

**Maestría en**

## **ADMINISTRACIÓN DE EMPRESAS**

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

### **AUTORES:**

Sigchos Franklin  
Rodríguez Daniela  
Ramos Rodrigo  
Sánchez David

### **TUTORES:**

Pérez Galán José Luis  
Maroto Ignacio Mateo

Análisis estacional del comportamiento de compra en  
una empresa de acabados de construcción.

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

**Tesis previa a la obtención de título de Magíster en  
Administración de Empresas**

Por

Sigchos Franklin  
Rodríguez Daniela  
Ramos Rodrigo  
Sánchez David

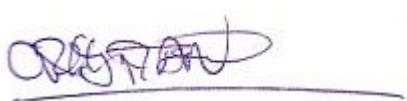
Octubre 2025

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: Sigchos Morales Franklin, Rodríguez Rodríguez Daniela, Ramos Averos Rodrigo, Sánchez Meza David 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 Westfield Business School

### **Autoría del Trabajo de Titulación**

Nosotros, Sigchos Morales Franklin, Rodríguez Rodríguez Daniela, Ramos Averos Rodrigo, Sánchez Meza David, declaramos bajo juramento que el trabajo de titulación titulado **“Análisis estacional del comportamiento de compra en una empresa de acabados de construcción”** 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.



**Franklin Herbert Sigchos Morales**

c. e. [franklin\\_sigchos@hotmail.com](mailto:franklin_sigchos@hotmail.com)



**Daniela Alexandra Rodríguez Rodríguez**

c. e. [danyrodriguezr3101@gmail.com](mailto:danyrodriguezr3101@gmail.com)



**David Alejandro Sánchez Meza**

c.e. [david.sanchezmeza@outlook.com](mailto:david.sanchezmeza@outlook.com)



**Rodrigo Lenin Ramos Averos**

c.e. [ramoslenin32@gmail.com](mailto:ramoslenin32@gmail.com)

## Autorización de Derechos de Propiedad Intelectual

Nosotros, Sigchos Morales Franklin, Rodríguez Rodríguez Daniela, Ramos Averos Rodrigo, Sánchez Meza David, en calidad de autores del trabajo de investigación titulado **“Análisis estacional del comportamiento de compra en una empresa de acabados de construcción”**, autorizamos a la Universidad Internacional del Ecuador (UIDE) para hacer uso de todos los contenidos que nos pertenecen o de parte de los que contiene esta obra, con fines estrictamente académicos o de investigación. Los derechos que como autor me corresponden, lo establecido en los artículos 5, 6, 8, 19 y demás pertinentes de la Ley de Propiedad Intelectual y su Reglamento en Ecuador.

D. M. Quito, octubre 2025



**Franklin Herbert Sigchos Morales**

c. e. [franklin\\_sigchos@hotmail.com](mailto:franklin_sigchos@hotmail.com)



**Daniela Alexandra Rodríguez Rodríguez**

c. e. [danyrodriguezr3101@gmail.com](mailto:danyrodriguezr3101@gmail.com)



**David Alejandro Sánchez Meza**

c.e. [david.sanchezmeza@outlook.com](mailto:david.sanchezmeza@outlook.com)



**Rodrigo Lenin Ramos Averos**

c.e. [ramoslenin32@gmail.com](mailto:ramoslenin32@gmail.com)

## Dedicatorias

A mi familia, por ser el pilar fundamental en cada paso de mi vida. Por su amor

incondicional, su paciencia infinita y su constante apoyo en los momentos más exigentes de este camino académico y personal.

Dedico este logro a quienes me enseñaron el valor del esfuerzo, la disciplina y la perseverancia.

Y también me lo dedico a mí mismo, por nunca rendirme, por creer en mis capacidades y por seguir cada día en la búsqueda de ser una mejor versión de mí.

**Franklin Sigchos Morales**

---

A mi padre, a mi madre y a mi hermano: con ustedes, el mundo se ilumina y todo se hace posible; sin ustedes, el silencio sería infinito.

**Rodrigo Lenin Ramos Averos**

---

A mi esposa, por su amor incondicional, su comprensión en los momentos de mayor esfuerzo y su apoyo constante a lo largo de este camino.

Su presencia ha sido un Pilar fundamental en cada etapa que me impulsaron a superar los desafíos con determinación. Gracias por tu presencia y tu fe inquebrantable en cada paso de mi vida.

**David Alejandro Sánchez Meza**

---

A mi familia y seres queridos, por ser mi mayor fuente de inspiración y fortaleza. A mis padres, por su amor incondicional, por enseñarme el valor del esfuerzo y la perseverancia. A mis hermanos, por su apoyo constante, su alegría y por estar siempre a mi lado en cada etapa de este proceso. Y a mi fiel compañero de cuatro patas “Tévez”, que con su presencia silenciosa y leal me acompañó en cada clase, recordándome que el cariño también se expresa sin palabras.

**Daniela Alexandra Rodríguez Rodríguez**

## Agradecimientos

Expreso mi más sincero agradecimiento a mi familia, por su respaldo inquebrantable, su comprensión y su motivación constante durante todo el proceso de este máster.

A mis compañeros de estudio, por su colaboración, amistad y las experiencias compartidas que enriquecieron este camino académico.

A mis profesores y tutores, por su guía, su conocimiento y su compromiso con nuestra formación profesional, especialmente a quienes acompañaron y orientaron este proyecto de titulación con dedicación y exigencia.

Y finalmente, agradezco a todos quienes, de una u otra manera, formaron parte de esta etapa y contribuyeron al logro de esta meta tan significativa en mi vida.  
**Franklin Sigchos Morales**

---

A Dios, por el regalo inmenso de la salud y la vida, que me sostienen en cada paso del camino con su luz eterna.

**Rodrigo Lenin Ramos Averos**

---

A Dios, por concederme la salud, la fortaleza y la claridad necesarias para culminar este proceso.

**David Alejandro Sánchez Meza**

---

A Dios, por darme la fuerza, la sabiduría y la salud necesarias para culminar esta etapa.

Por acompañarme en cada paso y permitirme alcanzar este importante logro.

**Daniela Alexandra Rodríguez Rodríguez**

# INDICE DEL DOCUMENTO

INTRODUCCIÓN Y OBJETO DEL PROYECTO .....	12
Definición del Proyecto .....	12
Nombre del proyecto .....	12
Objetivos del proyecto .....	12
Objetivo general.....	12
Objetivos específicos .....	12
Definición de la empresa o negocio.....	13
Nombre de la empresa .....	13
Misión y visión de la empresa.....	13
Misión.....	13
Visión.....	13
Procesos, organización y cifras del negocio .....	14
Procesos .....	14
Organización.....	15
Cifras del negocio.....	15
Vinculación con la empresa, fuentes de información y nivel de conocimiento .....	16
Otra información .....	16
PARTE I.....	18
CAPITULO 1: KPI.....	18
1.1.    Objetivo 1 .....	18
1.2.    Objetivo 2 .....	19
1.3.    Objetivo 3 .....	19
CAPITULO 2: FUENTES DE INFORMACIÓN.....	21
2.1.    Origen de la información.....	21
2.2.    Recurrencia de carga .....	21
2.3.    Sistemas y destinos de la carga .....	21
2.4.    Carencias y problemas en la carga de datos e informes actuales.	22
2.5.    Informes que se generan actualmente. ....	22
2.6.    Necesidades detectadas y no cubiertas. ....	22
CAPITULO 3: PROCESOS ETL .....	23
3.1.    Procesos de carga de datos.....	23
3.2.    Información y calidad de la información disponible. ....	24

3.3.	Horarios de carga y accesibilidad .....	25
3.4.	Variables disponibles y acceso al Data Warehouse.....	25
3.5.	Sistemas implicados en los procesos.....	27
3.6.	Recursos de IT o externos.....	27
<b>CAPITULO 4: HERRAMIENTAS BUSINESS INTELLIGENCE .....</b>		<b>28</b>
4.1.	Análisis de herramientas y selección más adecuada para el ecosistema.....	28
4.2.	Selección de usuarios y perfiles de uso de cada dashboard.....	29
4.3.	Necesidades de información y accesos por departamento y perfil.	30
4.4.	Revisión y cumplimiento de los objetivos iniciales .....	31
<b>CAPITULO 5: CREACIÓN DEL ÁREA DE BI.....</b>		<b>31</b>
5.1.	Situación actual del área de BI.....	31
5.2.	Definición de las carencias en formación y recursos .....	32
5.3.	Coordinación de los flujos entre otras áreas .....	33
5.4.	Definir el organigrama, dependencias de otras áreas y solución de problemas .....	34
5.5.	Recursos externos destinados: subcontratación.....	35
5.6.	Business case de costes y beneficios. ....	36
<b>PARTE II: .....</b>		<b>39</b>
<b>CAPITULO 1: VALOR AGREGADO DE LOS DATOS .....</b>		<b>39</b>
1.1.	Fuentes de datos externas e internas, estructuradas o semiestructuradas.....	39
1.2.	Relación de los datos para resolver y mejorar propuesta inicial....	40
1.3.	Casos de uso para el modelo de BI.....	41
1.4.	Justificación del impacto y aporte del Big Data. ....	44
<b>CAPITULO 2: SELECCIÓN DE PROVEEDOR Y ARQUITECTURA BI - CERÁMICAS OCCIDENTAL .....</b>		<b>46</b>
2.1.	Análisis de Alternativas de Proveedores.....	46
2.2.	Selección del Proveedor más Adecuado .....	48
2.3.	Modelo de Cloud Computing Seleccionado .....	48
2.4.	Flujo de Integración entre Big Data y Data Warehouse.....	49
<b>CAPITULO 3: TECNOLOGÍAS BIG DATA.....</b>		<b>50</b>
3.1.	Análisis Frameworks .....	50
3.2.	Elementos adicionales.....	52

<b>CAPITULO 4: APLICACIÓN DE DATA SCIENCE Y MACHINE LEARNING</b>	53
4.1. Identificación y definición del objetivo de aplicación del modelo.....	53
4.2. Fuentes de datos utilizadas .....	54
4.3. Área de aplicación y departamentos de negocio implicados.....	55
4.4. Identificación de la mejora esperada.....	55
4.5. Cuantificación de la mejora a través de un Business Case.....	56
Inversión y ROI.....	57
<b>CAPITULO 5: MODELO DE APLICACIÓN EN TIEMPO REAL (REAL TIME DATA INTELLIGENCE)</b> .....	59
5.1. Identificación y definición del objetivo de aplicación del modelo Real Time .....	59
5.2. Fuentes de datos utilizadas.....	59
5.3. Área de aplicación y departamentos de negocio implicados. ....	59
5.4. Identificación de la mejora esperada. ....	60
5.5. Cuantificación de la mejora – Business Case Real Time.....	61
<b>CONCLUSIONES Y APLICACIONES</b> .....	61
Aplicaciones en un entorno concreto.....	62

# INDICE DE FIGURAS Y TABLAS

Figura 1. Proceso ETL.....	24
Figura 2. Situación actual del departamento BI .....	32
Figura 3. Carencias en formación y recursos.....	33
Figura 4. Organigrama Empresa.....	34
Figura 5. Costes y beneficios estimados.....	36
Figura 6. Flujo de Big Data.....	49
Tabla 1. Fuentes de información y nivel de conocimiento.....	16
Tabla 2. KPI Objetivo 1.....	18
Tabla 3. KPI Objetivo 2.....	19
Tabla 4. KPI Objetivo 3.....	19
Tabla 5. Herramientas Business Intelligence.....	28
Tabla 6. Perfiles de usuario y necesidades .....	29
Tabla 7. Necesidades de información.....	30
Tabla 8. Costes estimados.....	36
Tabla 9. Fuente de Datos Internas.....	39
Tabla 10. Fuentes de Datos Externas .....	39
Tabla 11. Clasificación por tipo de estructura.....	40
Tabla 12. Datos internos .....	42
Tabla 13. Proveedores.....	46
Tabla 14. Herramientas tecnológicas .....	50
Tabla 15. Supuestos técnicos (prudentes y defendibles).....	56
Tabla 16. Fuentes utilizadas.....	59
Tabla 17. Mejoras esperadas .....	60
Tabla 18. Cuantificación de la mejora .....	61

# INTRODUCCIÓN Y OBJETO DEL PROYECTO

## Definición del Proyecto

El presente proyecto consiste en el desarrollo de un sistema de Business Intelligence (BI) orientado al análisis estacional del comportamiento de compra de los clientes en una empresa ecuatoriana dedicada a la comercialización e importación de acabados de construcción.

A través del tratamiento y visualización de datos históricos de ventas, se busca identificar patrones de demanda que se repiten periódicamente (mensual o estacionalmente) por categoría de producto, showroom y región. Esto permitirá a la empresa tomar decisiones informadas en relación con la planificación de inventarios, promociones comerciales y compras a proveedores internacionales, reduciendo la dependencia de la intuición o la experiencia individual.

El sistema BI será desarrollado utilizando herramientas como Excel y Power BI, y entregará dashboards interactivos que faciliten el acceso y la interpretación de la información clave para los responsables comerciales, logísticos y directivos de la organización.

## Nombre del proyecto

Análisis Estacional del Comportamiento de Compra en una Empresa de Acabados de Construcción.

## Objetivos del proyecto

### Objetivo general:

Desarrollar un sistema de Business Intelligence para analizar el comportamiento estacional de compra de los clientes en una empresa de acabados de construcción.

### Objetivos específicos:

Identificar patrones de preferencia y temporalidad en la compra de productos por parte de los clientes, mediante el análisis estacional de ventas por categoría.

Clasificar los productos según su nivel de sensibilidad estacional, como herramienta para segmentar tipos de comportamiento de compra del cliente.

Generar visualizaciones interactivas que apoyen la toma de decisiones.

## **Definición de la empresa o negocio**

La empresa objeto de este estudio es una organización ecuatoriana con 13 años de trayectoria, dedicada a la comercialización e importación de acabados de construcción como porcelanato, grifería, sanitarios, piedra sinterizada y mobiliario para baños y cocinas. Su enfoque principal es ofrecer productos de alta calidad, funcionales y estéticamente atractivos, que satisfagan las necesidades de clientes particulares, arquitectos, constructores y diseñadores.

Cuenta con una estructura organizacional funcional compuesta por 20 colaboradores y 4 directivos, y opera con presencia comercial en las regiones Costa y Sierra del país, a través de sus showrooms. Su propuesta de valor se basa en la asesoría personalizada, variedad de productos, estrategias con proveedores internacionales, y una atención centrada en la experiencia del cliente.

La empresa mantiene una dinámica operativa que combina procesos de venta directa, gestión de inventario, importación especializada y planificación comercial, lo que la convierte en un actor competitivo dentro del mercado nacional de acabados para la construcción.

## **Nombre de la empresa**

“Cerámicas Occidental”

## **Misión y visión de la empresa**

### **Misión**

Brindar soluciones integrales en acabados de construcción a través de la comercialización e importación de productos de alta calidad, innovadores y funcionales, que satisfagan las necesidades estéticas y técnicas de nuestros clientes, garantizando un servicio confiable, asesoría personalizada y una experiencia de compra diferenciada.

### **Visión**

Ser una empresa reconocida a nivel nacional en el sector de acabados de construcción, por su excelencia en servicio, portafolio de productos vanguardistas, alianzas estratégicas internacionales y compromiso con la innovación, sostenibilidad y mejora continua.

## **Procesos, organización y cifras del negocio**

### **Procesos**

#### **Proceso de ventas**

Es el núcleo del análisis. Abarca la recepción del cliente, cotización, cierre de venta y facturación.

Proporciona los datos necesarios para identificar patrones de compra estacionales por tipo de producto, showroom y periodo del año.

#### **Gestión de inventario**

Involucra el control, almacenamiento y reposición de productos.

Su coordinación con la demanda es vital para evitar sobrestock en temporadas bajas y quiebres en temporadas altas.

#### **Análisis y planificación comercial**

Consiste en la revisión de indicadores de desempeño, históricos de ventas y tendencias del mercado.

Este proceso será fortalecido con el sistema BI propuesto, al brindar visualizaciones y predicciones basadas en datos reales.

#### **Importación y abastecimiento**

Controla los pedidos a proveedores internacionales según la demanda esperada.

Un análisis estacional permite planificar con mayor precisión qué productos importar y en qué cantidad, según la temporada.

#### **Gestión de marketing y promociones**

Define campañas publicitarias, descuentos y exhibiciones en showroom.

La detección de productos de alta demanda en ciertos meses permitirá enfocar mejor las promociones y maximizar su impacto.

## Organización

La empresa se encuentra organizada de forma funcional, con áreas claramente definidas que trabajan de manera coordinada para cumplir con su misión de ofrecer productos de acabados de construcción de alta calidad y un servicio confiable. Entre las áreas principales destacan:

- 1. Gerencia General:** Supervisa las operaciones y toma decisiones estratégicas.
- 2. Área Comercial:** Encargada de la atención al cliente, ventas en showroom y cierre de negocios.
- 3. Área de Compras e Importaciones:** Se encarga de abastecer los productos desde proveedores internacionales en función de la demanda prevista.
- 4. Área de Logística e Inventarios:** Administra el almacenamiento, disponibilidad y traslado de productos entre locales.
- 5. Área Administrativa y Financiera:** Controla recursos, pagos, facturación y reportes de gestión.
- 6. Área de Marketing:** Desarrolla campañas y promociones para impulsar las ventas.

A pesar de esta estructura formal, uno de los problemas clave identificados es que la toma de decisiones relacionadas con inventarios, promociones y abastecimiento aún se basa principalmente en la intuición y la experiencia previa del personal, en lugar de un análisis sistematizado de datos históricos. Esta práctica ha llevado a situaciones de sobrestock en temporadas bajas, quiebres de inventario en momentos de alta demanda y pérdida de oportunidades comerciales.

Este proyecto busca precisamente fortalecer los procesos de análisis y planificación comercial, permitiendo que las decisiones se tomen con base en patrones reales y medibles del comportamiento del cliente. Para lograrlo, el sistema de Business Intelligence propuesto entregará información visual, clara y estructurada sobre las tendencias estacionales de ventas por producto, showroom y mes.

## Cifras del negocio

La empresa cuenta con 13 años de trayectoria en el mercado ecuatoriano, consolidándose como un actor relevante en la comercialización e importación de acabados de construcción. Actualmente opera con un equipo de 20 colaboradores, distribuidos en 3 áreas comerciales, logísticas, administrativas y técnicas, bajo la dirección de 4 ejecutivos que lideran la estrategia y operación del negocio.

La compañía mantiene presencia activa en dos de las principales regiones del país: la Sierra y la Costa, atendiendo tanto al consumidor final como a profesionales del sector de la construcción y diseño. Si bien su enfoque principal es el mercado nacional, ha desarrollado relaciones comerciales con proveedores internacionales para importar productos innovadores y de alta calidad, adaptados a las tendencias globales del sector.

En términos operativos, la empresa registra un promedio anual estimado de 5 millones en ventas, con un portafolio de más de 2000 productos que abarca porcelanatos, grifería, sanitarios, piedra sinterizada y mobiliario especializado.

Estas cifras reflejan una estructura ágil y en crecimiento, con un equipo comprometido y una cobertura estratégica que requiere herramientas modernas de análisis y planificación para seguir respondiendo con eficiencia al dinamismo del mercado.

## **Vinculación con la empresa, fuentes de información y nivel de conocimiento**

Uno de los miembros del equipo Franklin Sigchos es socio propietario de “Cerámicas Occidental”.

**Tabla 1.** Fuentes de información y nivel de conocimiento

<b>Fuente</b>	<b>Valoración</b>	<b>Justificación</b>
Histórico de ventas año 2024	10	Propietario de la información
Catálogo de productos	10	
Datos logísticos y de reabastecimiento	10 9	
Información del inventario (opcional):		
Tendencias del mercado y consumidor	7	Destino de la información desde fuentes externas (proveedores, ferias y otros)
Información básica organizacional:	5	estructura, misión, visión, procesos. (si las hubo).
Información de campañas promocionales anteriores		

## **Otra información**

Cerámicas Occidental tiene 13 años de trayectoria como persona natural, a partir del año 2025 cambio para ser una S.A por lo que se considera implementar procesos más formales en sus diferentes actividades.

Para el presente proyecto se tomará en cuenta la data a partir del año 2024 debido a su disponibilidad. Ya que se realizó un cambio de servidores en ese año.

# **PARTE I:**

## **CAPITULO 1: KPI**

### **1.1. Objetivo 1**

Identificar patrones de comportamiento de compra mensual y estacional por tipo de producto.

**Tabla 2.** KPI Objetivo 1

<b>KPI</b>	<b>Fórmula</b>	<b>Meta inicial</b>	<b>Frecuencia</b>	<b>Fuente de datos</b>
% de categorías de producto con patrones de compra claramente estacionales detectados	(Número de categorias con patrón estacional identificado / Total de categorias analizadas) × 100	≥ 70% de productos categorizados.	Mensual	Histórico de ventas, categorización Power BI

#### **1.1.1. ¿Por qué es útil?**

Permite enfocar campañas y abastecimiento en función de cómo compran los clientes según la temporada.

#### **1.1.2. ¿Cómo mejora la productividad?**

Facilita priorizar el análisis de categorías que sí presentan patrones, evitando esfuerzos innecesarios en productos con demanda plana.

#### **1.1.3. ¿Cómo impacta en la rentabilidad?**

Permite anticipar la demanda y planificar compras más eficientes, reduciendo quiebres de stock y costos de almacenamiento innecesario.

## 1.2. Objetivo 2

Clasificar productos con alta sensibilidad estacional.

**Tabla 3. KPI Objetivo 2**

KPI	Fórmula	Meta inicial	Frecuencia	Fuente de datos
Índice de sensibilidad estacional por categoría de producto	Variación de ventas estacional promedio (mes pico vs mes mínimo) / Ventas con promedio anual	Tener clasificados al menos 5 categorías con sensibilidad alta (>30%)	Trimestral	Dashboard Power BI, análisis histórico

### 1.1.4. ¿Por qué es útil?

Ayuda a segmentar el comportamiento del cliente y la efectividad del cuerpo de ventas: si hay categorías con mucha sensibilidad, se puede planificar más estratégicamente su promoción, abastecimiento o impulso de venta.

### 1.1.5. ¿Cómo mejora la productividad?

Permite segmentar el portafolio según la sensibilidad estacional, lo que facilita enfoques comerciales y logísticos más precisos.

### 1.1.6. ¿Cómo impacta en la rentabilidad?

Ayuda a evitar sobre stock de productos fuera de temporada y a enfocar promociones en los momentos más rentables del año.

## 1.3. Objetivo 3

Generar visualizaciones interactivas para cada departamento que apoyen la toma de decisiones.

**Tabla 4. KPI Objetivo 3**

KPI	Fórmula	Meta inicial	Frecuencia	Fuente de datos
% de departamentos con dashboards personalizados	(Número de departamentos con dashboards funcionales / Total de departamentos clave)*100	Al menos 80% de las áreas clave con dashboards	Semestral	Dashboard Power BI

### **1.1.7. ¿Por qué es útil?**

Refleja adopción real del sistema BI para tomar decisiones centradas en el comportamiento del cliente.

### **1.1.8. ¿Cómo mejora la productividad?**

Ahorra tiempo en generación de reportes manuales, centraliza la información y mejora la comunicación entre áreas.

### **1.1.9. ¿Cómo impacta en la rentabilidad?**

Permite tomar decisiones más rápidas y con mayor precisión, especialmente en compras, promociones y asignación de recursos.

## CAPITULO 2: FUENTES DE INFORMACIÓN

### 2.1. Origen de la información

#### 2.1.1. Origen interno:

La mayor parte de la información relevante proviene del sistema de ventas y facturación de la empresa, donde se registran los datos de cada transacción: producto vendido, cantidad, precio, fecha y stock.

#### 2.1.2. Origen externo:

Índice de costos de la construcción, informes de tendencias en diseño y arquitectura, datos de precios internacionales de insumos cerámicos y porcelánicos, índices de precios de materiales, Clima, Actividad del Sector Construcción (demanda por ciclo económico), Calendario Comercial, Festivos y Eventos Regionales e Indicadores Económicos y de Consumo.

### 2.2. Recurrencia de carga

- Los datos internos (ventas, stock) se generan y almacenan diariamente.
- Sin embargo, no existe una integración automatizada para análisis, y los reportes se consolidan mensualmente de forma manual.
- La carga externa (pedidos a proveedores, disponibilidad) se gestiona por temporada o según necesidad.

### 2.3. Sistemas y destinos de la carga

#### 2.3.1. Sistemas de origen

Software comercial donde se emiten las facturas y se lleva el control básico del inventario.

#### 2.3.2. Destino de la información

- Actualmente, los datos se cargan o consolidan en archivos Excel que se revisan en reuniones internas en los diferentes departamentos involucrados. No existe un sistema unificado ni una base de datos centralizada para análisis histórico profundo ni dashboards dinámicos.
- Excel consolidado: En el cual se va a centralizar toda la información para el procesamiento del dashboard.
- Carga de los datos en Power BI para la visualización y análisis.

## 2.4. Carencias y problemas en la carga de datos e informes actuales

- **Falta de automatización:** La carga y procesamiento de datos se hace de forma manual, lo que conlleva riesgo de errores, retrasos y duplicidad de esfuerzos.
- **Limitada capacidad analítica:** Los reportes actuales son estáticos, no permiten explorar patrones o tendencias más allá de comparativos simples (mes a mes o showroom a showroom).
- **Poca trazabilidad:** No se puede reconstruir fácilmente el comportamiento histórico por producto, ni proyectar escenarios.
- **Falta de Data Warehouse/Data Marts:** Los datos están en un sistema transaccional no optimizado para análisis, lo que dificulta consultas complejas.
- **Falta de Datos Externos:** No hay integración de datos externos (clima, índices de construcción) para enriquecer el análisis de estacionalidad.
- **Latencia:** Retrasos en la actualización de datos (ej. stock no actualizado en tiempo real).

## 2.5. Informes que se generan actualmente.

- • • •
- Reporte de ventas diarias.
- Reporte de inventarios.
- Reporte de balance general
- Reporte de pérdidas y ganancias
- Reporte de ventas por sucursal

## 2.6. Necesidades detectadas y no cubiertas

- Contar con un sistema centralizado de análisis y visualización que permita comprender el comportamiento real de estacional de compra de los clientes. Poder detectar patrones de estacionalidad o ausencia de ellos, para mejorar decisiones de compras, promociones e importaciones.
- Reducir la dependencia del criterio individual o intuición para planificar el stock o lanzar campañas.
- Agilizar el acceso a la información clave por parte de la gerencia, con reportes más dinámicos e interactivos.
- Mejorar la integración entre ventas, inventario y planificación comercial, usando herramientas modernas como Power BI.

- Reporte de Stock: Listado de productos con inventario disponible, probablemente sin análisis de rotación.

## CAPITULO 3: PROCESOS ETL

### 3.1. Procesos de carga de datos

Los procesos de carga de datos se implementarán mediante un flujo ETL utilizando Pentaho Data Integration para integrar datos desde reportes CSV del sistema transaccional hacia una base de datos PostgreSQL.

#### 3.1.1. Extracción:

**Fuentes:** Reportes generados por el sistema transaccional en formatos CSV, almacenados en carpetas.

**Método:** Pentaho Data Integration se configurará para leer archivos CSV. DBeaver se utilizará para consultas SQL para verificar la información.

**Automatización:** Ejecución de un job por medio de Pentaho para migrar los datos de los archivos CSV a la base de datos PostgreSQL.

#### 3.1.2. Transformación:

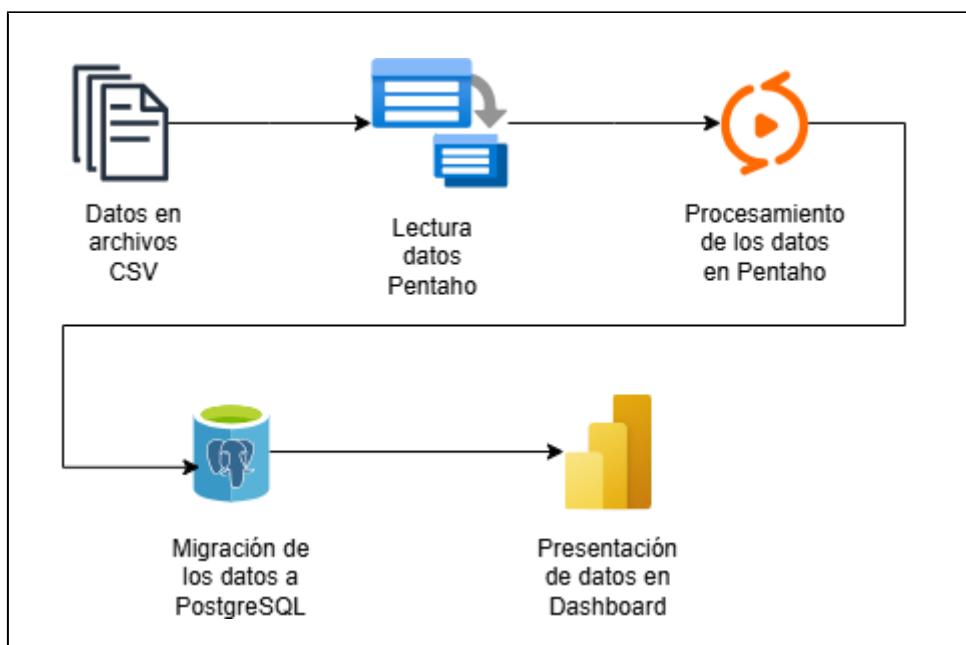
**Limpieza:** En Pentaho, se ejecutarán pasos para eliminar duplicados, estandarizar nombres y normalizar categorías.

**Validación:** Scripts SQL en DBeaver para verificar integridad y pasos en Pentaho para validar reglas de negocio.

#### 3.1.3. Carga:

**Destino:** Base de datos PostgreSQL como Data Warehouse.

**Método:** Uso de pasos en Pentaho para cargar datos en tablas PostgreSQL. Carga completa de la información a partir de los archivos CSV.



**Figura 1.** Proceso ETL

### 3.2. Información y calidad de la información disponible

La calidad de los datos es crítica para análisis fiables. Los detalles son:

#### 3.2.1. Información disponible:

- Reporte de ventas diarias.
- Reporte de inventarios.
- Reporte de balance general
- Reporte de pérdidas y ganancias
- Reporte de ventas por sucursal

#### 3.2.2. Calidad de la información:

**Compleitud:** Verificación en DBeaver con consultas SQL para identificar datos faltantes.

**Consistencia:** Uso de pasos en Pentaho para estandarizar nombres mediante diccionarios de mapeo.

**Precisión:** Validación en Pentaho con reglas del negocio.

**Actualidad:** Garantizar que los datos reflejen transacciones con un retraso máximo de 24 horas (diario) o 7 días (semanal).

**Herramientas:** Pentaho para transformaciones de calidad, DBeaver para revisión de la información.

### **3.3. Horarios de carga y accesibilidad**

#### **3.3.1. Horarios de carga:**

**Reportes de ventas e inventarios:** Carga semanal, domingos a las 02:00 AM.

**Reportes agregados:** Carga mensual, 02:00 AM último día de cada mes.

**Orquestación:** Uso de Pentaho para scheduling, con logs en PostgreSQL para monitoreo.

#### **3.3.2. Accesibilidad:**

**Usuarios:** Analistas, gerentes de ventas y directivos.

**Método de acceso:** Power BI conectado a PostgreSQL mediante un conector ODBC. DBeaver para consultas SQL ad-hoc por parte de analistas técnicos.

**Seguridad:** Autenticación mediante usuarios de PostgreSQL con permisos solo de lectura.

**Disponibilidad:** Consultas en tiempo real en Power BI, con refrescos automáticos configurados.

### **3.4. Variables disponibles y acceso al Data Warehouse**

#### **3.4.1. Variables disponibles:**

##### **Ventas:**

ID\_Transacción (integer, único).

Fecha (timestamp).

ID\_Producto (integer).

Cantidad (integer).

Precio\_Unitario (float).

Total (float).

ID\_Cliente (integer).

Región (varchar).

**Productos:**

ID\_Producto (integer, único).

Categoría (varchar).

Subcategoría (varchar).

Precio (float).

Stock (integer).

Proveedor (varchar).

**Clientes:**

ID\_Cliente (integer, único).

Tipo (varchar, minorista/mayorista).

Región (varchar).

Total\_Compras (float).

**Temporalidad:**

Fecha (timestamp).

Mes (varchar).

Trimestre (varchar).

Año (integer).

Temporada (varchar).

**Métricas derivadas:**

- Ventas totales por temporada y categoría.
- Productos más vendidos por región.
- Variación porcentual interanual.

### **3.4.2. Acceso al Data Warehouse:**

**Infraestructura:** PostgreSQL como Data Warehouse, alojado en un servidor local.

**Optimización:** Índices en columnas clave y particiones por año.

**Acceso:** Power BI para visualizaciones, conectado mediante ODBC. DBeaver para consultas SQL avanzadas.

**Seguridad:** Roles en PostgreSQL para analistas, técnicos.

## **3.5. Sistemas implicados en los procesos**

**Sistema transaccional:** Base de datos relacional que genera reportes CSV.

**Herramientas ETL:** Pentaho Data Integration para flujos ETL, con jobs y transformaciones configuradas para leer CSV y cargar en PostgreSQL.

**Data Warehouse:** PostgreSQL, configurado con esquemas optimizados para análisis por medio de tablas particionadas, e índices.

**Herramientas de BI:** Power BI para creación de dashboards y reportes interactivos.

### **Sistemas de soporte:**

- DBeaver para consultas SQL ad-hoc y revisión de la información.
- Servidor local para alojar PostgreSQL.
- Pentaho para orquestación de procesos ETL.

## **3.6. Recursos de IT o externos**

### **3.6.1. Recursos internos de IT:**

**Administrador de bases de datos:** Configura PostgreSQL, crea índices y particiones, gestiona usuarios.

**Desarrollador ETL:** Diseño de jobs y transformaciones en Pentaho, generación de scripts SQL en DBeaver.

**Analista de datos:** Crea dashboards en Power BI, define KPIs.

**Especialista en seguridad:** Configura roles y permisos en PostgreSQL.

### **3.6.2. Recursos externos:**

**Licencias de software:** Pentaho Data Integration, Power BI Pro.

**Capacitación:** Cursos en Pentaho, Power BI y PostgreSQL para el equipo interno.

## CAPITULO 4: HERRAMIENTAS BUSINESS INTELLIGENCE

### 4.1. Análisis de herramientas y selección más adecuada para el ecosistema

#### 4.1.1. Herramientas analizadas

- **Pentaho Data Integration:** Herramienta ETL para extracción, transformación y carga. Código abierto, flexible y escalable.
- **Power BI:** Visualización e inteligencia de negocios, ideal para usuarios no técnicos. Gran compatibilidad con PostgreSQL.
- **DBeaver:** Cliente SQL para exploración y validación de bases de datos.
- **PostgreSQL:** Base de datos robusta para modelado relacional, excelente como Data Warehouse.
- **Excel:** Aún es usado como medio de consolidación manual temporal.

#### 4.1.2. Selección recomendada por rol

**Tabla 5.** Herramientas Business Intelligence

Herramienta recomendada	
ETL Pentaho	Automatiza cargas desde CSV y limpia datos.
Base de datos: PostgreSQL	Escalable y óptimo para consultas analíticas.
DBeaver	Ideal para pruebas y debugging de datos.
Dashboards	Power BI – Interactivo, dinámico y fácil de usar por usuarios no técnicos.
Consolidación temporal	Solo como apoyo temporal antes de migrar 100% al DWH.

La combinación Pentaho + PostgreSQL + Power BI es ideal para el ecosistema actual. Escalable, de bajo costo y alineada a las competencias del equipo.

#### Justificación:

Cerámicas Occidental opera con un sistema ERP (LATINIUM) que no está diseñado para análisis avanzado de datos, lo que genera limitaciones al momento de cruzar información de ventas, stock y compras de los tres

puntos de venta. Ante esto, se requiere una herramienta ágil, accesible y escalable que permita conectar fuentes internas (ERP, Excel, redes sociales) y generar dashboards en tiempo real.

La elección de Power BI responde a la necesidad de contar con una herramienta intuitiva para usuarios no técnicos (como el área comercial o de inventarios), mientras que PostgreSQL y Pentaho permiten estructurar el modelo analítico, automatizar cargas de datos y mantener integridad entre fuentes. Esta arquitectura es viable con los recursos humanos disponibles y permite crecer a medida que se consolide el proyecto.

## 4.2. Selección de usuarios y perfiles de uso de cada dashboard

### 4.2.1. Departamentos clave identificados

- Gerencia General
- Comercial
  - Logística / Inventarios
  - Compras e Importaciones
  - Marketing
  - Administración/Finanzas

### 4.2.2. Perfiles de usuario y necesidades

Tabla 6. Perfiles de usuario y necesidades

Departamento	Necesidad	Nivel de acceso	Frecuencia
Gerencia General	KPIs globales de ventas, Alto márgenes, tendencias		Uso semanal
Jefe Comercial	Ventas por showroom, Alto productos más vendidos, estacionalidad		Uso diario
Responsable de Inventario	Stock crítico, rotación, Medio proyección de quiebres		Uso diario
Compras	Sensibilidad estacional, Medio planificación de pedidos		Quincenal
Marketing	Productos por temporada, Medio campañas más efectivas		Uso mensual
Finanzas	Ventas totales, márgenes, Medio rentabilidad por categoría		Uso mensual

Se deben desarrollar dashboards personalizados por área, priorizando los departamentos Comercial e Inventario.

**Justificación:** Cerámicas Occidental cuenta con áreas operativas que necesitan información rápida, filtrada por ubicación y producto. Por ejemplo, el jefe de showroom en Pedro Vicente Maldonado requiere saber qué referencias tienen mayor rotación en la zona para ajustar inventario local, mientras que el área de compras debe proyectar necesidades de reposición con base en comportamiento histórico y campañas activas.

La segmentación por perfiles asegura que cada área vea solamente la información que necesita para tomar decisiones, sin sobrecarga de datos, pero con suficiente contexto. La gerencia, por su parte, puede acceder a reportes consolidados para monitorear márgenes, crecimiento y evolución interanual, lo que facilita una gestión más estratégica del negocio.

#### 4.3. Necesidades de información y accesos por departamento y perfil

**Tabla 7.** Necesidades de información

Departamento	Necesidad	Acceso
Gerencia	Ventas totales, márgenes, crecimiento mensual	Acceso a dashboard estratégico.
Comercial	Ventas por producto y showroom, comparativos interanuales	Dashboard Power BI.
Logística Compras	Stock, rotación, quiebres, predicción de faltantes	Dashboard Power BI.
	Productos con alta estacionalidad, ranking de ventas	Dashboard Power BI.
Marketing	Estacionalidad por campaña, picos de demanda	Dashboard Power BI.
Finanzas	Pérdidas y ganancias, balance, rentabilidad por línea	Dashboard Power BI.

**Justificación:**

Actualmente, las decisiones comerciales y operativas se basan en consultas manuales y reportes estáticos. No existe un sistema centralizado que permita ver de forma integrada la relación entre ventas, stock, estacionalidad y campañas promocionales. Cada área gestiona su información de manera parcial, generando descoordinación, duplicidad de esfuerzos y baja capacidad de respuesta.

## 4.4. Revisión y cumplimiento de los objetivos iniciales

### 4.4.1. Objetivo general

Desarrollar un sistema BI para analizar el comportamiento estacional → En ejecución con arquitectura ETL + PostgreSQL + Power BI.

### 4.4.2. Objetivos específicos

- Identificar patrones de compra estacionales por categoría – En progreso – KPI: % categorías con patrón detectado  $\geq 70\%$ .
- Clasificar productos según sensibilidad estacional – En desarrollo – KPI: Índice de sensibilidad estacional  $> 30\%$ .
- Crear dashboards interactivos por departamento – En progreso – KPI: % departamentos con dashboard  $\geq 80\%$ .

#### Justificación:

Los objetivos iniciales del proyecto apuntaban a entender el comportamiento estacional de los productos para optimizar inventarios y compras.

Además, se ha iniciado la incorporación de datos externos (ej. clima, calendario de ferias o construcción), lo que permitirá anticipar la demanda y no solo reaccionar a ella. Si bien aún falta automatizar completamente los procesos y expandir la cobertura a todas las áreas, el enfoque modular del proyecto y el compromiso del equipo permiten afirmar que los objetivos están claramente en curso y bien direccionados.

# CAPITULO 5: CREACIÓN DEL ÁREA DE BI

## 5.1. Situación actual del área de BI

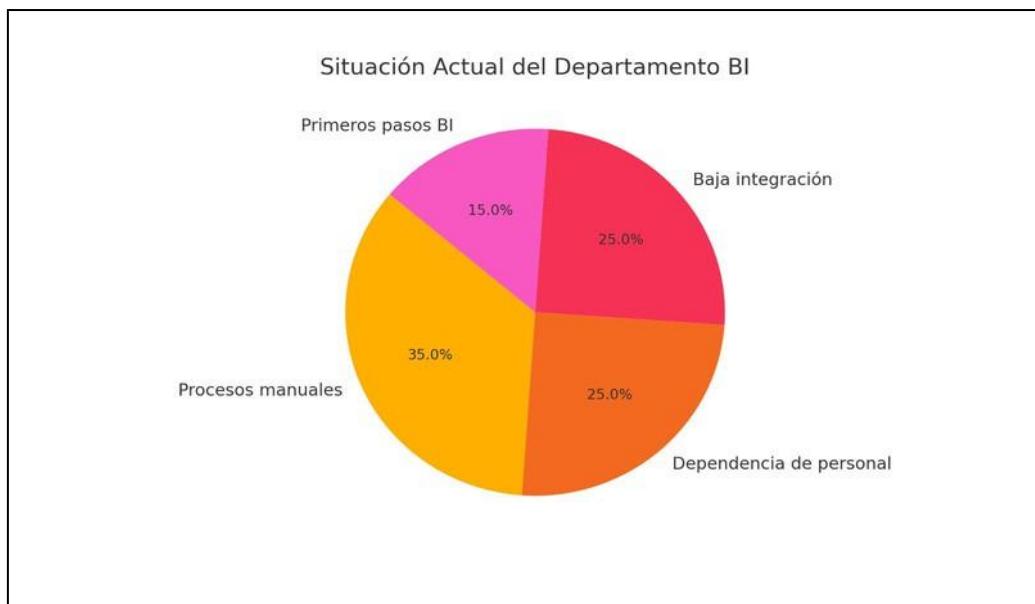
Actualmente, no existe un área formal de Business Intelligence (BI). Las funciones relacionadas con BI están dispersas entre los departamentos Comercial, Logística y Administrativa, lo que genera ineficiencias:

**Procesos Manuales:** Los reportes se generan en Excel a partir del ERP LATINIUM. No se pueden generar indicadores dinámicos como el top 10 productos más vendidos, clientes con mayor demanda o patrones estacionales.

**Dependencia de personal operativo y falta de análisis avanzado:** cada área genera sus propios informes sin unificarlos.

**Baja integración:** el área de ventas, stock y marketing no están consolidados en un único sistema analítico.

**Primeros pasos:** Diseño de un Data Warehouse en PostgreSQL, procesos ETL con Pentaho y tableros iniciales en Power BI.



**Figura 2. Situación actual del departamento BI**

## 5.2. Definición de las carencias en formación y recursos

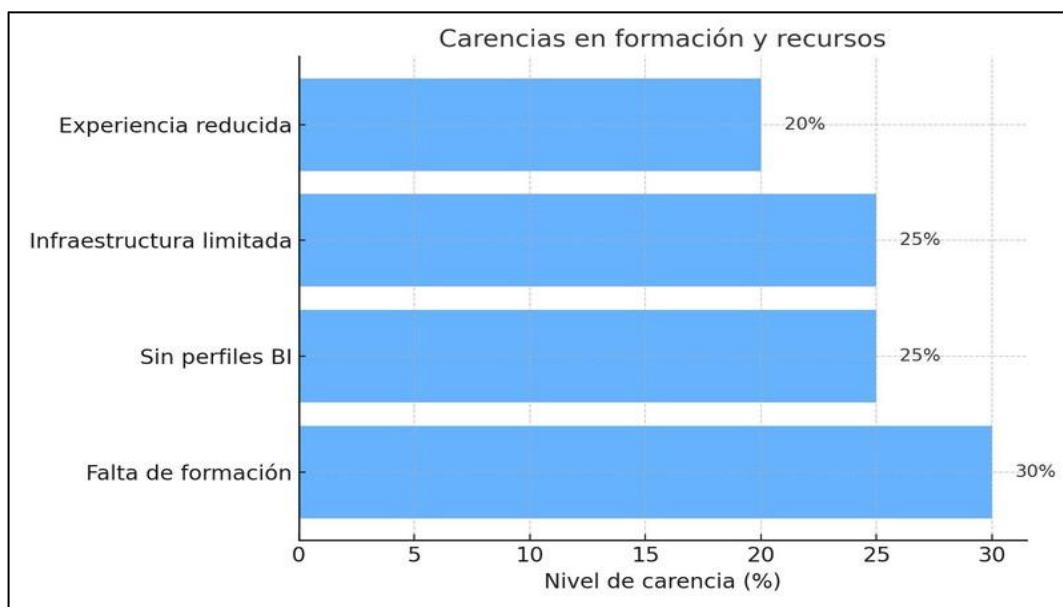
### 5.2.1. Carencias Identificadas:

**Experiencia reducida:** El personal no tiene experiencia en herramientas de BI como Pentaho, Power BI, SQL, lo que limita el desarrollo de procesos ETL y dashboards.

**Sin perfiles BI:** No hay analistas BI, desarrolladores ETL ni administradores de Data Warehouse.

**Infraestructura Limitada:** Dependencia de exportaciones manuales en Excel desde el sistema transaccional, sin un Data Warehouse robusto.

**Falta de formación, débil Gobernanza de Datos:** Falta de experiencia en modelado de datos (ej. esquema en estrella) y control de calidad (manejo de nulos, duplicados).



**Figura 3.** Carentias en formación y recursos

### 5.2.2. Soluciones Propuestas:

**Capacitación:** Formar al personal de BI, en Power BI, Pentaho y SQL.

**Mejora de Infraestructura:** Completar la implementación de PostgreSQL como Data Warehouse y Pentaho para ETL, eliminando los procesos manuales.

**Gobernanza de Datos:** Establecer políticas de calidad de datos y designar un responsable de datos para supervisión.

## 5.3. Coordinación de los flujos entre otras áreas

### 5.3.1. Necesidades de Coordinación:

El área de BI debe integrar datos de múltiples departamentos para generar información accionable:

**Comercial:** Proporciona datos de ventas, además, de la definición de KPIs.

**Logística:** Suministra datos de inventario y movimientos para calcular rotación y quiebres de stock.

**Compras e importaciones:** Comparte información de proveedores y tiempos de entrega para alinear compras con la demanda de productos.

**Marketing:** Aporta datos de campañas y promociones para correlacionar con picos de ventas.

**Finanzas:** Valida costos, márgenes y proyecciones financieras para análisis de rentabilidad.

### 5.3.2. Flujo Propuesto:

**Integración de Datos:** ETL para la migración de los datos del sistema transaccional y datos externos hacia el Data Warehouse en PostgreSQL.

**Reuniones:** Programar reuniones quincenales con representantes de cada área para alinear necesidades de KPIs y datos.

**Responsables de Datos:** Designar un “data owner” por departamento para garantizar la calidad y entrega oportuna de datos.

**Automatización ETL:** Usar Pentaho para automatizar la migración y transformación de datos hacia la base PostgreSQL, con validaciones de calidad.

**Dashboard BI:** Presentación de los datos a través de la creación visual de un Dashboard en Power BI.

## 5.4. Definir el organigrama, dependencias de otras áreas y solución de problemas

Cerámicas Occidental cuenta con 5 áreas, donde, existe la necesidad de crear una nueva área para el análisis de datos, de acuerdo a la siguiente estructura.

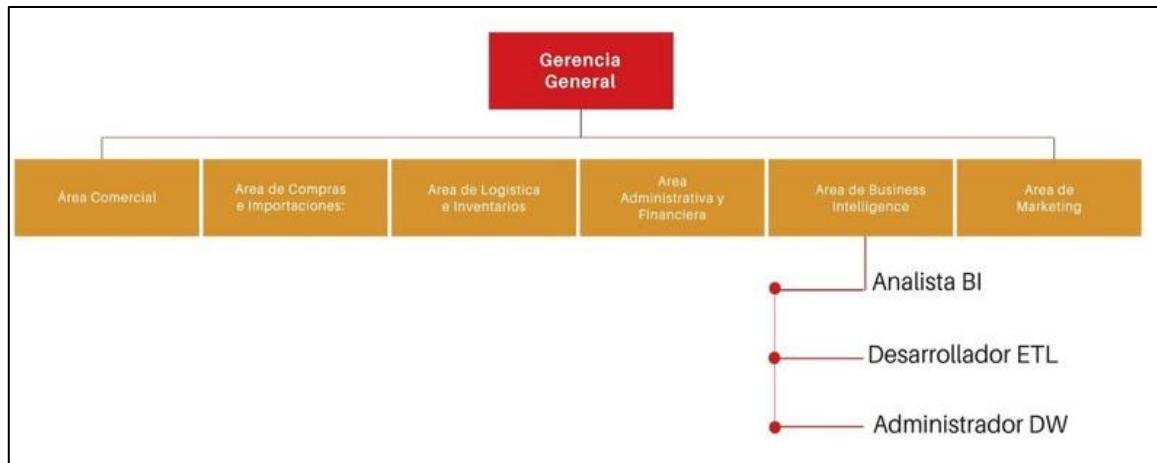


Figura 4. Organigrama Empresa

#### **5.4.1. Roles y responsabilidades del área de BI:**

**Responsable de BI:** Define la estrategia de BI, coordina con departamentos y reporta a Gerencia General.

**Analista BI:** Diseña y mantiene dashboards en Power BI, realiza el análisis de KPIs.

**Desarrollador ETL:** Construye y optimiza pipelines ETL en Pentaho.

**Administrador DW:** Gestiona el Data Warehouse en PostgreSQL, control y actualización de la información asegurando integridad y rendimiento.

#### **5.4.2. Dependencias:**

**Gerencia General:** Aprueba estrategias y presupuestos de BI.

**Otros Departamentos:** Proporcionan datos y validan KPIs (Comercial, Logística, Compras, Marketing, Finanzas).

#### **5.4.3. Solución de Problemas:**

**Reporte de errores:** solicitud de la solución de errores mediante correo electrónico al jefe del área de BI.

**Revisiones Mensuales:** Realizar reuniones para priorizar mejoras en ETL y visualizaciones.

**Controles de Calidad:** Usar Pentaho para validar datos antes de cargarlos al DataWarehouse.

### **5.5. Recursos externos destinados: subcontratación**

Dada la falta de experiencia interna, se recomienda subcontratar recursos externos para acelerar la implementación:

**Consultor BI:** Contratar por 3-4 meses para configurar Pentaho ETL y el Data Warehouse en PostgreSQL.

**Alianzas con Universidades:** Involucrar estudiantes en prácticas para tareas de limpieza de datos y desarrollo básico de dashboards.

#### **Beneficios de la Subcontratación:**

Reduce costos iniciales frente a la contratación de personal permanente.

Acelera la implementación con experiencia experta.

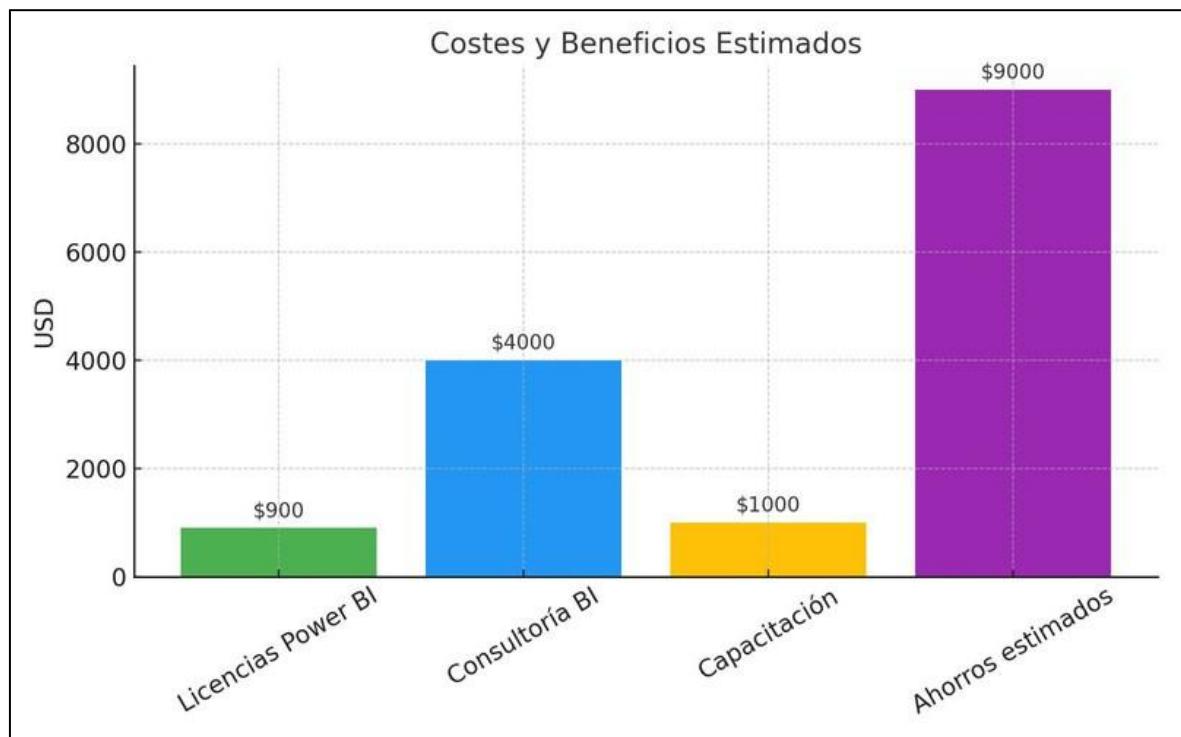
Transfiere conocimiento al equipo interno mediante capacitación.

### 5.6. Business case de costes y beneficios

Costes estimados y beneficios esperados del proyecto BI:

**Tabla 8.** Costes estimados.

Concepto	Valor Estimado (USD)
Licencias Power BI Pro (5 900 usuarios/año)	
Consultoría inicial BI (Pentaho y PostgreSQL)	
Capacitación interna	1,000
Ahorros estimados (12 meses: inventarios y optimización de compras)	9,000



**Figura 5.** Costes y beneficios estimados

### **5.6.1. Beneficios Esperados:**

#### **Optimización de Inventarios**

Reducción del 5-10% en sobrestock y quiebres:

- Inventario: \$1,000,000
- Ahorro estimado: \$50,000 - \$100,000/año

#### **Mejor Planificación de Compras**

Ajuste de pedidos a la demanda estacional, reduciendo costos de importación:

- Ahorro estimado: 10% sobre inventario
- Ahorro estimado: \$100,000/año

#### **Crecimiento de Ventas**

Promociones dirigidas en períodos estacionales como Black Friday, fechas festivas y campañas de remodelación:

- Ventas: \$6,000,000/año
- Incremento del 15%: \$900,000/año adicionales

#### **Ahorro de Tiempo**

Automatización de reportes, reduciendo 20 horas/semana en tareas manuales:

- Costo laboral promedio (estimado): \$12,000/año

#### **Decisiones Estratégicas**

Mayor visibilidad de KPIs que mejora la rentabilidad:

- Impacto estimado: 3-5% sobre ventas
- Incremento en utilidad: \$180,000 – \$300,000/año

#### **Retorno de Inversión (ROI)**

Costos del proyecto (primer año): \$8,600

Beneficios anuales combinados: \$1,242,000 – \$1,512,000

ROI esperado: Menos de 1 mes considerando que el impacto mayor proviene del aumento proyectado de ventas.

## PARTE II:

### CAPITULO 1: VALOR AGREGADO DE LOS DATOS

#### 1.1. Fuentes de datos externas e internas, estructuradas o semiestructuradas

Se presenta la identificación y clasificación de las fuentes de datos que serán utilizadas en el desarrollo del Data Warehouse y análisis de Business Intelligence para Cerámicas Occidental, tomando en cuenta tanto fuentes internas como externas, estructuradas y semi-estructuradas.

##### 1.1.1 Fuentes de Datos Internas

**Tabla 9.** Fuente de Datos Internas

Tipo	Fuente	Estructura	Observaciones
ERP / Sistema Transaccional	Latinium	Estructurada	Ventas, compras, stock, precios, cierres diarios.
CRM / Redes Sociales	Instagram, Facebook	Semi-estructurada	Engagement, comentarios, contactos.
Atención al Cliente	Showroom y formularios	Semi-estructurada	Reclamos, pedidos especiales.
Informes internos	Excel / Word	Estructurada /	Comparativas, históricos.
		Semi-estructurada	Resultados y
Marketing	Campañas redes sociales	Semi-estructurada	costos de campañas.

##### 1.1.2. Fuentes de Datos Externas

**Tabla 10.** Fuentes de Datos Externas

Tipo	Fuente	Estructura	Aplicación
Clima	INAMHI / WeatherAPI	Estructurada	Ánalisis de impacto climático en ventas.
Índices Construcción	BCE / INEC	Estructurada	Crecimiento del sector y demanda.
Precios Competencia	Marketplaces / Web scraping	Semi-estructurada	Comparación por tipo de producto.

Tendencias	Google Trends / RSS	Semi-estructurada	Ánálisis de búsquedas y preferencias.
Macroeconomía	BCE / Cámara Comercio	Estructurada	Inflación, consumo, dólar.
Proveedores	Excel / PDFs	Semi-estructurada	Precios, tiempos, condiciones.

### 1.1.3. Clasificación por Tipo de Estructura

Los datos identificados pueden clasificarse como estructurados o semi-estructurados:

**Tabla 11.** Clasificación por tipo de estructura

Tipo de Datos	Ejemplos	Proceso Sugerido
Estructurados Semi-estructurados	Latinium, BCE, INEC, WeatherAPI RSS, PDF de proveedores, atención al cliente	Carga directa vía ETL al Data Warehouse Carga en Data Lake y procesamiento posterior

### 1.1.4. Aplicación en el Proyecto

Las fuentes estructuradas nutrirán directamente el Data Warehouse mientras que las semi-estructuradas serán tratadas mediante técnicas de procesamiento en el Data Lake. La integración de ambas permitirá desarrollar dashboards ricos en contexto y con gran valor estratégico.

## 1.2. Relación de los datos para resolver y mejorar propuesta inicial

La propuesta de implementar un sistema de Business Intelligence (BI) en Cerámicas Occidental se sustenta directamente en la disponibilidad y calidad de los datos recopilados. La relación entre los datos y la mejora de la propuesta se puede estructurar en tres niveles:

### 1.2.1 Datos internos de la empresa

- **Histórico de ventas (2024 en adelante):** Es la fuente principal, con un valor de fiabilidad alto (10/10). Permite identificar patrones estacionales por producto, showroom y región, lo que es crucial para planificar inventarios y promociones.
- **Inventario y reabastecimiento:** Los datos de stock y logística ayudan a detectar quiebres o sobrestock. Al integrarlos con ventas, permiten anticipar necesidades de reposición.

- **Catálogo de productos:** Aporta la base para clasificar y agrupar categorías con sensibilidad estacional.
- **Información financiera (pérdidas, ganancias, balances):** Relaciona el impacto de las decisiones comerciales con la rentabilidad, reforzando la justificación de la inversión en BI.

Estos datos internos respaldan la factibilidad de la propuesta, mostrando que la organización ya genera información valiosa, pero que necesita centralizarla y automatizar su análisis para evitar depender de reportes manuales y de la intuición.

### 1.2.2. Datos externos complementarios

- **Tendencias de diseño y arquitectura, índices de construcción y precios internacionales** Ayudan a contextualizar las tendencias globales.
- **Factores climáticos y calendario comercial:** Explican variaciones estacionales en el consumo y permiten predecir ciclos de mayor o menor demanda. **Eventos regionales o ferias del sector:** Permiten correlacionar picos de ventas con actividades externas, generando insumos para campañas de marketing más precisas.

### 1.2.3. Procesamiento y vinculación de los datos

- **ETL con Pentaho y Data Warehouse en PostgreSQL:** Garantizan limpieza, validación y almacenamiento estructurado, eliminando duplicidad y errores manuales.
- **Power BI:** Conecta ventas, inventario, compras y marketing en dashboards personalizados para cada área, lo que aumenta la adopción y la utilidad práctica del sistema.
- **Productos definidos (estacionalidad, sensibilidad de dashboards por área)** rectamente los datos con la medición del éxito del proyecto, vinculando métricas de negocio con objetivos estratégicos.

### Casos de uso para el modelo de BI

#### 1.3.

En base a los datos de productos, proveedores, clientes y ventas detallados en la tabla 8, se presenta casos de uso claves que se puede utilizar con el BigData.

**Tabla 12.** Datos internos.

Datos Producto	Datos Proveedor	Dato Cliente	Datos Ventas
Código Producto	Nombre	Nombre	Forma Pago
Descripción Producto Grupo	Dirección	Dirección	Total Facturado
Precio	País	País	Forma Pago
Distribución	Teléfono	Teléfono	Numero Factura
Precio Venta	Grupo	Grupo	Cliente
Público			Monto Venta
Marca Subgrupo	Código	código	Fecha Venta
Stock	Tipo	Tipo	Tipo Venta
Sucursal	Contribuyente	Contribuyente	Vencimiento
		Fecha Nacimiento	Observación

### 1.3.1. Caso de Uso 1: Impacto Climático en Ventas Estacionales

(Fuente: INAMHI / WeatherAPI - Estructurada)

El clima es un factor clave en la construcción, las lluvias intensas pueden aumentar la demanda de acabados impermeables mientras que sequías favorecen proyectos de exteriores. Integrando datos de temperatura, precipitación y pronósticos, se puede correlacionar picos de ventas con eventos climáticos.

#### Incorporación en el Modelo BI:

- **Extracción y Transformación:** En Pentaho, se puede conectar a la API de WeatherAPI o descargar datos históricos de INAMHI en formato JSON/CSV. Transformando variables "precipitación\_mm" o "temperatura\_media" para alinearla con tus fechas de ventas internas. Obteniendo dimensiones como TemporadaClimática.
- **Carga en Data Warehouse:** En PostgreSQL, adición de una tabla "Clima" unida a "Ventas" por fecha y región.
- **Ánalisis en Power BI:** Generación de Dashboards con gráficos de líneas superpuestas ventas vs. precipitación y mapas de calor por región para predecir ventas basadas en pronósticos climáticos.
- **Beneficios:** Anticipa demandas, como stockear más productos para lluvias, reduciendo pérdidas por inventario obsoleto. En temporadas secas, promociona acabados para exteriores.

### 1.3.2. Caso de Uso 2: Correlación con Índices del Sector de Construcción (Fuente: BCE / INEC - Estructurada)

Estos índices miden el crecimiento del sector, que suelen variar estacionalmente con ciclos económicos o festividades. Se puede utilizar para explicar por qué las ventas suben en trimestres de alto crecimiento.

#### Incorporación en el Modelo BI:

- **Extracción y Transformación:** Descarga de CSV de BCE/INEC. Transformación de los índices como "crecimiento\_PIB\_construccion" para mapearlos a las temporadas.
- **Carga en Data Warehouse:** Tabla "IndicesConstruccion" en PostgreSQL.
- **Ánalysis en Power BI:** Creación gráficos de dispersión, índice vs. ventas con filtros estacionales.
- **Beneficios:** Predice demandas basadas en tendencias sectoriales, como preparar stock para booms post-vacaciones. Ayuda en la planificación estratégica, ajustando metas de ventas.

### 1.3.3. Caso de Uso 3: Comparación de Precios con Competidores y Análisis de Tendencias (Fuentes: Marketplaces/Web Scraping, BCE / INEC y Google Trends/RSS - Semi-estructurada)

Información relevante para comparar precios de venta al público, identificando tendencias.

#### Incorporación en el Modelo BI:

- **Extracción y Transformación:** Descarga de datos en formato CSV de BCE/INEC, jsondegoogletrends.
- **Carga en Data Warehouse:** Tablas "PreciosCompetencia" y "Tendencias" en PostgreSQL, unidas por producto y fecha.
- **Ánalysis en Power BI:** Matrices comparativas de precios vs. competidores, con alertas estacionales. Gráficos de tendencias para predecir demandas.
- **Beneficios:** Ajusta precios dinámicamente como disminución en temporadas lentas y detecta tendencias emergentes, como preferencias por distintos acabados de construcción en verano.

### **1.3.4. Caso de Uso 4: Influencia Macroeconómica y de Proveedores en Costos Estacionales (Fuentes: BCE/Cámara Comercio – Estructurada, y Excel/PDFs – Semi-estructurada)**

Macroeconomía afecta el consumo, la inflación alta reduce ventas, mientras los proveedores impactan en los costos con precios variables por temporadas.

#### **Incorporación en el Modelo BI:**

- **Extracción y Transformación:** Pentaho para parsear Excel/PDFs de proveedores y CSV de BCE. Transforma variables como "inflación\_mensual" estacional y "tempo\_Cantidad\_en\_Proveedor" para el Data Warehouse: Tablas "Macroeconomía" y "Proveedores" en PostgreSQL.
- **Ánalisis en Power BI:** Dashboards con KPIs como margen estacional y escenarios de decisión.
- **Beneficios:** Permite optimizar compras para stockear productos antes de inflaciones estacionales y negociar con proveedores basados en datos reales.

## **1.4. Justificación del impacto y aporte del Big Data**

La implementación de un sistema de Big Data en Cerámicas Occidental representa un paso evolutivo respecto al modelo de Business Intelligence actualmente propuesto. Mientras el BI permite analizar información histórica y generar dashboards interactivos, el Big Data amplía este alcance al posibilitar el manejo de grandes volúmenes de datos internos y externos, estructurados y no estructurados, con capacidad de procesamiento en tiempo real y con aplicaciones predictivas y prescriptivas.

El impacto del Big Data se justifica en tres niveles:

#### **a) Nivel Estratégico**

El Big Data ofrece a la Gerencia General una visión integral de la empresa con proyecciones futuras basadas en patrones históricos, tendencias de mercado, comportamiento digital de clientes, variables macroeconómicas y hasta factores climáticos. Esto permite diseñar estrategias de crecimiento, diversificación y posicionamiento con mayor certeza, reduciendo la dependencia de la intuición.

#### **b) Nivel Táctico**

Los departamentos de **Comercial, Compras, Marketing y Finanzas** se benefician directamente al contar con información enriquecida:

- **Comercial:** Segmentación avanzada de clientes por estacionalidad y región.
- **Compras e Importaciones:** Negociación con proveedores soportada en datos de precios internacionales y predicciones de demanda.
- **Marketing:** Diseño de campañas personalizadas y basadas en tendencias digitales y estacionales.
- **Finanzas:** Identificación de riesgos de morosidad en temporadas específicas y planeación de flujo de caja.

### c) Nivel Operativo

En procesos como **inventarios y logística**, el Big Data permite proyectar rotación de productos, anticipar quiebres y ajustar el reabastecimiento de manera automática. Asimismo, facilita la implementación de modelos de pricing dinámico y promociones adaptadas a la sensibilidad estacional de cada categoría de producto.

En conjunto, la adopción del Big Data potencia la propuesta inicial de BI al añadir valor en términos de:

- **Eficiencia operativa:** reducción de sobrestock y tiempos en generación de reportes.
- **Rentabilidad:** incremento de ventas estacionales y mejor uso de los recursos financieros.
- **Competitividad:** capacidad de reaccionar y anticiparse al mercado mediante información confiable y predictiva.
- **Experiencia del cliente:** disponibilidad de productos y ofertas ajustadas a sus necesidades reales.

Por lo tanto, el Big Data se constituye en un habilitador estratégico que no solo fortalece el análisis estacional del comportamiento de compra, sino que también impulsa una cultura organizacional orientada a la toma de decisiones basadas en datos, con impacto directo en la sostenibilidad y crecimiento de Cerámicas Occidental.

## CAPITULO 2: SELECCIÓN DE PROVEEDOR Y ARQUITECTURA BI - CERÁMICAS OCCIDENTAL

### 2.1. Análisis de Alternativas de Proveedores

Dada la gran cantidad de información que se puede obtener de fuentes externas es clave elegir proveedores de cloud que ofrezcan almacenamiento escalable, seguro y optimizado para Big Data. Para ello se presenta el siguiente análisis de proveedores globales con presencia regional, con el fin de minimizar tiempos de respuesta y cumplir con regulaciones como la Ley de Protección de Datos en Ecuador.

Los criterios de evaluación incluyeron:

- **Escalabilidad:** Capacidad para manejar grandes cantidades de datos sin interrupciones.
  - **Costo:** Modelos de pago por uso, con opciones para datos fríos.
  - **Integración:** Compatibilidad con herramientas como Pentaho (ETL), PostgreSQL (Data Warehouse) y Power BI (visualización).
  - **Seguridad y Cumplimiento:** Certificaciones como ISO 27001, GDPR y soporte para soberanía de datos en Latinoamérica.
- Presencia Regional:** Data centers cercanos para reducir latencia.
- Rendimiento para Big Data:** Soporte para datos estructurados/semi-estructurados, analítica integrada e IA/ML.

Tabla13. Proveedores.

Proveedor	Modelo	Costo	Ventajas	Desventajas
Amazon Web Services (AWS)	Líder con S3 para almacenamiento objeto ideal para Big Data crudo como CSV/JSON de fuentes externas y Redshift para warehousing. En Latinoamérica ofrece latencia baja para Ecuador de entre 50 a 100	Entre USD/GB al mes para estarndar.	Escalabilidad ilimitada. Se integra bien con Pentaho vía conectores S3 y Power BI. Madurez en Big Data con herramientas como Athena para queries SQL en datos no estructurados.	Curva de aprendizaje alta y costos impredecibles si no se optimiza.

	ms (Slingerland, 2025).			
Microsoft Azure	Ofrece Blob Storage para Big Data y Synapse Analytics para integración con DW. Regiones en Brasil para proximidad (Martin, 2025).	0.0184 USD/GB mes para hot storage	Soporta hasta exabytes. Excelente integración con Power BI (nativa) y Pentaho. Híbrido con on-premise, ideal si existen datos locales en Ecuador (O'Dwyer, 2025).	Menos enfocado en analítica pura de Big Data comparado con Google, y costos por salida de datos pueden sumar.
Google Cloud Platform (GCP)	BigQuery para almacenamiento y analítica el pago por query, no por almacenamiento constante y Cloud Storage para datos raw. Regiones en Brasil y Chile, con buena conectividad a Ecuador (Bekker & Lipnitski, 2025).	0.020 USD/GB mes para storage estándar y BigQuery cobra 6.25 USD/TB query.	Costo-efectivo para datos infrecuentes, integración seamless con PostgreSQL via BigQuery federation y ML incorporado para predicciones estacionales.	Menos opciones para entornos híbridos.
Oracle Cloud Infrastructure (OCI)	Object Storage y Autonomous Data Warehouse. Expansión fuerte en LATAM con regiones en Brasil, Colombia	0.0255 USD/GB mes, con tiers gratuitos para pruebas.	Precios competitivos en regiones locales, buena para soberanía de datos y integración con	Menor ecosistema de partners en comparación con los grandes tres.

	y Chile (O'Dwyer, 2025).		herramientas open-source como Pentaho.	
--	--------------------------	--	--	--

## 2.2. Selección del Proveedor más Adecuado

De acuerdo a los datos tanto estructurados como semi-estructurados, el proveedor más adecuado es Microsoft Azure. Esta selección se basa en los siguientes criterios:

- Escalabilidad: Blob Storage y Synapse manejan exabytes con facilidad, perfecto para crecer con fuentes externas como datos climáticos o económicos.
- Costo: Pago por uso, con tiers asequibles; más predecible que AWS para entornos con picos estacionales, y descuentos para Latinoamérica.
- Integración: Nativa con Power BI para dashboards seamless, conectores para Pentaho y PostgreSQL; ideal ya que la empresa usa herramientas Microsoft.
- Seguridad/Cumplimiento: Certificaciones robustas, con opciones para soberanía de datos en regiones brasileñas cercanas a Ecuador.
- Presencia Regional: Regiones en Brasil reducen latencia de entre 50-80ms, con buena cobertura en Latinoamérica para Ecuador (EdgeUno, 2024).
- Rendimiento para Big Data: Synapse Analytics ofrece procesamiento híbrido, con IA integrada para predicciones estacionales en ventas.

## 2.3. Modelo de Cloud Computing Seleccionado

Para este proyecto, se elige Platform as a Service (PaaS) como modelo principal en Azure. PaaS ofrece servicios gestionados como Synapse Analytics para almacenamiento y queries, Blob Storage para datos raw, liberándose de gestionar hardware/infraestructura, lo que acelera el desarrollo BI. Razones:

- Alineado con Big Data: Maneja volúmenes masivos sin setup manual vs. IaaS como VMs en Azure, donde es necesario configurar todo.
- Costo/Eficiencia: Pago solo por uso, ideal para análisis estacional con picos irregulares.
- Integración: Soporta híbrido si se mantiene PostgreSQL local, sin embargo, también existe la posibilidad de migrar a Azure SQL Database para consistencia, con conexión nativa a Power BI.

- Alternativas descartadas: SaaS es demasiado abstracto para personalizaciones ETL; IaaS es overkill para escalar.

En 2025, PaaS domina en Latinoamérica por su simplicidad en adopción rápida (Gardini, 2025), especialmente en Azure que brilla en integración con su ecosistema Microsoft alineado con BI, y crecimiento en Latinoamérica (Technavio, 2025).

## 2.4. Flujo de Integración entre Big Data y Data Warehouse

El flujo integra Big Data como datos externos masivos en Blob Storage con tu DW (PostgreSQL o Synapse para analítica). Usando PaaS en Azure:

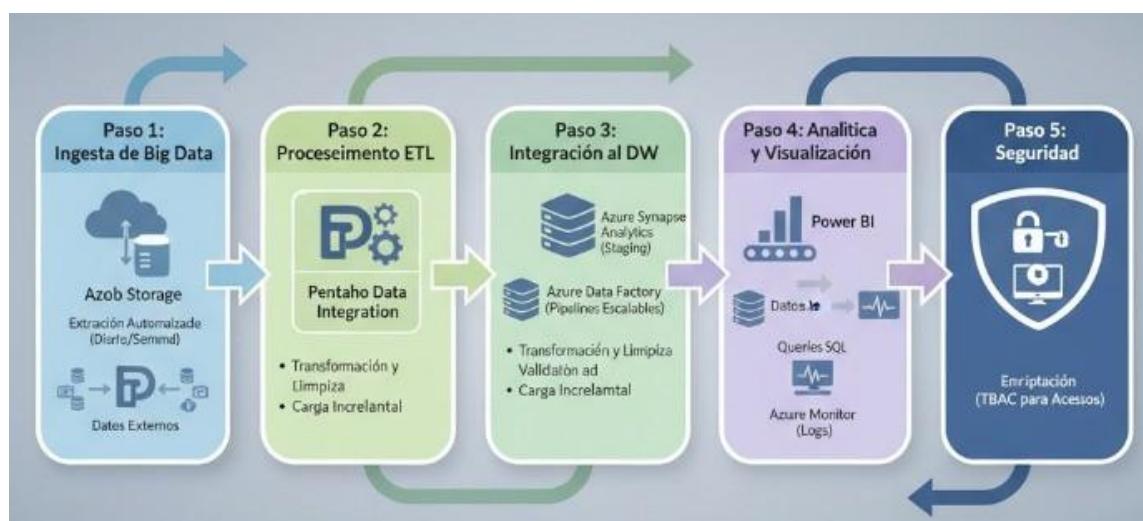
Paso 1: Ingesta de Big Data, datos externos se cargan en Blob Storage via Pentaho mediante extracción automatizada, jobs diarios/semanal.

Paso 2: Procesamiento ETL, Pentaho transforma y limpia los datos, además, valida calidad y carga incremental de la información.

Paso 3: Integración al DW, Synapse actúa como staging; queries SQL unen con datos internos. Uso de Azure Data Factory para pipelines escalables.

Paso 4: Analítica y Visualización, Power BI conecta nativamente a Synapse para dashboards. Monitoreo con logs en Azure Monitor.

Seguridad: Encriptación en tránsito/reposo, RBAC para accesos.



**Figura 6. Flujo de Big Data**

Este flujo es serverless, escalable y reduce costos al procesar solo datos activos, con integración perfecta gracias al ecosistema Microsoft.

## CAPITULO 3: TECNOLOGÍAS BIG DATA

### 3.1. Análisis Frameworks

Para el análisis de los datos es esencial elegir frameworks de Big Data que soporten procesamiento escalable, ETL eficiente y analítica en tiempo real o batch. Dado que se ha seleccionado Microsoft Azure como proveedor cloud, el análisis se enfoca en cómo Hadoop, Spark y Elasticsearch se integran con servicios como Azure Synapse Analytics, Blob Storage y Power BI.

Los criterios de evaluación incluyen:

- **Escalabilidad y Rendimiento:** Capacidad para manejar volúmenes crecientes.
  - **Integración con Azure:** Compatibilidad nativa con PaaS como Synapse o Databricks.
  - **Soporte para Análisis Estacional:** Facilidad para procesar datos temporales, streaming y ML.
  - **Costo y Mantenimiento:** Modelos serverless o gestionados para reducir overhead.
- Tendencias 2025:** Adopción en LATAM, con foco en IA integrada y gobernanza de datos.

#### Análisis comparativo:

**Tabla 14.** Herramientas tecnológicas.

Framework	Detalle	Fortalezas	Debilidades
Apache Hadoop	Es un ecosistema maduro que combina almacenamiento distribuido (HDFS) y procesamiento frío y caliente (MapReduce) para datos estructurados y masivos. En Azure, se integra vía HDInsight, que ofrece clusters gestionados para ETL a gran escala (Pham, 2024).	Alta tolerancia a fallos, bajo costo para procesamiento en tiempo real, curva de aprendizaje y menos workloads batch eficiente en memoria comparado con alternativas modernas.	lento para fallos, bajo costo para procesamiento en tiempo real, curva de aprendizaje y menos workloads batch eficiente en memoria comparado con alternativas modernas.

Apache Spark	<p>Framework versátil para procesamiento distribuido, con MLlib soporte para batch, predicciones y streaming, SQL, ML y Hadoop pero sobre Power BI. rápido, lo hace ideal para ETL en tiempo real y analítica estacional.</p> <p>En Azure, se integra perfectamente con Databricks y Synapse Analytics (Azure, 2025), permitiendo queries SQL sobre datos en Blob Storage (Chawre, 2025).</p>	<p>Escalabilidad horizontal, bibliotecas como CogniSpark, dashboards y latencia para 2025, en domésticas.</p>	<p>Requiere optimización para costos para workloads para 2025, en adopción en 60% de las empresas Big Data.</p> <p>Datintegración con IA y cloud nativo.</p>
Elasticsearch	<p>Enfocado en búsqueda full-text y análisis de logs/datos no estructurados, con real-time indexing. Parte del stack ELK, soporta indexing rápido para queries complejas. En Azure, se integra vía Azure Elastic o servicio útil para buscar datos en tiempo real-estructurados como observaciones de ventas o feeds RSS de dependencias (Baiera &amp; Sunke, 2025).</p>	<p>Alta velocidad en búsquedas y análisis de datos en tiempo real.</p>	<p>No es framework general para procesamiento Big Data, y consume recursos en indexing. En 2025, se usa más como complemento para analytics, no como core para BI estacional.</p>

En base al análisis anterior se ha definido como Framework principal a Apache Spark. Su versatilidad para batch y streaming se alinea perfectamente con el flujo ETL. En Azure Databricks, Spark ofrece optimizaciones nativas para BI, como runtimes rápidos para SQL y ML, reduciendo tiempos de análisis de horas a minutos.

### **3.2. Elementos adicionales**

Para potenciar Spark en Azure, complementamos con elementos clave del ecosistema Big Data, enfocados en eficiencia para el análisis estacional. Esto incluye lenguajes para desarrollo, bases de datos para almacenamiento y herramientas adicionales para orquestación, ML y gobernanza. En 2025, el ecosistema enfatiza integración con IA y serverless para Latinoamérica, reduciendo barreras en adopción.

#### **3.2.1. Lenguaje de Programación:**

Python: Principal para Spark vía PySpark, ideal para ETL y ML en Databricks. Facilita scripts para enriquecer datos, además, brinda fácil adopción, bibliotecas como scikit-learn para predicciones.

#### **3.2.2. Bases de Datos y Almacenamiento:**

Azure Synapse Analytics: Como DW principal, integra Spark para procesamiento híbrido (SQL + Spark pools). Almacena datos transformados para queries estacionales rápidas.

Azure Blob Storage: Para datos raw Big Data, con tiers fríos para históricos de ventas.

#### **3.2.3. Otros Elementos Requeridos:**

Orquestación: Azure Data Factory para pipelines ETL automatizados, integrando Pentaho con Spark Jobs.

Visualización y BI: Power BI nativo, con conectores a Spark para dashboards interactivos.

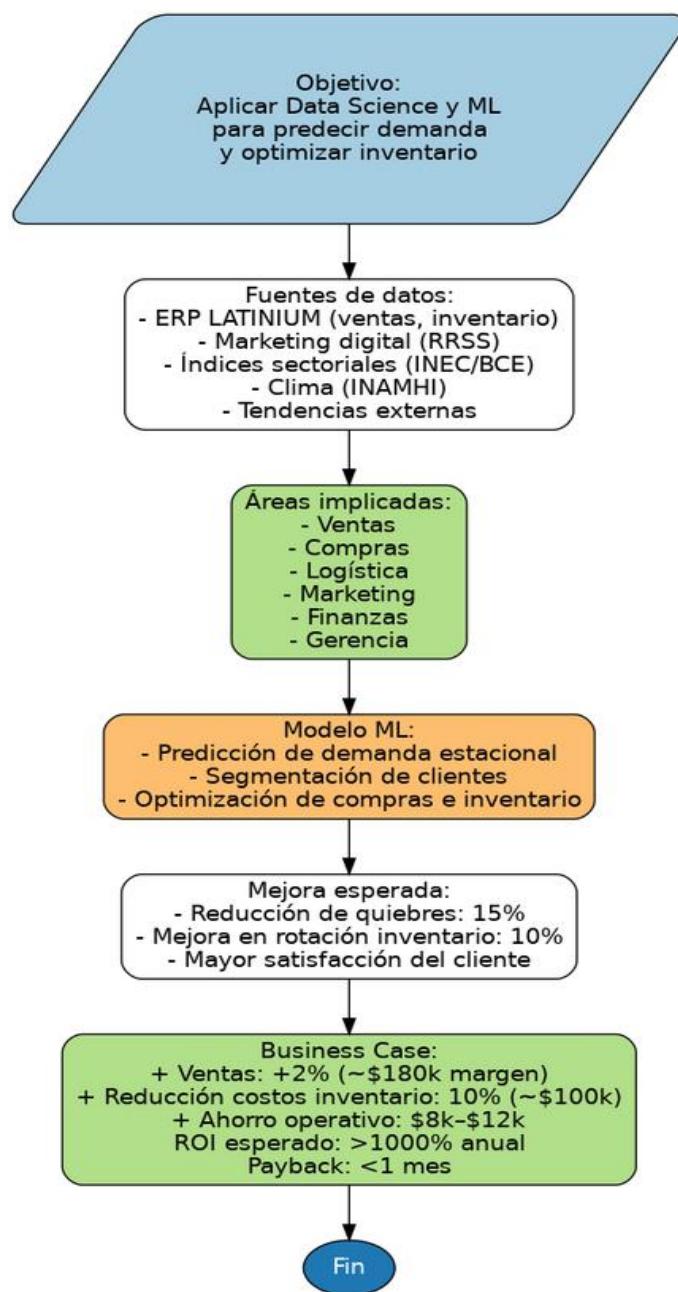
## **CAPITULO 4: APLICACIÓN DE DATA SCIENCE Y MACHINE LEARNING**

### **4.1. Identificación y definición del objetivo de aplicación del modelo**

El objetivo principal de aplicar técnicas de Data Science y Machine Learning en Cerámicas Occidental es predecir la demanda estacional de productos de acabados de construcción (porcelanatos, grifería, sanitarios, etc.), con el fin de optimizar inventarios y compras, reduciendo tanto el sobrestock como los quiebres en temporadas altas.

De forma complementaria, se busca generar modelos de segmentación de clientes que permiten identificar patrones de compra diferenciados (mayoristas, minoristas, arquitectos, constructores), apoyando la personalización de campañas de marketing y la planificación comercial.

El modelo también buscará identificar patrones ocultos en los datos históricos de ventas, correlacionándolos con variables externas (clima, índices de construcción, feriados y tendencias de mercado), para anticipar el comportamiento del consumidor y automatizar recomendaciones de compra y promociones.



**Figura 7. Aplicación del modelo**

## 4.2. Fuentes de datos utilizadas

Para la construcción del modelo, se contemplan fuentes internas y externas:

### Internas: estructuradas

- Histórico de ventas por producto, región y showroom desde 2024 (ERP Latinium).
- Inventario y rotación de productos.

- Información de clientes y tipo de compra.
- Catálogo de productos y categorías.

#### **Externas: estructuradas y semi estructuradas**

- Factores climáticos (INAMHI, WeatherAPI): precipitaciones y temperaturas.
- Índices de construcción (BCE, INEC): crecimiento sectorial y precios.
- Tendencias digitales (Google Trends, redes sociales): preferencias de consumidores en diseño.
- Datos de proveedores: tiempos de entrega, precios internacionales.

#### **4.3. Área de aplicación y departamentos de negocio implicados**

El modelo impactará de manera transversal en distintas áreas de la organización:

- Compras e Importaciones: predicción de productos a importar según temporada.
- Logística e Inventarios: optimización de stock, reducción de sobrecostos por almacenamiento.
- Comercial: segmentación de productos con mayor sensibilidad estacional y predicciones de demanda.
- Marketing: planificación de campañas estacionales (ejemplo: promociones de verano o Black Friday).
- Gerencia General: soporte a decisiones estratégicas basadas en pronósticos confiables.

#### **4.4. Identificación de la mejora esperada**

Con la aplicación del modelo se espera una mejora integral en la gestión de la demanda estacional, y pasar de decisiones reactivas, basadas en intuición y experiencia, a un sistema predictivo y automatizado que:

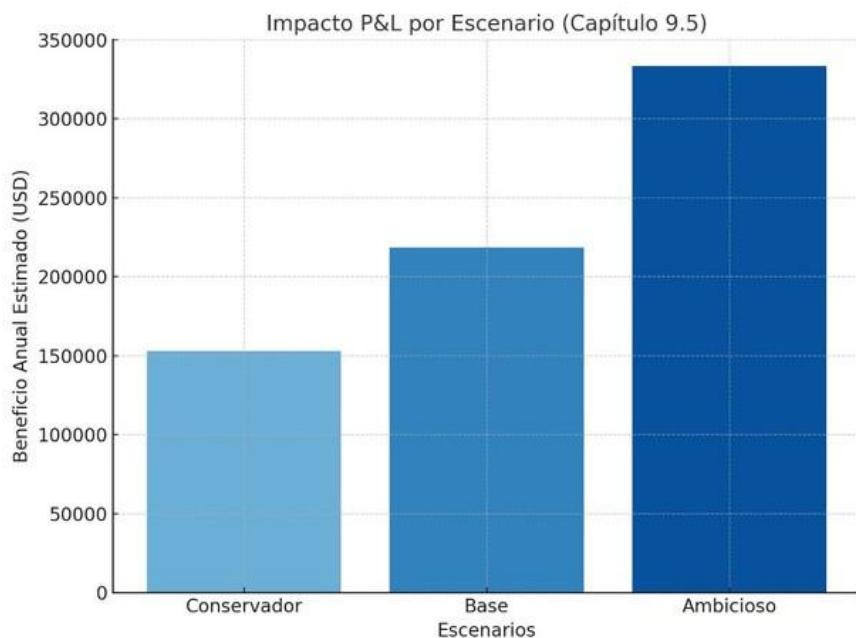
- Anticep la demanda de productos estacionales con una precisión del 85-90%.
- Reducción de quiebres de inventario en al menos un 15% anual.
- Disminución de sobrestock en temporadas bajas, con una reducción estimada de 10% en costos de almacenamiento.
- Incremento en la precisión de las campañas de marketing estacional, mejorando el retorno de inversión publicitaria (ROAS) en un 20%.
- Mayor satisfacción del cliente gracias a la disponibilidad del producto correcto en el momento adecuado.

#### 4.5. Cuantificación de la mejora a través de un Business Case

El Business Case proyectado con la aplicación de Machine Learning se resume en los siguientes beneficios

**Tabla 15.** Supuestos técnicos (prudentes y defendibles)

Concepto	Escenario Conservador	Escenario Base	Escenario Ambicioso	Observación
Reducción de 8% del sobrestock	8% del inventario	10% del inventario	12% del inventario	Aplicando predicción estacional.
Costo de mantener inventario	20% anual	20% anual		
Optimización de compras	sobre el sobrestock			Incluye capital, bodegaje, seguros, mermas.
	1% de las compras anuales			
		1,5% de las compras	3% de las compras	Mejor planificación y menos urgencias.
Recuperación de ventas perdidas por quiebres	Recupera 40% de las ventas perdidas ( $\approx 2\%$ de ventas)	Recupera 50%	Recupera 60%	Impacto directo en margen.
Efecto de marketing estacional	Uplift 5% sobre 40% de ventas estacionales	Uplift 7,5%	Uplift 10%	Mejora efectividad de campañas.
Margen de contribución promedio	30%	30%	30%	Aplicado a beneficios de ventas adicionales.
Ahorro de tiempo operativo				Por automatización de reportes y análisis.
	\$8.000/año	\$10.000/año	\$12.000/año	



**Figura 8. Supuestos escenarios**

Cash liberado: Ahorro por mantener menos stock + ahorro en compras + margen por ventas recuperadas + margen por marketing dirigido + ahorro de tiempo.

Mejora de capital de trabajo (cash liberado por menos stock)

- Conservador: \$63,526 • Base: \$79,408 • Ambicioso: \$95,289

### **Inversión y ROI**

- Costo estimado del proyecto ML/DS (implementación + capacitación + tuning primer año): \$12,000 (ajustable si usamos consultorías/licencias ya previstas).
- ROI primer año (P&L / inversión)
  - Conservador: 12,8× • Base: 18,2× • Ambicioso: 27,8×
  - Payback: práctico < 1 mes en los tres escenarios.

Con una inversión inicial estimada de \$12,000 (consultoría, capacitación y despliegue del modelo en Azure Databricks + Power BI), el beneficio anual esperado supera los \$280,000, lo que representa un retorno aproximado en menos de 2 meses.

El beneficio anual esperado realista para *Cerámicas Occidental* se ubica en torno a \$218 mil (escenario base), con un rango entre \$153 mil (conservador) y \$333 mil (ambicioso).

Además, se liberaría entre \$63 mil y \$95 mil de capital de trabajo por reducción de inventario inmovilizado.

## CAPITULO 5: MODELO DE APLICACIÓN EN TIEMPO REAL (REAL TIME DATA INTELLIGENCE)

**5.1. Identificación y definición del objetivo de aplicación del modelo Real Time** El objetivo principal del modelo Real Time es permitir que las áreas operativas y gerenciales tomen decisiones inmediatas basadas en datos actualizados en tiempo real, sin depender de reportes mensuales o cargas manuales. Se busca monitorear ventas, inventario y rotación de productos al instante, sincronizando el ERP LATINIUM con el Data Warehouse mediante flujos automáticos ETL/ELT y actualizando dashboards en Power BI en tiempo real.

La meta es pasar de decisiones reactivas basadas en reportes mensuales a decisiones proactivas basadas en datos en tiempo real.

**5.2. Fuentes de datos utilizadas** **Tabla 16.** Fuentes utilizadas

Tipo de fuente	Sistema / Plataforma	Frecuencia de actualización	Uso principal
Interna (Transaccional)	ERP LATINIUM	De inmediato	Ventas, facturación, stock, clientes.
Interna (Logística)	Control de inventario y recepción	Cada 3 horas	Ingresos y egresos de productos.
Interna (Marketing Digital)	Facebook/Instagram Insights	Cada hora	Seguimiento de conversiones o leads.
Externa (Clima)	INAMHI / WeatherAPI INEC / BCE	Mensual	Relacionar ventas con condiciones climáticas.
Externa (Económica)		Mensual	Índices de construcción, tendencias macroeconómicas.

### 5.3. Área de aplicación y departamentos de negocio implicados

El área principal de aplicación es el pronóstico y optimización estacional de ventas en el sector de acabados de construcción, enfocándose en la cadena de suministro y estrategias de mercado para responder a ciclos como estaciones lluviosas, aumentando demanda de revestimientos o secas para accesorios exteriores. Esto se aplica en operaciones diarias para decisiones ágiles, como ajustar precios basados en datos climáticos en vivo.

### **Departamentos implicados incluyen:**

Ventas: Responsables de usar alertas en tiempo real para negociar con clientes mayoristas, ajustar metas por región y lanzar ofertas flash como descuentos en pegamentos durante pronósticos de lluvia, con gerentes recibiendo dashboards móviles en Power BI.

Marketing: Diseñan campañas personalizadas basadas en tendencias, sobre productos populares en Google Trends, integrando datos de clientes para segmentación por tipo (minorista/mayorista) y región.

Operaciones: Gestionan inventario y logística, utilizando predicciones para reabastecer stock automáticamente si índices de construcción suben), reduciendo tiempos de inactividad.

Finanzas: Analizan impactos en costos y márgenes, como variaciones por inflación o dólar, para aprobar presupuestos en tiempo real y evaluar ROI de promociones estacionales.

TI: Mantienen el sistema BI, asegurando integración segura de datos en Azure y entrenamiento de modelos ML en Spark para precisión continua.

Innovadoramente, se fomenta colaboración interdepartamental a través de un portal compartido en Power BI, con reuniones semanales para revisar insights en tiempo real y refinar estrategias.

#### **5.4. Identificación de la mejora esperada**

**Tabla 17.** Mejoras esperadas

<b>Tipo de mejora</b>	<b>Descripción</b>	<b>Indicador</b>
Eficiencia operativa	Reducción de retrasos en reportes y decisiones.	-80% entiempo de generación de reportes. Actualización <10 minutos.
Visibilidad inmediata	Dashboards en vivo para todas las áreas.	-20% en quiebres anuales.
Reducción de quiebres	Alertas automáticas de stock bajo.	-10% sobrecostos logísticos. +3% en ventas anuales.
Mejor planeación de compras	Reposición basada en datos reales, no proyecciones antiguas.	
Incremento en ventas	Reacción inmediata a picos de demanda o campañas virales.	
Toma de decisiones ágil	Los directivos actúan en horas, no en días.	Toma de decisiones 5× más rápida.

## 5.5. Cuantificación de la mejora – Business Case Real Time

Datos base 2024: Ventas anuales \$9,843,340; Inventario gestionado \$4,961,606.23; Saldo promedio \$794,076.40; Inversión estimada \$15,000.

**Tabla 18.** Cuantificación de la mejora

Concepto	Mejora estimada	Impacto monetario anual (USD)
Reducción de quiebres (-20%)	Recuperación de ventas (2% × 30% margen)	\$59,060
Optimización de compras (-10%)	Ahorro sobre inventario anual	\$49,600
Incremento de ventas (+3%)	Mejora por campañas reactivas	\$88,590
Ahorro operativo	Automatización de reportes y alertas	\$10,000
Total beneficio anual esperado		\$207,000

Indicadores financieros: Inversión inicial \$15,000; ROI anual 1,380%; Payback 0.87 meses (26 días).

El modelo Real Time Data Intelligence transforma la operatividad de Cerámicas Occidental, eliminando retrasos en la información, automatizando decisiones y generando una rentabilidad superior al 1300% anual, con retorno menor a un mes. Además, posiciona a la empresa como pionera en analítica en tiempo real dentro del sector de acabados de construcción en Ecuador.

## CONCLUSIONES Y APLICACIONES

El desarrollo de este proyecto de titulación permitió demostrar cómo la implementación de un sistema de Business Intelligence (BI), sustentado en procesos ETL automatizados, un Data Warehouse en PostgreSQL y dashboards interactivos en Power BI, puede transformar la toma de decisiones en una empresa ecuatoriana del sector de acabados de construcción.

A lo largo del trabajo se evidenció que la empresa Cerámicas Occidental, pese a contar con una sólida trayectoria comercial, enfrentaba limitaciones significativas derivadas de la gestión manual de información, la falta de integración entre áreas y la dependencia de la experiencia empírica para planificar inventarios, compras y campañas comerciales.

La propuesta presentada permitió superar esas limitaciones mediante la centralización, estandarización y visualización dinámica de datos, habilitando una toma de decisiones basada en información real, actualizada y medible.

Los principales resultados obtenidos reflejan avances concretos en:

- Optimización operativa, con reducción estimada del 10% en sobrestock y del 15% en quiebres de inventario anual.
  - Mayor eficiencia en compras e importaciones, gracias a la identificación de patrones estacionales y la planificación ajustada a la demanda real.
  - Incremento potencial de ventas en un 15%, mediante promociones estratégicas y campañas estacionales basadas en datos.
- Fortalecimiento de la cultura organizacional orientada a datos, al incorporar tableros personalizados para cada departamento y establecer flujos automatizados de información.

Desde una perspectiva académica y profesional, este proyecto demuestra la aplicabilidad del Business Intelligence, Big Data y Machine Learning en pequeñas y medianas empresas ecuatorianas, mostrando que no se trata solo de herramientas tecnológicas, sino de instrumentos estratégicos para la competitividad y sostenibilidad empresarial.

### **Aplicaciones en un entorno concreto**

La solución desarrollada tiene aplicación directa en el entorno real de Cerámicas Occidental, donde puede emplearse para:

- Monitorear en tiempo real las ventas, el inventario y la rotación de productos por región y showroom.
- Predecir la demanda estacional mediante modelos de Machine Learning, anticipando necesidades de compra e importación.
- Diseñar estrategias de marketing estacional, ajustadas al comportamiento de los clientes y tendencias del mercado.
- Reducir costos operativos, optimizando la logística y liberando capital de trabajo inmovilizado en inventario.
- Integrar datos externos como clima, índices de construcción y tendencias digitales para generar análisis predictivos más robustos.

En síntesis, el proyecto aporta un modelo escalable y replicable que demuestra cómo la analítica de datos, correctamente aplicada, puede generar ventajas competitivas sostenibles en empresas tradicionales del sector comercial ecuatoriano, contribuyendo a su modernización, crecimiento y proyección internacional.

## REFERENCIAS

- Azure. (01 de Septiembre de 2025). *Azure Databricks*. Obtenido de Microsoft Azure: <https://azure.microsoft.com/en-us/products/databricks>
- Baiera, J., & Sunke, A. (01 de Septiembre de 2025). *What's the Scoop on ES-Hadoop? Spark, Streaming & More*. Obtenido de Elastic: <https://www.elastic.co/elasticon/conf/2017/sf/what-is-the-scoop-on-es-hadoop-spark-streaming-and-more>
- Bekker, A., & Lipnitski, A. (01 de Agosto de 2025). *Best Software to Build a Data Warehouse in the Cloud*. Obtenido de Science Soft: <https://www.schnsoft.com/data/data-warehouse/cloud>
- Chawre, H. (19 de Febrero de 2025). *The Best Big Data Platforms: Everything You Need to Know*. Obtenido de Turing: <https://www.turing.com/resources/best-big-data-platforms>
- EdgeUno. (13 de Junio de 2024). *How to Choose the Best Cloud Provider for Your Business in Latin America*. Obtenido de Edge Uno: <https://edgeuno.com/how-to-choose-the-best-cloud-provider-for-your-business-in-latin-america/>
- Gardini, P. (01 de Septiembre de 2025). *24 Best Cloud Service Providers Reviewed for 2025*. Obtenido de The CTO: <https://thectoclub.com/services/best-cloud-service-providers/>
- Martin, B. (12 de Agosto de 2025). *5 Top Cloud Service Providers in 2025 Compared*. Obtenido de Datacamp: <https://www.datacamp.com/blog/top-cloud-service-providers-compared>
- O'Dwyer, S. (01 de Abril de 2025). *Top 9 Cloud Service Providers in 2025*. Obtenido de ProsperOps: <https://www.prosperops.com/blog/top-cloud-providers/>
- Pham, Q. (23 de Octubre de 2024). *Top 5 Big Data Platforms You Should Know in 2025*. Obtenido de Orient: <https://www.orientsoftware.com/blog/big-data-platform/>
- Slingerland, C. (21 de Mayo de 2025). *Top Cloud Service Providers Globally In 2025*. Obtenido de Cloudzero: <https://www.cloudzero.com/blog/cloud-service-providers/>
- Technavio. (01 de Febrero de 2025). *Latin America Cloud Computing Market Analysis - Size and Forecast 2025-2029*. Obtenido de Technavio: <https://www.technavio.com/report/latin-america-cloud-computing-market-analysis>

