemas de información gerencial: Gestión de la empresa digital* (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. (2023). *SQL Server Integration Services (SSIS) documentation*. Microsoft Learn. https://learn.microsoft.com/es-es/sql/integration-services/

Microsoft. (2024). *SQL Server Integration Services (SSIS)*. Microsoft Learn. https://learn.microsoft.com/sql/integration-services

Microsoft. (2025). *Power BI documentation*. Microsoft Learn.

Microsoft. (2026a). *Azure y Power BI*. Microsoft Learn.

Microsoft. (2026b). *Introducción a Azure Data Factory*. Microsoft Learn.

Microsoft. (2026c). *Preguntas más frecuentes sobre Azure Data Factory*. Microsoft Learn.

Microsoft. (2026d). *¿Qué es Azure Synapse Analytics?* Microsoft Learn.

OKUNDATA. (2026). *Costo de business intelligence*. https://okundata.com/novedades/costo-de-business-intelligence

Oracle. (2025). *Oracle Business Intelligence Applications documentation*.

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

SAGA ERP. (s. f.). *Pricing*. https://sagaerp.com/pricing

Sharda, R., Delen, D., & Turban, E. (2014). *Business intelligence and analytics: Systems for decision support* (10.ª ed.). Pearson.

Sharda, R., Delen, D., & Turban, E. (2024). *Business intelligence, analytics, data science, and AI* (5.ª ed.). Pearson.

Turban, E., Sharda, R., Delen, D., & King, D. (2021). *Inteligencia empresarial, análisis y ciencia de datos: Una perspectiva gerencial* (5.ª ed.). Pearson.