

## Maestría en

# ADMINISTRACIÓN DE EMPRESAS

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

**AUTORES:** Kevin Daniel Álvarez López,  
María Verónica Valle Reyes,  
Yael Elizabeth Fierro Guillén,  
Gabriela Natali Romero González,  
Eunice Leonor Valdiviezo Ruiz

**TUTORES:** José Luis Pérez Galán, Ing. Ignacio  
Maroto

**TUTORES:** José Luis Pérez Galán, Ing. Ignacio Maroto

**Estrategias de Análisis del Comportamiento del Cliente y Personalización de Ofertas en el Comercio Electrónico: Caso de Estudio en la Industria Cárnica**

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

Estrategias de Análisis del Comportamiento del Cliente y Personalización de Ofertas en el Comercio Electrónico: Caso de Estudio en la Industria Cárnica

Por

Kevin Daniel Álvarez López, María Verónica Valle Reyes, Yael Elizabeth Fierro Guillén, Gabriela Natali Romero González, Eunice Leonor Valdiviezo Ruiz

Lic. José Luis Pérez Galán  
Lic. Ignacio Maroto

octubre 2025

Aprobado



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

Nosotros, Cristian Melo e Ignacio Maroto, en calidad de responsables, declaramos que, personalmente conocemos que los graduandos: Kevin Daniel Álvarez López, María Verónica Valle Reyes, Yael Elizabeth Fierro Guillén, Gabriela Natali Romero González y Eunice Leonor Valdiviezo Ruiz, son los autores exclusivos de la presente investigación y que ésta es original, auténtica y de su autoría.



---

Cristian Melo  
Coordinador MBA UIDE



---

Ignacio Maroto  
Provost WBS

### Autoría del Trabajo de Titulación

Nosotros, Kevin Daniel Álvarez López, María Verónica Valle Reyes, Yael Elizabeth Fierro Guillén, Gabriela Natali Romero González y Eunice Leonor Valdiviezo Ruiz, declaramos bajo juramento que el trabajo de titulación titulado **"Estrategias de Análisis del Comportamiento del Cliente y Personalización de Ofertas en el Comercio Electrónico: Caso de Estudio en la Industria Cárnica"** 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.



Kevin Daniel Álvarez López  
[kevin163457@gmail.com](mailto:kevin163457@gmail.com)



Yael Elizabeth Fierro Guillén  
[yael\\_david4@hotmail.com](mailto:yael_david4@hotmail.com)



María Verónica Valle Reyes  
[vero\\_14alsu@hotmail.com](mailto:vero_14alsu@hotmail.com)



Gabriela Natali Romero González  
[gaby\\_romerog14@outlook.com](mailto:gaby_romerog14@outlook.com)



Eunice Leonor Valdiviezo Ruiz  
[eunicevaldiviezo@hotmail.com](mailto:eunicevaldiviezo@hotmail.com)

## Autorización de Derechos de Propiedad Intelectual

Nosotros, Kevin Daniel Álvarez López, María Verónica Valle Reyes, Yael Elizabeth Fierro Guillén, Gabriela Natali Romero González y Eunice Leonor Valdiviezo Ruiz, en calidad de autores del trabajo de investigación titulado *"Estrategias de Análisis del Comportamiento del Cliente y Personalización de Ofertas en el Comercio Electrónico: Caso de Estudio en la Industria Cárnica"*, autorizamos a la Universidad Internacional del Ecuador (UIDE) para hacer uso de todos los contenidos que nos pertenecen o que contienen esta obra, con fines estrictamente académicos o de investigación. Los derechos que nos corresponden como autores se ajustan a lo establecido en los artículos 5, 6, 8, 19 y demás pertinentes de la Ley de Propiedad Intelectual y su Reglamento en Ecuador.

Quito, 07 octubre de 2025.



Kevin Daniel Álvarez López  
[kevin163457@gmail.com](mailto:kevin163457@gmail.com)



María Verónica Valle Reyes  
[vero\\_14alsu@hotmail.com](mailto:vero_14alsu@hotmail.com)



Yael Elizabeth Fierro Guillén  
[Yael\\_david4@hotmail.com](mailto:Yael_david4@hotmail.com)



Gabriela Natali Romero González  
[gaby\\_romerog14@outlook.com](mailto:gaby_romerog14@outlook.com)



Eunice Leonor Valdiviezo Ruiz  
[eunicevaldiviezo@hotmail.com](mailto:eunicevaldiviezo@hotmail.com)

## **Dedicatorias y Agradecimientos**

A nuestras familias, que, con su apoyo incondicional, paciencia y motivación, han sido el pilar fundamental para alcanzar este logro académico. Este trabajo es un reflejo del esfuerzo conjunto, la inspiración y la confianza que nos han brindado en cada etapa de nuestra formación. A todos ellos, dedicamos este logro con profundo agradecimiento y gratitud.

Expresamos nuestro más sincero agradecimiento a todas las personas que, de manera directa o indirecta, contribuyeron a la culminación de esta tesis. En primer lugar, queremos agradecer a nuestras familias, por su apoyo, comprensión y estímulo constante, los cuales han sido esenciales en nuestro proceso de formación.

Asimismo, reconocemos la guía, asesoramiento y disposición de nuestros docentes, profesores y asesores académicos, cuyas recomendaciones y conocimientos enriquecieron esta investigación. Además, agradecemos a la empresa Carnivery por facilitar la información necesaria y colaborar en la realización de este trabajo.

Por último, extendemos nuestra gratitud a nuestros compañeros y amigos por su apoyo, entusiasmo y motivación en los momentos que requirieron mayor esfuerzo. Este logro es un resultado colectivo, que refleja el compromiso y la colaboración de todos los que creyeron en nuestro potencial y nos acompañaron en este proceso.

# INDICE DEL DOCUMENTO

## Contenido

INTRODUCCIÓN Y OBJETO DEL PROYECTO .....	11
PARTE I:.....	12
CAPITULO 1: Presentación de la Empresa, Diagnóstico del Problema y Análisis de las Necesidades Informativas .....	12
1.1.    Presentación de la Empresa .....	12
1.2.    Identificación de los Objetivos Iniciales del Proyecto y Definición del Problema.....	12
1.3.    Áreas y Personas Demandantes de Información y Uso Actual de la Mismo ....	13
1.4.    Objetivos de las Áreas, Información Actual y Necesidades Detectadas.....	13
1.5.    Definición de los KPI's Relevantes .....	14
CAPITULO 2: Gestión y Procesamiento de Datos en el Data Warehouse .....	16
2.1.    Origen de la Información y Recurrencia de Carga: Interna y Externa.....	16
2.2.    Sistemas de Destino de la Carga.....	16
2.3.    Almacenamiento en Data Warehouse y Data Marts .....	17
2.4.    Carencias y Problemas en la Carga de Datos .....	18
2.5.    Informes Actuales, Necesidades Detectadas y Brechas.....	19
CAPITULO 3: Procesos ETL .....	20
3.1.    Procesos de Carga de Datos (ETL, Data Lake).....	20
3.2.    Información y Calidad de la Información Disponible .....	20
3.3.    Horarios de Carga y Accesibilidad.....	21
3.4.    Variables Disponibles y Acceso al Data Warehouse .....	22
3.5.    Sistemas Implicados en los Procesos.....	23
3.6.    Recursos de IT o Externos Empleados .....	24
CAPITULO 4: Herramientas de visualización de datos.....	26
4.1.    Evaluación y selección de herramientas de visualización en función del ecosistema de Carnivery.....	26
4.2.    Propuesta de usuarios y perfiles de uso para dashboards.....	27
4.3.    Necesidades de información y accesos por departamentos y perfil .....	28
4.4.    Revisión y cumplimiento de los objetivos iniciales .....	30
CAPITULO 5: Creación del área de BI en una organización .....	33
5.1.    Identificar la situación actual del departamento de Business Intelligence en Carnivery .....	33

5.2.	Definición de las carencias en formación y recursos.....	34
5.3.	Coordinación de los flujos entre otras áreas .....	34
5.4.	Organigrama, dependencias de otras áreas y solución de problemas.....	35
5.5.	Recursos externos destinados: subcontratación .....	37
5.6.	Business Case de Costes y Beneficios .....	37
PARTE II.....		41
CAPITULO 1: Sistemas de Información .....		41
1.1.	Fuentes de información internas y externas.....	41
1.2.	Relación de los datos con la propuesta de BI .....	43
1.3.	Casos de usos y aplicabilidad.....	46
1.4.	Justificación del Impacto del Big Data .....	48
CAPITULO 2: Creación del Área de BI.....		50
2.1.	Analizar las alternativas de diferentes proveedores propuestos .....	50
2.2.	Hacer una selección de nuestra alternativa mejor de proveedor .....	50
2.3.	Definir nuestro modelo de cloud computing: SaaS, PaaS o IaaS .....	51
2.4.	Flujo de la integración de nuestra arquitectura Big Data con el data warehouse o estructura de BI alternativa.....	52
CAPITULO 3: FRAMEWORKS APLICABLES.....		54
3.1.	Selección del Framework Óptimo para Carnivery.....	55
3.2.	Lenguajes de Programación en el Ecosistema Big Data .....	57
3.3.	Tecnologías de Almacenamiento de Datos: Bases Relacionales vs. NoSQL .....	59
3.4.	Integración de Fuentes de Datos y Herramientas de Soporte .....	62
CAPITULO 4 MODELO DE PROPENSIÓN DE COMPRA CON MACHINE LEARNING (NON REAL TIME) .....		67
4.1.	Objetivo .....	67
4.2.	Fuentes de datos utilizadas .....	68
4.3.	Área de aplicación y departamentos de negocio implicados .....	70
4.4.	Identificación de la mejora esperada.....	73
4.5.	Cuantificación de la mejora (Business Case).....	75
CAPITULO 5 MODELO DE RECOMENDACIÓN PERSONALIZADA EN TIEMPO REAL.....		79
5.1.	Objetivo del modelo .....	79
5.2.	Fuentes de datos utilizadas .....	79
5.3.	Área de aplicación y departamentos de negocio implicados .....	81
5.4.	Identificación de la mejora esperada.....	83



5.5. Cuantificación de la mejora (Business Case).....	84
CONCLUSIONES Y APLICACIONES.....	88
REFERENCIAS .....	91

## INDICE DE FIGURAS Y TABLAS

### Índice de Tablas

Tabla 1: Análisis de Criterios Estratégicos para Carnivery .....	26
Tabla 2: Clasificación de Perfiles y Niveles de Acceso .....	28
Tabla 3: Análisis de Costos por Categoría .....	38
Tabla 4: Clasificación de Fuentes de Datos.....	42
Tabla 5: Clasificación de Fuentes y Tipos de Datos en BI.....	43
Tabla 6: Relación entre KPI y Fuentes de Impacto.....	44
Tabla 7: Evaluación de Proveedores y Capacidades Tecnológicas.....	50
Tabla 8: Análisis de Frameworks: Rendimiento y Casos de Uso .....	54

### Índice de Gráficos

Gráfico No. 1: Ejemplo de posibles Dashboards a desarrollar .....	30
Gráfico No. 2: Unidades vendidas por categoría y por mes.....	30
Gráfico No. 3: Coordinación parcial de Flujos entre áreas funcionales .....	35
Gráfico No. 4: Relaciones entre Carnivery y las distintas áreas de Carnevoros.....	36
Gráfico No. 5: Clasificación de Fuentes internas y externas de datos .....	43
Gráfico No. 6: Carnivery IA & Data Framework .....	66

## INTRODUCCIÓN Y OBJETO DEL PROYECTO

En el contexto actual, el comercio electrónico ha experimentado un crecimiento exponencial, consolidándose como un canal imprescindible para la expansión y competitividad de las empresas en múltiples sectores. En particular, la industria cárnica ha comenzado a reconocer la importancia de integrar estrategias digitales que permitan mejorar la experiencia del cliente, incrementar la fidelidad y potenciar las ventas a través de la personalización de ofertas. En este sentido, la implementación de técnicas de análisis del comportamiento del consumidor se revela como una herramienta importante para comprender las preferencias, patrones de consumo y decisiones de compra de los usuarios en plataformas digitales.

Asimismo, es perceptible que la capacidad de ofrecer propuestas ajustadas a las necesidades específicas de cada cliente, mediante la personalización de ofertas, representa una ventaja competitiva significativa. Por ello, en el presente proyecto de investigación, se propone analizar en profundidad el comportamiento del cliente en el ámbito del comercio electrónico, específicamente en la industria cárnica, con el fin de identificar las variables que influyen en su proceso de decisión de compra, así como los mecanismos mediante los cuales dichas variables pueden ser utilizadas para diseñar estrategias de marketing digital más efectivas y direccionadas.

De manera concreta, el objetivo principal de este estudio radica en aplicar técnicas de análisis de datos y modelado del comportamiento del consumidor, con el fin de diseñar propuestas de personalización de ofertas que incrementen la satisfacción del cliente y, en consecuencia, mejoren los resultados económicos de las plataformas de comercio electrónico del sector cárnico. Por otro lado, este trabajo busca proporcionar un marco teórico y aplicado que contribuya tanto a la comunidad académica como a las empresas del sector, en la medida en que facilite la adopción de estrategias basadas en datos y en el conocimiento profundo del cliente, además de promover la innovación en las prácticas de marketing digital.

En síntesis, el presente proyecto se orienta a desarrollar un análisis integral del comportamiento del consumidor en el comercio electrónico de la industria cárnica, con el propósito de optimizar la personalización de ofertas y, en última instancia, potenciar la interacción y fidelización del cliente mediante el uso de herramientas analíticas avanzadas, en un entorno cada vez más competitivo y digitalizado.

## PARTE I:

### CAPITULO 1: Presentación de la Empresa, Diagnóstico del Problema y Análisis de las Necesidades Informativas

#### 1.1. Presentación de la Empresa

La empresa objeto de estudio se denomina *Carnivery*, una organización dedicada a la comercialización de productos cárnicos a través de plataformas de comercio electrónico. Según datos internos y reportes oficiales publicados en su página web (Carnivery, 2024), la compañía ha logrado posicionarse en el mercado local, ofreciendo una amplia variedad de productos de alta calidad, respaldados por estándares de seguridad alimentaria y un servicio de atención al cliente eficiente. Como parte de su estrategia de crecimiento, la empresa ha invertido en el desarrollo de canales digitales para ampliar su alcance, adaptándose a las tendencias globales que indican la creciente preferencia del consumidor por las compras en línea. Es importante mencionar que Carnivery responde a un sector altamente competitivo, donde la innovación y la personalización de ofertas constituyen elementos diferenciadores para la fidelización del cliente y la consolidación de una propuesta de valor única en el mercado.

#### 1.2. Identificación de los Objetivos Iniciales del Proyecto y Definición del Problema

El presente proyecto surge ante la necesidad de evaluar y optimizar las estrategias digitales orientadas a personalizar las ofertas comerciales en el canal de comercio electrónico, con énfasis en el comportamiento del consumidor. En primera instancia, los objetivos iniciales del proyecto se centraron en entender el patrón de consumo de los clientes, identificar las variables que influyen en sus decisiones de compra y determinar cómo utilizar estos datos para desarrollar ofertas personalizadas (Shapiro & Varian, 1999).

El problema principal que se busca abordar radica en la baja tasa de conversión en las plataformas digitales, atribuida, en parte, a la falta de conocimiento profundo acerca del comportamiento del cliente y a la insuficiente utilización de los datos disponibles para personalizar las ofertas. Este escenario genera una oportunidad de mejora significativa, dado que, al aplicar técnicas de análisis de datos, es posible crear estrategias más ajustadas a las preferencias del usuario, aportando valor tanto a la organización como a sus clientes (Davenport, 2014).

### 1.3. Áreas y Personas Demandantes de Información y Uso Actual de la Mismo

En la estructura organizacional de Carnivery, diversas áreas demandan información, ya que esta resulta esencial para la toma de decisiones estratégicas y operativas. En primer lugar, el área de ventas y marketing requiere datos sobre tendencias de consumo, patrones de compra y preferencias del cliente para diseñar campañas promocionales dirigidas (Chaffey & Ellis-Chadwick, 2016). Por otro lado, el área de atención al cliente utiliza información para mejorar la experiencia del usuario, resolver incidencias y fidelizar a los clientes mediante ofertas personalizadas.

Asimismo, el departamento de logística y operaciones accede a ciertos datos de ventas para gestionar inventarios y optimizar los procesos de distribución. Sin embargo, actualmente, la utilización de la información es todavía limitada en términos de análisis predictivos y segmentación avanzada, lo cual representa una oportunidad para potenciar la toma de decisiones basada en datos y mejorar la eficiencia en todas las áreas. Algunos de los desafíos identificados en este contexto incluyen la dispersión de los datos y la falta de una gestión integrada que permita un análisis en tiempo real para responder a las demandas del mercado (Provost & Fawcett, 2013).

### 1.4. Objetivos de las Áreas, Información Actual y Necesidades Detectadas

De acuerdo con los diagnósticos realizados, las áreas sustanciales y sus objetivos, junto con las necesidades detectadas, son las siguientes:

- *Área de ventas y marketing:* Su objetivo principal es incrementar la tasa de conversión y mejorar la personalización de campañas. Actualmente, utiliza reportes de ventas y análisis básicos que no permiten una segmentación avanzada ni la predicción del comportamiento. La necesidad detectada es contar con un sistema de análisis de datos que posibilite identificar segmentos de clientes con mayor potencial y diseñar ofertas específicas para cada uno (Delen et al., 2016).
- *Atención al cliente:* Busca ofrecer soluciones inmediatas y personalizadas para fortalecer la fidelización. La información que dispone se limita a registros históricos de atención y encuestas de satisfacción. La necesidad radica en contar con un sistema que permita anticipar preferencias y resolver solicitudes antes de que ocurran, mediante análisis predictivos.
- *Logística y operaciones:* Pretende mejorar la gestión de inventarios y tiempos de entrega. Actualmente, se basa en datos históricos de ventas, con poca capacidad de predicción a corto plazo. La necesidad consiste en adoptar

técnicas de análisis avanzadas para prever picos de demanda y optimizar recursos logísticos.

### 1.5. Definición de los KPI's Relevantes

Para garantizar una evaluación rigurosa del desempeño de las estrategias propuestas y medir de manera precisa el impacto de las acciones implementadas en el contexto del proyecto de maestría, se han establecido distintos indicadores clave de rendimiento (KPI's). La utilización de estos indicadores permite, en primer lugar, realizar un seguimiento detallado y sistemático, además de apoyar la toma de decisiones fundamentadas en datos objetivos y empíricos. De este modo, los KPI's seleccionados son los siguientes:

- **Tasa de conversión:** Este indicador mide el porcentaje de visitantes que, tras interactuar con la plataforma, realizan una compra efectiva. La relevancia de este KPI radica en que refleja directamente la efectividad de las acciones de personalización y segmentación, dado que una tasa elevada indica que las ofertas y comunicaciones están alineadas con las preferencias del cliente (Chaffey y Ellis-Chadwick, 2016). En el marco del proyecto, la mejora en la tasa de conversión permitirá validar si las estrategias analíticas implementadas fomentan una mayor probabilidad de ingreso y fidelización del cliente, contribuyendo así a alcanzar los objetivos comerciales de Carnivery.
- **Valor promedio del pedido (AOV):** Este KPI, que evalúa cuánto gasta en promedio un cliente en cada transacción, resulta fundamental dado que permite analizar la eficiencia de las acciones de cross-selling y upselling. Incrementar el valor del pedido, por medio de la oferta de productos complementarios ajustados al perfil del usuario, tiene un impacto directo en la rentabilidad y sostenibilidad del negocio (García, 2018). Desde la perspectiva de la tesis, esta métrica es clave para verificar si las técnicas de análisis de datos y personalización aportan a maximizar los ingresos por cliente, lo cual, además, favorece la escalabilidad de las estrategias comerciales.
- **Tasa de retención de clientes:** Este indicador refleja la capacidad de la organización para mantener y fortalecer su base de clientes, siendo un indicador de satisfacción y fidelidad derivados de experiencias personalizadas y de calidad. La importancia de este KPI radica en que, en un entorno competitivo y en constante cambio, la retención de clientes es esencial para reducir costos de adquisición y garantizar ingresos recurrentes (Reichheld y Sasser, 1990). En relación con la investigación, una mayor tasa de retención confirmará que las acciones de personalización generan lealtad, elemento crucial para la sostenibilidad a largo plazo de Carnivery.

En definitiva, la incorporación de estos KPI's en el proceso de monitoreo no solo permite realizar mediciones objetivas y consistentes, sino que también facilita la detección temprana de oportunidades de mejora, en línea con los objetivos de esta tesis de maestría. Adicionalmente, dicha evaluación contribuirá a consolidar una cultura basada en la analítica de datos, promoviendo una gestión dinámica y adaptativa frente a las necesidades del mercado digital.

## **CAPITULO 2: Gestión y Procesamiento de Datos en el Data Warehouse**

### **2.1. Origen de la Información y Recurrencia de Carga: Interna y Externa**

El sistema de gestión de la información en Carnivery se fundamenta en un esquema dual, en el cual se recopilan datos tanto desde fuentes internas como externas, con una frecuencia de actualización que varía desde la operación diaria hasta la extracción en tiempo real. En lo que respecta a las fuentes internas, se obtiene información relevante sobre transacciones diarias, incluyendo ventas por producto y categoría, rotación de inventarios y datos asociados a la base de clientes, que se generan automáticamente a través de los sistemas de gestión empresariales integrados (ERP). La disponibilidad de estos datos en tiempo real y bajo demanda posibilita una trazabilidad exhaustiva de cada transacción, contribuyendo a una toma de decisiones informada y oportuna (García, 2018).

Por otro lado, desde una perspectiva externa, la empresa ha sabido incorporar herramientas como Google Analytics, redes sociales (TikTok, Instagram, Facebook), Zoho PageSense y WhatsApp Business, que facilitan la recopilación de datos sobre el comportamiento del usuario en diferentes canales digitales. Dichas fuentes permiten obtener información acerca de las preferencias del usuario, interacciones, consultas y el ciclo de vida del cliente, integrándose en una arquitectura de inteligencia de negocios (BI) que posibilita análisis en casi tiempo real, con una frecuencia de carga diaria.

No obstante, se ha identificado que, en la actualidad, no existe un sistema de monitoreo automatizado que verifique la disponibilidad o continuidad de estas fuentes de datos. Asimismo, la frecuencia de carga no se encuentra formalmente estandarizada: mientras que la información transaccional se registra en tiempo real, los datos provenientes de redes sociales y otras fuentes externas se cargan por lotes (batch) o según necesidad, lo que podría generar asimetrías temporales en los reportes integrados.

Como señala García (2018), “la integración de fuentes de datos internas y externas fortalece la base analítica de la organización, permitiéndole responder con mayor agilidad a las dinámicas del mercado digital”.

### **2.2. Sistemas de Destino de la Carga**

El ecosistema tecnológico de Carnivery está compuesto por plataformas que automatizan y gestionan los procesos comerciales y operativos. En primer lugar, destaca la utilización de Zoho One, que incluye módulos de CRM y gestión de



inventarios, conectados directamente con la tienda online a través de WooCommerce. Esta integración automatiza funciones críticas, desde la recepción del pedido hasta la entrega, permitiendo un flujo de información eficiente y coherente (Rodríguez, 2019). Además, se emplea Contifico, un sistema ERP que abarca la facturación y la gestión contable, facilitando la consolidación de los procesos administrativos.

Estas plataformas están integradas mediante enlaces que automatizan el ingreso y almacenamiento de datos, aunque, no obstante, todavía existen oportunidades para mejorar la consolidación de los datos, así como la creación de un entorno analítico unificado.

En la actualidad, no se utiliza una herramienta de procesamiento ETL (Extract, Transform, Load) especializada que permita estructurar el flujo de datos desde las distintas fuentes hacia un repositorio analítico. Tampoco se cuenta con un sistema automatizado de detección de errores o generación de alertas en caso de fallos en la carga de datos. Esta carencia representa un riesgo para la integridad de los reportes y evidencia la necesidad de incorporar soluciones que permitan la trazabilidad, auditoría y control de calidad de los datos cargados.

La asociación de estas herramientas permite una gestión eficiente y una respuesta rápida a las demandas del mercado, pero también revela la necesidad de avanzar hacia sistemas que faciliten análisis integrados y en tiempo real. La incorporación de un data warehouse (almacén de datos) formal facilitaría este proceso, centralizando la información y optimizando la generación de informes (García, 2018).

### **2.3. Almacenamiento en Data Warehouse y Data Marts**

Actualmente, Carnivery no mantiene un data warehouse tradicional o formal. En su lugar, utiliza Zoho One como sistema centralizado para almacenar y gestionar los datos operativos, tales como ventas, inventarios y clientes, configurándose como una especie de datamart operativo. Sin embargo, el uso de SharePoint como repositorio de archivos externos y la gestión específica de bases de datos para campañas digitales, redes sociales y feedback de clientes, evidencia una distribución fragmentada de la información (Rodríguez, 2019).

En este entorno, no existe una zona de staging documentada para realizar validaciones o pruebas antes de cargar los datos al sistema principal. Tampoco se ha definido un mecanismo de versionado o de respaldo automático de los datos (backups), lo que limita la capacidad de recuperación en caso de fallos. La retención

de datos históricos y la escalabilidad del sistema, es decir, su capacidad de crecer con nuevas fuentes o volúmenes de información, no están formalizadas ni contempladas en una arquitectura BI robusta.

Este escenario hace evidente la necesidad de consolidar la información en un entorno analítico unificado, mediante la implementación de una pequeña datamart analítica que permita realizar análisis avanzados, integración de indicadores y generación de dashboards interactivos. La utilización de plataformas como Power BI o Google Looker Studio facilitaría la integración de datos provenientes de diferentes fuentes, generando un análisis más profundo y efectivo para la toma de decisiones estratégicas, conforme a las recomendaciones de García (2018).

## **2.4. Carencias y Problemas en la Carga de Datos**

A pesar de la infraestructura tecnológica existente, en Carnivery se detectan varias carencias y problemas relacionados con la carga de datos. En primer lugar, aunque los sistemas están integrados, frecuentemente se presentan desajustes, atribuibles a actualizaciones o errores en las configuraciones de las herramientas, lo cual afecta la calidad y consistencia de los reportes (García, 2018).

Asimismo, existe una limitada capacidad para generar reportes específicos y ajustados a las necesidades reales de las áreas, particularmente en aspectos como la atención al cliente, el seguimiento del feedback y la medición de métricas críticas, tales como porcentaje de recompra, frecuencia de compra y carritos abandonados. La falta de indicadores automatizados y dashboards dinámicos limita la visualización en tiempo real, provocando dependencia de reportes manuales y dificultando decisiones ágiles (Rodríguez, 2019).

Adicionalmente, no existen métricas sistematizadas que evalúen la calidad de los datos. No se mide, por ejemplo, el porcentaje de duplicados, valores nulos o retrasos en la carga, lo cual impide detectar fallos estructurales antes de que afecten los reportes. Tampoco hay alertas automáticas que notifiquen sobre latencias en la ingestión de datos o desconexiones de las fuentes. Esta situación compromete la confiabilidad del sistema de información y limita su capacidad de respuesta en entornos de alta demanda.

El proceso de carga de datos presenta además dificultades para obtener información estructurada sobre la experiencia del cliente, lo que impide una evaluación integral y predictiva. La capacitación en manejo de herramientas analíticas y la optimización de los procesos de integración y limpieza de datos emergen como elementos necesarios para mejorar sustancialmente esta problemática

## **2.5. Informes Actuales, Necesidades Detectadas y Brechas**

En la actualidad, Carnivery genera reportes quincenales sobre inventarios, rotación de productos y ventas por categorías y sucursales. Sin embargo, estos reportes resultan insuficientes para abordar aspectos estratégicos, como la medición de la satisfacción del cliente, tasas de recompra y análisis de comportamiento en carritos de compra, además de no contar con dashboards interactivos que permitan visualizar la información en tiempo real (García, 2018).

A esto se suma la falta de automatización en los procesos de generación de reportes. Los análisis actuales son manuales o semi-manuales, lo que incrementa los tiempos de procesamiento y limita la agilidad organizacional. Tampoco existen dashboards operacionalizados que se mantengan actualizados con información en tiempo real ni alertas que adviertan sobre desviaciones relevantes.

Ante esta realidad, la empresa identifica la necesidad de construir una infraestructura analítica que integre los datos de todas las fuentes, automatice los reportes y genere dashboards dinámicos, con el propósito de facilitar la toma de decisiones y mejorar la capacidad de respuesta ante las variaciones del mercado digital. La implantación de dichas soluciones permitiría no solo cerrar las brechas existentes, sino también potenciar el análisis predictivo y la personalización de ofertas, en línea con las tendencias actuales de la inteligencia de negocios (Rodríguez, 2019).

## **CAPITULO 3: Procesos ETL**

### **3.1. Procesos de Carga de Datos (ETL, Data Lake)**

En la actualidad, Carnivery no dispone de un proceso formalizado de carga de datos mediante herramientas ETL (Extract, Transform, Load) ni de un entorno estructurado como un data lake. Las operaciones de carga de datos se desarrollan de manera directa, sin transformaciones intermedias. Los datos generados a partir de interacciones en el canal físico —como órdenes de venta en el local— son registrados manualmente, mientras que en el canal digital, los datos ingresan automáticamente al sistema Zoho CRM.

Si bien Zoho One ofrece funcionalidades de integración y automatización de flujos de datos, estas herramientas no han sido aún implementadas con fines analíticos. En consecuencia, los datos recopilados desde el e-commerce, el CRM o el sistema de facturación (Contifico) son almacenados sin pasar por procesos previos de limpieza, normalización o validación, lo que implica que actualmente se opera sin un pipeline de datos robusto. Este panorama evidencia una oportunidad para incorporar soluciones ETL que faciliten el diseño de un modelo analítico escalable, así como la consolidación de un repositorio centralizado de información.

Actualmente, el sistema de Carnivery gestiona datos provenientes de diversas fuentes como la página web, el sistema de inventario y plataformas contables, pero no cuenta con un monitoreo automatizado ni una estandarización formal en la frecuencia de carga. Esto implica que la información transaccional se registra en tiempo real, mientras que otros datos se integran por lotes según necesidad, lo que puede generar asimetrías temporales en los reportes. Además, la falta de herramientas avanzadas ETL limita la automatización y centralización de los datos, obligando a realizar descargas manuales y procesamiento adicional para la generación de paneles e informes analíticos. La implementación de procesos ETL robustos y la consolidación en un Data Lake permitirían mejorar la trazabilidad, la calidad y la explotación estratégica de los datos, alineando la gestión de la información con las mejores prácticas en Big Data y facilitando la toma de decisiones basada en datos integrados y confiables. (Guevara, comunicación personal, 15 de julio de 2025, 09:00 h) .

### **3.2. Información y Calidad de la Información Disponible**

“El sistema web impide el avance de registros incompletos y la creación de cuentas duplicadas, mientras que Zoho no permite generar productos con códigos repetidos, lo que ayuda a evitar inconsistencias elementales en el inventario.” (Guevara, comunicación personal, 15 de julio de 2025, 09:00 h) .

La calidad de la información en Carnivery presenta un enfoque limitado a controles estructurales básicos proporcionados por las plataformas digitales. Por ejemplo, el sistema web no permite el avance de registros incompletos ni la creación de cuentas duplicadas. Asimismo, dentro de Zoho, no es posible generar productos con códigos repetidos, lo que evita inconsistencias elementales en inventario.

No obstante, la empresa no cuenta con un proceso sistemático de aseguramiento de la calidad de los datos, ni con herramientas automáticas que permitan validar, corregir o auditar los registros ingresados. No se monitorean indicadores de calidad como el porcentaje de registros nulos, duplicados, desactualizados o erróneos, lo cual representa un riesgo frente a la generación de informes confiables. En este sentido, se identifica como línea prioritaria la implementación de mecanismos de monitoreo y depuración que garanticen la confiabilidad y precisión de la información utilizada para la toma de decisiones estratégicas.

Luego de conversar con los expertos de la empresa se pudo advertir que si bien se dispone de una cantidad considerable de datos provenientes de diferentes sistemas (como el inventario, la página web y plataformas contables), existe una falta de procesos formales para la validación, limpieza y auditoría de la información. Actualmente, los datos se almacenan en sistemas independientes y, aunque se generan informes personalizados sobre ventas, inventario y trazabilidad de productos, no se cuenta con mecanismos automatizados para evaluar la calidad de los datos ni para asegurar su integridad a lo largo del proceso de integración. Esta situación puede derivar en inconsistencias, duplicidades o pérdida de valor analítico. Por ello, es fundamental implementar prácticas de control y aseguramiento de la calidad de la información, como la definición de métricas, la depuración periódica y la integración de herramientas de monitoreo, para fortalecer la base de datos y maximizar el potencial de las soluciones de Big Data en la organización. (Guevara, comunicación personal, 15 de julio de 2025, 09:00 h) .

### **3.3. Horarios de Carga y Accesibilidad**

Los procesos de sincronización de datos entre los sistemas utilizados por Carnivery —especialmente entre Zoho One y WooCommerce— se realizan con una frecuencia de dos veces al día, específicamente en horas de la mañana y la tarde. Esta programación aplica principalmente a los datos de inventario. En el caso de las órdenes de compra y la información del cliente, la sincronización ocurre con cada transacción registrada.

En cuanto al acceso a la información, el CEO de la empresa es quien posee control total sobre los datos almacenados en los sistemas operativos. Por su parte, el personal encargado del punto de venta no interactúa con el sistema de información

más allá de los procesos básicos, debido a la falta de formación especializada. Esta estructura sugiere un enfoque altamente centralizado en la gestión de datos, con acceso restringido o poco distribuido, lo cual limita el aprovechamiento transversal del conocimiento generado. (Guevara, comunicación personal, 15 de julio de 2025, 09:00 h) .

### **3.4. Variables Disponibles y Acceso al Data Warehouse**

Cada cliente que realiza una transacción, ya sea en línea o presencial, queda registrado con un conjunto de variables relevantes, entre las cuales se incluyen: nombre completo, correo electrónico, número de cédula o RUC, dirección de entrega (con posibilidad de múltiples direcciones), teléfono de contacto y datos de facturación.

En el plano comercial, se almacenan variables relacionadas con la compra, tales como: código y nombre del producto adquirido, categoría del producto, precio unitario, cantidad solicitada, cantidad comprometida (producto reservado pero no entregado), sucursal de despacho (matriz o local) y bodega correspondiente. Esta información es crucial para realizar segmentaciones de clientes, análisis de rentabilidad por categoría y proyecciones de demanda.

De manera particular, el CEO enfatizó el valor estratégico de la frecuencia de compra como variable central para la construcción de campañas de fidelización, remarketing y diseño de ofertas personalizadas. No obstante, esta variable no se encuentra actualmente sistematizada a nivel global ni integrada a una plataforma de análisis predictivo, sino que su consulta es puntual y por cliente específico.

A pesar de contar con una buena cantidad de datos transaccionales y de cliente, Carnivery no dispone de un data warehouse formal ni de un inventario de variables documentado (diccionario de datos). Esto dificulta la estandarización, reutilización y análisis masivo de los datos generados.

Finalmente, se considera que en el entorno actual de Carnivery, los horarios de carga de datos presentan una falta de estandarización formal, lo que genera asimetrías temporales en la integración y disponibilidad de la información para análisis. Mientras que los datos transaccionales se registran en tiempo real, otros conjuntos de datos provenientes de sistemas externos o de la web se cargan de manera manual o por lotes, según la necesidad operativa. Esta ausencia de un pipeline automatizado y programado para la ingestión de datos limita la accesibilidad oportuna y homogénea a la información, dificultando la consolidación en un repositorio centralizado como un Data Lake. Para optimizar la arquitectura Big Data, es imprescindible definir ventanas de carga periódicas y mecanismos de automatización que aseguren la sincronización eficiente de las distintas fuentes, permitiendo así la disponibilidad continua y confiable de los datos para los procesos

analíticos y de toma de decisiones. (Guevara, comunicación personal, 15 de julio de 2025, 09:00 h) .

### **3.5. Sistemas Implicados en los Procesos**

El entorno tecnológico de Carnivery está compuesto por un conjunto de plataformas que operan de forma integrada para la gestión de procesos comerciales, logísticos y administrativos. Entre los sistemas utilizados destacan:

- Zoho One, que incluye módulos de CRM y gestión de inventarios;
- WooCommerce, como plataforma de comercio electrónico;
- Contifico, orientado a facturación y contabilidad;
- SharePoint, utilizado para almacenamiento de archivos administrativos;
- Página web transaccional, utilizada como canal directo de ventas y recolección de datos.

La integración de estos sistemas permite automatizar procesos clave del negocio; sin embargo, la empresa aún no cuenta con herramientas específicas para Business Intelligence, procesamiento avanzado de datos o desarrollo de dashboards dinámicos, lo que limita el aprovechamiento analítico del ecosistema tecnológico existente.

En el ecosistema de Carnivery, los procesos de integración y gestión de datos involucran múltiples sistemas interconectados, cada uno con funciones específicas dentro del flujo de información. El sistema principal de inventario y gestión operativa es Zoho Inventory, que actúa como núcleo para la consolidación de datos de productos, ventas y clientes, mientras que la plataforma contable Contífico recibe información sincronizada para el control financiero y la trazabilidad de transacciones. Además, la tienda virtual basada en WooCommerce funciona como fuente primaria de datos de ventas en línea y tráfico de clientes, cuyos registros son migrados mediante integraciones programadas hacia Zoho para su procesamiento y análisis. Sin embargo, la ausencia de una arquitectura de orquestación ETL centralizada y la dependencia de transferencias manuales o semiautomatizadas generan cuellos de botella en la consistencia y disponibilidad de la información. Para alcanzar un entorno Big Data robusto, es fundamental implementar pipelines automatizados que integren de manera eficiente estos sistemas heterogéneos, estandarizando formatos, gestionando la calidad y permitiendo la ingesta continua hacia un Data Lake centralizado, donde los datos puedan ser explotados mediante herramientas analíticas avanzadas y modelos de inteligencia de negocio. (Guevara, comunicación personal, 15 de julio de 2025, 09:00 h) .



### 3.6. Recursos de IT o Externos Empleados

Carnivery no cuenta actualmente con un equipo especializado en tecnologías de la información (TI) ni en analítica avanzada. La gestión digital y de datos está centralizada en el CEO, quien dirige los procesos clave sin personal interno capacitado en herramientas de BI o en la implementación de plataformas como Power BI o Tableau. Además, no se han contratado servicios externos para desarrollar infraestructura analítica o consultoría tecnológica, lo que representa una limitación estructural para avanzar hacia un modelo de gestión basado en datos.

La operación digital de Carnivery se apoya en aplicaciones específicas del ecosistema Zoho, combinadas con herramientas como Contifico y WooCommerce. Esto ha permitido una operación automatizada sin desarrollar software propio ni tener infraestructura local de TI. Sin embargo, la empresa solo utiliza aplicaciones puntuales de Zoho, como CRM, Forms, Social, Survey, Campaigns, PageSense e Inventory, lo que limita las capacidades integradas de automatización y analítica. Al no emplear el conjunto completo de módulos disponibles, se evidencia un uso parcial de la plataforma que, si bien es funcional, restringe el potencial completo de la herramienta (Zoho One, 2025).

Aunque Carnivery no ejecuta un proceso ETL formal, su entorno realiza flujos automatizados entre Zoho CRM, Inventory y WooCommerce, lo que permite la recolección y visualización de datos. Sin embargo, carece de las etapas avanzadas de transformación y limpieza que caracterizan a herramientas ETL especializadas, como Apache Nifi, Talend o Pentaho (Davenport & Harris, 2017). Esto indica que, aunque no maneja volúmenes de Big Data, sí necesita gestionar datos dispersos y se beneficiaría de adoptar principios ETL para garantizar la integridad y valor analítico de la información.

La falta de un departamento formal de IT y de perfiles especializados en ingeniería de datos significa que la gestión de herramientas digitales recae en usuarios avanzados en áreas operativas. Para tareas técnicas, Carnivery recurre a soporte externo, lo que limita su capacidad para escalar analíticamente y desarrollar un entorno de inteligencia de negocios automatizado. Esta dependencia también conlleva riesgos operativos, ya que las configuraciones críticas quedan fuera de su control.

A pesar de haber implementado un ecosistema digital funcional, la falta de un equipo interno de TI y el uso parcial de módulos restringen el potencial analítico de Carnivery. Para avanzar hacia una gestión basada en datos, se recomienda fortalecer las capacidades internas con personal capacitado en BI, ampliar el uso de módulos complementarios de Zoho, implementar procesos básicos de ETL mediante herramientas low-code, y evaluar la creación de dashboards interactivos. Además,



se sugiere formalizar alianzas con proveedores externos bajo contratos de nivel de servicio para asegurar la continuidad técnica. El desarrollo de una cultura empresarial basada en datos no requiere gestionar Big Data, sino integrar y aplicar información de manera estratégica y en tiempo real (Chaffey & Ellis-Chadwick, 2022).

Para terminar, en la arquitectura actual de Carnivery, los recursos de TI empleados para la gestión de datos se limitan principalmente a las capacidades nativas de las plataformas SaaS utilizadas, como Zoho Inventory, Contífico y WooCommerce, sin la incorporación de infraestructura dedicada para procesamiento masivo ni servicios avanzados de integración. La ausencia de clusters de procesamiento distribuido, almacenamiento escalable tipo Data Lake y herramientas especializadas de orquestación ETL restringen la capacidad de escalar y automatizar los flujos de datos conforme a los principios de Big Data.

Además, la dependencia de integraciones externas proporcionadas por los propios proveedores de software, en lugar de soluciones personalizadas o frameworks de código abierto (por ejemplo, Apache Spark, Kafka o Airflow), limita la flexibilidad, el control y la optimización de los pipelines de datos. Para evolucionar hacia una arquitectura Big Data madura, es imprescindible incorporar recursos tecnológicos que soporten la ingestión, procesamiento y análisis de grandes volúmenes de datos en tiempo real, así como la integración de servicios cloud y herramientas de automatización que permitan la gobernanza, seguridad y escalabilidad de la infraestructura de datos. (Guevara, comunicación personal, 15 de julio de 2025, 09:00 h) .

## CAPITULO 4: Herramientas de visualización de datos

### 4.1. Evaluación y selección de herramientas de visualización en función del ecosistema de Carnivery

La elección de una herramienta de visualización de datos en Carnivery debe responder a un principio sustancial manifestado por Camila Guevara: "menos pasos, menos clics y menos carga manual". Guevara, comunicación personal, 17 de julio de 2025, 09:00 h).

Actualmente, el ecosistema tecnológico incluye Zoho (CRM, Surveys, Inventario), WooCommerce (canal de ventas e-commerce) y Contífico (facturación y reportes financieros). Sin embargo, estos sistemas operan de forma fragmentada, lo que impide una visión integrada del cliente y de los indicadores clave del negocio (KPIs).

Para remediarlo, se ha considerado la adopción de herramientas como Power BI o Looker Studio, cuya evaluación debe centrarse en los siguientes criterios:

**TABLA 1: ANÁLISIS DE CRITERIOS ESTRATÉGICOS PARA CARNIVERY**

<b>Criterio estratégico</b>	<b>Descripción aplicada al contexto de Carnivery</b>
Integración	Capacidad de conectarse con Zoho, WooCommerce y Contífico mediante APIs o ETL.
Costo	Preferencia por opciones de bajo costo o gratuitas (Looker Studio) con funcionalidad sólida.
Interfaz intuitiva	Uso sencillo para usuarios no técnicos, especialmente Camila Guevara.
Capacidades de automatización	Dashboards actualizados en tiempo real para minimizar carga operativa.
Flexibilidad de visualización	Segmentación dinámica por cliente, categoría o canal de compra.

En esta línea, se propone una arquitectura de BI que integre los datos de los tres sistemas mediante procesos ETL hacia un Data Warehouse, que alimente dashboards en Power BI o Looker Studio, permitiendo una vista unificada de:

- Ticket promedio por canal, producto y segmento.
- Frecuencia de compra, analizada por ciclos temporales y eventos promocionales.
- Tasa de recompra, con alertas automáticas de clientes inactivos o patrones de abandono.

La visualización deja de ser un simple tablero gráfico para convertirse en un motor de decisiones comerciales automatizadas, donde el comportamiento del consumidor se mide, gestiona y acciona de forma predictiva.

#### **4.2. Propuesta de usuarios y perfiles de uso para dashboards**

Actualmente, la gestión de datos y dashboards recae casi exclusivamente en la CEO, Camila Guevara, quien opera Zoho de forma integral. En el punto de venta físico, solo una persona adicional ingresa órdenes de venta, y en Contífico también se utiliza un único usuario. No existen usuarios externos ni proveedores con acceso a dashboards, y no hay una división clara por departamentos en el uso de la información.

Esto revela una centralización extrema del acceso y uso de dashboards. La estructura actual no contempla perfiles diferenciados ni accesos segmentados, lo que limita la escalabilidad y la colaboración interdepartamental. Para evolucionar, sería recomendable definir perfiles de acceso según funciones (ventas, marketing, administración), permitiendo que cada área visualice los indicadores relevantes sin comprometer la seguridad ni la simplicidad operativa.

Actualmente, el acceso es sumamente restringido:

- Solo una persona (la responsable principal) tiene acceso completo a SOHO y a la información crítica.
- Una colaboradora en el local solo ingresa órdenes de venta.
- El equipo de marketing externo no tiene acceso directo; recibe reportes quincenales generados por la responsable.
- En Contífico, solo la colaboradora del local puede ver facturas, pero no tiene acceso a dashboards operativos.

No existen usuarios externos ni proveedores con acceso directo a dashboards. La responsable centraliza la toma de decisiones y la generación de reportes, la operación es altamente centralizada en la CEO.

La propuesta busca distribuir progresivamente el acceso a otras áreas con dashboards diferenciados y controlados.

Con base en la información recaba, se definen los siguientes perfiles y niveles de acceso:

**TABLA 2: CLASIFICACIÓN DE PERFILES Y NIVELES DE ACCESO**

Perfil	Área	Nivel de acceso propuesto	Tipo de decisiones o uso
<b>CEO (Camila Guevara)</b>	Dirección general	Acceso total y administración del BI	Toma de decisiones estratégicas, análisis de rentabilidad, compras
<b>Operador local</b>	Punto de venta	Acceso restringido solo a órdenes de venta	Registro de pedidos e información operativa básica
<b>Marketing y ventas</b>	Campañas y promociones	Visualización de productos más vendidos, AOV, rotación	Segmentación, promoción, gestión de productos novedosos
<b>Logística y compras</b>	Gestión de inventario	Acceso a rotación de inventario, stock y trazabilidad	Validación de compras, planificación de abastecimiento
<b>Atención al cliente</b>	Servicio y encuestas	Visualización de encuestas, mapas de calor, feedback	Mejora de experiencia y fidelización

#### **4.3. Necesidades de información y accesos por departamentos y perfil**

El manejo de big data permitirá generar información relevante para la toma decisiones informadas por cada una de las áreas y la gestión diaria es así como se evidencia las siguientes necesidades para implementación de big data y dashboards.

##### **4.3.1. Información requerida**

Con el fin de cumplir con las expectativas y necesidades de herramientas útiles para Carnivery es vital contar con la siguiente información:

- Base de datos total de clientes activos dentro del último año
- Base de ventas con información a detalle como:
  - Código y nombre del cliente
  - Producto y categoría del producto vendido
  - Unidades, costo unitario de la venta
  - Fecha de la venta
  - Canal de venta

La información antes mencionada nos permitirá definir los KPIS establecidos al inicio del proyecto y que son de gran relevancia para la toma de decisiones en Carnivery.

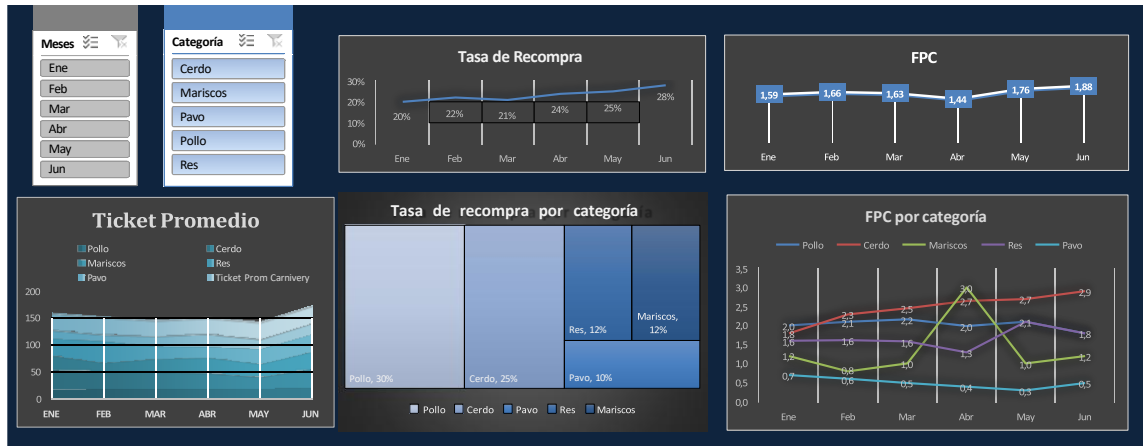
#### 4.3.2. Gerencia y Ventas

En Carnivery las decisiones más importantes son tomadas por su CEO Camila Guevara quien va realizando un análisis continuo de la aceptación en el mercado de los diferentes productos de forma que en coordinación el equipo de abastecimientos y bodega gestionan un inventario saludable que permita ofrecer a los clientes calidad y asegurar una vida útil óptima para cada producto, con el equipo de marketing impulsan campañas para apuntar a un crecimiento y sostenibilidad de algunas categorías de productos.

Es así que se propone para el área de Gerencia y Ventas los siguientes indicadores representados en dashboard interactivos enlazados a bases de datos de los programas existentes en Carnivery.

- **Ticket promedio:** el valor de ventas por ticket por categoría nos permite evaluar la inversión en sostenimiento del inventario por línea y al revisarlo en su evolución notaremos como la venta se marca por temporadas como: Viernes Santo, Día de la madre, padre y Navidad permitiendo realizar el abastecimiento correcto.
- **Tasa de recompra:** este indicador nos permite observar la fidelización de los clientes por categoría de producto y nos brinda la posibilidad de analizar a profundidad el perfil de los productos de menor tasa de recompra.
- **Frecuencia promedio de compra:** este indicador debe ser evaluado de la mano con el ticket promedio ya que un incremento en el valor del ticket promedio puede estar asociado a una reducción de frecuencia promedio de compra impulsado por la garantía de la vida útil de los diferentes productos de forma que las familias realizan compras con menor frecuencia, pero incrementan su ticket promedio.

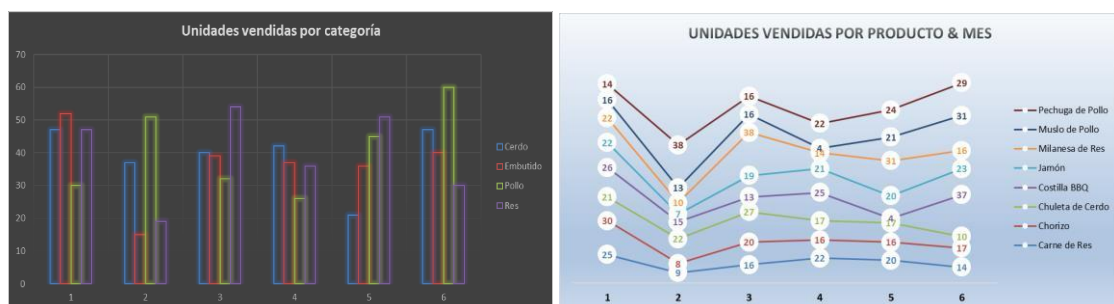
A continuación, se muestra un ejemplo de posibles Dashboards a desarrollar en el programa propuesto Power BI con el fin de entregar información relevante y que se relaciona entre sí para la toma de decisiones.



**GRÁFICO NO. 1: EJEMPLO DE POSIBLES DASHBOARDS A DESARROLLAR**

#### 4.3.3. Departamento de Abastecimientos y Bodega

En el caso del equipo encargado de abastecimiento, control de inventario es importante que mediante visualizadores gráficos tengan información relevante de las categorías y productos más vendidos por mes e incluso por semana para el manejo adecuado del inventario, trazabilidad y la gestión de compras.



**GRÁFICO NO. 2: UNIDADES VENDIDAS POR CATEGORÍA Y POR MES**

Con información confiable y dentro del tiempo optimo el área de abastecimiento tomará decisiones informadas y podrá remitir información valiosa hacia la Gerencia.

#### 4.4. Revisión y cumplimiento de los objetivos iniciales

La etapa de revisión y cumplimiento de los objetivos iniciales constituye un componente crítico dentro de todo proyecto de inteligencia de negocios (BI), ya que permite contrastar las metas planteadas en la fase de diagnóstico con los resultados

efectivamente alcanzados tras la implementación de soluciones tecnológicas. En el caso específico de Carnivery, esta revisión debe orientarse a verificar si las acciones ejecutadas han permitido mejorar la personalización de ofertas, incrementar la tasa de recompra, y optimizar la fidelización del cliente, tal como se estableció en los objetivos del proyecto.

#### **4.4.1. Objetivos iniciales definidos**

De acuerdo con la propuesta inicial de análisis de comportamiento del cliente y personalización de ofertas, los objetivos estratégicos planteados fueron los siguientes:

- Implementar un sistema de análisis de comportamiento de clientes.
- Personalizar campañas de marketing basadas en hábitos de compra.
- Incrementar la tasa de recompra y la frecuencia de compra por cliente.
- Mejorar la fidelización mediante el uso de datos de navegación y compra integrados.
- Optimizar la toma de decisiones con base en dashboards interactivos y KPIs clave.

Estos objetivos responden directamente a la necesidad de Carnivery de transitar de una estrategia genérica de comunicación y ventas a un modelo centrado en datos, en el que la información obtenida del comportamiento del usuario permita anticiparse a sus necesidades y ofrecer productos o promociones relevantes.

#### **4.4.2. Indicadores para medir el cumplimiento**

Para evaluar el grado de cumplimiento de los objetivos, se deben revisar los KPIs establecidos al inicio del proyecto y verificar si han sido medidos, visualizados y utilizados para la toma de decisiones. Entre estos indicadores clave se encuentran:

- Tasa de recompra mensual y por categoría.
- Frecuencia promedio de compra por cliente.
- Ticket promedio por segmento.
- Porcentaje de carritos abandonados.
- ROI por campaña personalizada.
- Tiempo medio entre compras.

La existencia de reportes automatizados o dashboards que muestren la evolución de estos indicadores será evidencia clara de que la solución propuesta ha sido implementada correctamente. En caso contrario, será necesario identificar los cuellos de botella o limitaciones técnicas u organizativas que impiden la medición efectiva de estas métricas.

#### 4.4.3. Evaluación cualitativa de los resultados

Además del análisis cuantitativo de KPIs, es importante considerar una revisión cualitativa de los avances. Por ejemplo:

- ¿Los responsables de marketing están utilizando los datos para segmentar mejor las campañas?
- ¿La gerencia comercial toma decisiones más informadas sobre el portafolio de productos?
- ¿El área de atención al cliente puede anticipar quejas o detectar patrones de comportamiento?

En este sentido, la incorporación de herramientas como Zoho Campaigns, Survey, PageSense y CRM permite monitorear la interacción del cliente con las campañas y recopilar información valiosa para la toma de decisiones. Sin embargo, su uso debe ser medido no solo por su disponibilidad, sino por su efectividad en la operación diaria y su integración dentro del flujo de trabajo de cada área.

#### 4.4.4. Retroalimentación y mejora continua

La revisión de objetivos no debe ser un proceso estático. Debe contemplarse como un ciclo de retroalimentación continua, en el cual los resultados obtenidos retroinforman el diseño de nuevas estrategias. Si bien se hayan cumplido algunos objetivos iniciales, es probable que surjan nuevas necesidades de información, nuevos segmentos de clientes o nuevas oportunidades de automatización.

Este enfoque iterativo es esencial para asegurar que la solución de BI se mantenga relevante, alineada a los cambios en el entorno competitivo, y flexible para adaptarse a futuras fases de expansión o digitalización de la empresa (Davenport & Harris, 2017).

#### 4.4.5. Conclusión

En síntesis, la revisión y cumplimiento de los objetivos iniciales en el caso de Carnivery implica:

- Verificar la disponibilidad y uso de dashboards con KPIs relevantes.
- Contrastar los indicadores actuales con los objetivos propuestos.
- Evaluar el impacto organizacional de la solución implementada.
- Generar recomendaciones de mejora y escalabilidad.

Solo mediante una evaluación crítica y sistemática de estos aspectos será posible determinar el verdadero valor generado por el proyecto, y tomar decisiones informadas sobre su continuidad, mejora o expansión.



## **CAPITULO 5: Creación del área de BI en una organización**

### **5.1. Identificar la situación actual del departamento de Business Intelligence en Carnivery**

Actualmente, el departamento de Business Intelligence en Carnivery se encuentra en una etapa incipiente, con una estructura altamente centralizada y dependiente de una sola persona. Camila Guevara, quien lidera el área, es actualmente la única responsable con conocimientos sólidos en análisis de datos. Esta situación genera una alta dependencia operativa, donde la automatización de procesos y la disponibilidad de reportes estratégicos están condicionadas por su tiempo y capacidad individual.

El soporte que recibe proviene de un asistente sin formación específica en análisis de datos, lo que limita la capacidad de escalar o delegar tareas técnicas. No existe una infraestructura formal de BI —ni a nivel tecnológico (data warehouse, flujos ETL estandarizados), ni a nivel humano (equipo multidisciplinario de analistas)— lo cual impide el desarrollo continuo de analítica avanzada y dashboards actualizados.

Además, el flujo de información entre sistemas como Zoho, WooCommerce y Contífico sigue siendo fragmentado. La generación de reportes aún depende en gran medida de procesos manuales, con poco aprovechamiento de datos para predicciones o segmentaciones automáticas. Este estado genera una brecha importante entre el potencial analítico de la empresa y la realidad operativa del área de BI.

Finalmente, para sustentar el diagnóstico sobre la situación actual del departamento de Business Intelligence en Carnivery, se puede citar textualmente lo que mencionó Camila Guevara durante la entrevista:

“Ahorita todo el tema de BI lo manejo yo sola, y aunque tengo apoyo administrativo, no hay nadie más que realmente sepa de análisis de datos o de automatización, entonces todo depende de mi tiempo y de lo que yo pueda hacer.”<sup>1</sup>

Esta declaración confirma la alta centralización y dependencia operativa en una sola persona, así como la ausencia de un equipo técnico especializado. En conclusión, lo expuesto en mi análisis coincide plenamente con lo expresado por la empresa en la entrevista: la falta de estructura formal y recursos en el área de BI limita el

---

<sup>1</sup> C. Guevara, comunicación personal, 25 de julio de 2025.

desarrollo analítico y la capacidad de escalar procesos, lo que representa un reto importante para el crecimiento y la eficiencia de Carnivery.

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

Las carencias actuales en el departamento de BI se dividen en dos grandes bloques: formación técnica y recursos tecnológicos.

- **Carencias en formación:**

La ausencia de competencias analíticas en el personal de apoyo (asistente del área) implica una falta crítica en el desarrollo de capacidades internas. No se cuenta con habilidades en lenguajes de análisis como SQL, herramientas de visualización como Power BI o Looker Studio, ni en metodologías analíticas como RFM o modelos de CLV. Esta brecha no solo reduce la autonomía del equipo, sino que limita la innovación en el análisis de datos. Camila ha expresado directamente la necesidad de capacitaciones externas para fortalecer al equipo, lo cual es coherente con los objetivos del proyecto de personalización de ofertas y análisis del comportamiento del cliente.

- **Carencias en recursos:**

En términos de infraestructura, la empresa aún no ha implementado un entorno de datos consolidado. Los procesos de integración son mínimos y no se cuenta con un sistema de almacenamiento estructurado tipo data warehouse. Las conexiones entre Zoho, WooCommerce y Contífico no están automatizadas, lo que restringe la trazabilidad de KPIs como la frecuencia de compra, el ticket promedio y la tasa de recompra. Asimismo, no existe una política clara de mantenimiento de dashboards, ni se ha asignado un presupuesto fijo para herramientas de BI, lo que agrava el rezago tecnológico.

## **5.3. Coordinación de los flujos entre otras áreas**

La coordinación entre el área de BI y otras áreas funcionales es parcial y limitada. Actualmente, el equipo de BI trabaja en conjunto con áreas como logística y contabilidad, pero esta colaboración no está institucionalizada ni estandarizada mediante flujos de datos automatizados. La CEO confirmó que, aunque hay un manejo conjunto de procesos, no existen contratos formales con proveedores externos, ya que la mayoría de los procesos logísticos se realizan de forma interna.

Esto representa tanto una fortaleza como una limitación. Por un lado, tener el control logístico interno facilita la captura directa de datos operativos valiosos. Sin embargo, al no existir una estructura de intercambio fluido y automatizado entre áreas, los datos no fluyen hacia el sistema de BI de forma eficiente. Esto afecta la visibilidad global de indicadores estratégicos, como la rotación de productos por

canal, la estacionalidad en las compras o la relación entre promociones y ticket promedio.



**GRÁFICO NO. 3: COORDINACIÓN PARCIAL DE FLUJOS ENTRE ÁREAS FUNCIONALES**

Además, no hay claridad en los responsables de cada flujo de datos, ni en los protocolos de integración entre sistemas. La falta de procesos estandarizados impide que la información llegue de manera consistente al área de BI, afectando la calidad del análisis y la oportunidad de la toma de decisiones.

“Nos apoyamos mucho entre logística y contabilidad, pero todo es muy manual; no tenemos todavía un sistema donde la información pasa sola de un área a otra, entonces a veces se nos pierden datos o toca rehacer reportes.”<sup>2</sup>

Esta declaración evidencia la falta de automatización y la dependencia de procesos manuales para el intercambio de información entre áreas. En conclusión, aunque existe colaboración entre BI y otras áreas, la ausencia de flujos automatizados limita la eficiencia y la visibilidad integral de los datos estratégicos, lo que puede afectar la toma de decisiones en Carnivery.

#### **5.4. Organigrama, dependencias de otras áreas y solución de problemas.**

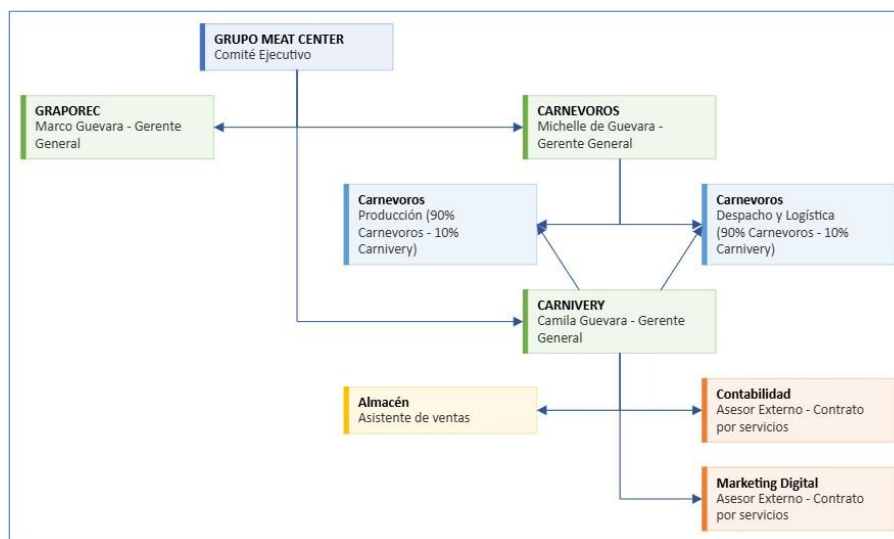
La empresa Carnivery forma parte del Grupo Empresarial Meat Center, un conjunto de tres sociedades que desarrollan toda la cadena de valor de productos cárnicos:

- Graporec se encarga del engorde de proteína animal para consumo humano.

<sup>2</sup> Ibidem, entrevista de 25 de julio de 2025.

- Carnevoros procesa esa proteína a nivel industrial y distribuye a clientes como hoteles, restaurantes y caterings.
- Carnivery, finalmente, comercializa y vende estos productos a través de su tienda online y un local físico.

Al operar como proveedores entre sí, estas compañías se consideran entidades relacionadas ante los organismos de control. Por ello, en el organigrama siguiente se muestran las relaciones entre Carnivery y las distintas áreas de Carnevoros encargadas de la producción y distribución.



**GRÁFICO NO. 4: RELACIONES ENTRE CARNIVERY Y LAS DISTINTAS ÁREAS DE CARNEVOROS**

### Relación operativa

Carnivery mantiene una vinculación estrecha con Carnevoros, particularmente con las áreas de producción. La Gerente General de Carnivery analiza las tendencias de ventas y determina los pedidos correspondientes, que luego se gestionan mediante la logística de Carnevoros para la entrega al cliente final. Gracias a este esquema, Carnivery utiliza los recursos de almacenamiento y transporte de Carnevoros, optimizando así la capacidad instalada de ambas empresas.

Actualmente, Carnivery emplea el 10 % del proceso de despacho y logística de Carnevoros.

### Ventajas del modelo de negocio

El modelo de Meat Center, en el que Carnevoros brinda apoyo operativo a Carnivery, le ha permitido a esta última consolidar sus actividades sin incurrir en la inversión que implicaría desarrollar su propia infraestructura de almacenamiento o

transporte para manejar mayores volúmenes y múltiples destinos. Lo esencial en este modelo es que Carnivery reconozca y contabilice como parte de sus resultados los costos de despacho y logística que Carnevoros le ofrece como servicio.

### 5.5. Recursos externos destinados: subcontratación

Carnivery contrata ciertos servicios externos especializados con el objetivo de optimizar su operación. Los principales recursos subcontratados son:

- **Contabilidad y Tributación:** Se cuenta con un servicio externo para la gestión contable y el cumplimiento de obligaciones tributarias. Este servicio es prestado por un profesional contratado a través del Grupo Empresarial.
- **Marketing y Soluciones Digitales:** Las actividades promocionales, estrategias de marketing y soluciones digitales están a cargo de la empresa Makinando Marketing. Esta compañía es responsable de la implementación, capacitación y desarrollo de plataformas como Zoho y WooCommerce para Carnivery.

Estos servicios forman parte de una línea de soporte no permanente en la operación diaria de la empresa, por lo que, respaldados mediante contratos de prestación de servicios, se establece una relación externa válida, eficiente y segura.

### 5.6. Business Case de Costes y Beneficios

El desarrollo de un sistema de Business Intelligence (BI) sólido y escalable en Carnivery representa una necesidad crítica, dada la situación actual del área, caracterizada por alta dependencia operativa de una sola persona, carencias formativas en el personal de apoyo, falta de automatización en flujos de datos, y ausencia de una infraestructura tecnológica de análisis. Por tanto, plantear un caso de negocio bien fundamentado permite justificar la inversión necesaria para cerrar estas brechas y potenciar el rendimiento empresarial mediante el uso estratégico de la información.

- **Situación actual: punto de partida**

Actualmente, el departamento de BI está liderado únicamente por Camila Guevara, quien asume la totalidad de las tareas de análisis y generación de reportes. El resto del equipo carece de formación técnica, lo cual limita la capacidad de escalar procesos o automatizar tareas rutinarias. No existe un data warehouse, ni se emplean herramientas avanzadas como Power BI o lenguajes de análisis como SQL. Además, los sistemas de origen de datos (Zoho, WooCommerce y Contífico) no están integrados automáticamente, lo que fragmenta la información y reduce su valor estratégico.

- **Costes estimados del proyecto**

La implementación de una solución de BI integral implicará una inversión en tres componentes principales: talento humano, tecnología y capacitación. A continuación, se presenta una estimación preliminar:

**TABLA 3: ANÁLISIS DE COSTOS POR CATEGORÍA**

Categoría	Detalle	Costo estimado (USD)
<b>1. Personal</b>	Contratación de un analista de datos junior con conocimientos en SQL y BI	\$15,000 anuales
	Soporte externo para implementación (consultoría de integración Zoho-WooCommerce-Contífico)	\$5,000
<b>2. Tecnología</b>	Suscripción Power BI Pro (10 usuarios)	\$1,200 anuales
	Implementación de un data warehouse (Google BigQuery, AWS Redshift o Zoho Analytics)	\$3,000
<b>3. Capacitación</b>	Cursos técnicos para el equipo (SQL, Power BI, RFM, CLV)	\$2,000
<b>Total estimado</b>		<b>\$26,200 USD/año</b>

**Nota.** Los valores son aproximados y pueden variar según proveedores, escalabilidad o alianzas estratégicas.

- **Beneficios esperados:**

La implementación del sistema de BI permitirá a Carnivery mejorar significativamente su capacidad analítica, lo cual se traducirá en beneficios directos e indirectos. A continuación, se detallan los más relevantes:

- **Incremento de la tasa de recompra:** La segmentación avanzada y las campañas personalizadas permitirán diseñar ofertas relevantes para cada tipo de cliente. De lograrse un incremento del 10 % en la tasa de recompra, se estima un aumento proporcional en ingresos.
- **Reducción del churn (clientes inactivos):** Con la aplicación de modelos RFM y la automatización de alertas, se podrá reactivar a clientes dormidos. Una recuperación del 5 % del churn mensual representa una fuente adicional de ventas sin adquirir nuevos clientes.
- **Aumento del ticket promedio:** Al cruzar datos entre productos, categorías y promociones, se podrán diseñar paquetes estratégicos para incrementar el ticket promedio por cliente. Se proyecta una subida de entre \$2 y \$4 por transacción, con impacto directo en la rentabilidad.
- **Optimización del tiempo operativo:** La automatización de reportes permitirá liberar tiempo de Camila Guevara y reducir la dependencia operativa. Actualmente, se estima que dedica el 70 % de su tiempo a tareas repetitivas. Al reducir esta carga en un 50 %, se potenciará su capacidad estratégica.

- **Toma de decisiones basada en datos reales y actualizados:** Los dashboards interactivos ofrecerán a las áreas comercial, marketing y logística una visibilidad en tiempo real de indicadores clave, como la rotación de inventario, la conversión por campaña y la evolución de la frecuencia de compra.
- **Mejora del posicionamiento competitivo:** Contar con una solución de BI consolidada permitirá a Carnivery anticiparse a los cambios del mercado, diseñar estrategias de fidelización más eficaces, y proyectarse como una empresa innovadora en el segmento cárnico digital de Ecuador.

- **Retorno de inversión (ROI)**

Aunque el cálculo exacto del ROI dependerá de variables futuras, puede estimarse un escenario conservador:

- Incremento proyectado en ingresos por fidelización y ventas adicionales: \$50,000/año
- Costes de implementación y mantenimiento del sistema de BI: \$26,200/año
- ROI estimado en primer año: ~91 %

En escenarios optimistas, este retorno puede superar el 150 % en un plazo de 12 a 18 meses.

## Conclusión

El business case para el desarrollo de un sistema de BI en Carnivery es sólido tanto en términos técnicos como financieros. La inversión, aunque moderada, permitirá resolver cuellos de botella actuales, reducir la dependencia operativa, mejorar la fidelización del cliente, y convertir los datos en una verdadera ventaja competitiva. A medida que el sistema se consolide y el equipo adquiera madurez analítica, la empresa podrá avanzar hacia modelos de analítica predictiva y automatización de decisiones, posicionándose como líder en su sector. La entrevista refuerza la urgencia y relevancia de invertir en un sistema de Business Intelligence en Carnivery. Como señaló Camila Guevara:

“Si tuviéramos un sistema más automatizado y un equipo que entienda de análisis, podríamos tomar decisiones mucho más rápido y aprovechar mejor la información que ya tenemos, pero ahorita todo recae en mí y en procesos manuales.”<sup>3</sup>

Esta afirmación evidencia cómo la dependencia operativa y la falta de automatización limitan el potencial de la empresa. En conclusión, lo expuesto en el caso de negocio se alinea plenamente con la experiencia y percepción interna: la inversión en BI no solo resolverá los cuellos de botella actuales, sino que permitirá

---

<sup>3</sup> Ibid. Entrevista de 25 de julio de 2025.

a Carnivery transformar sus datos en una ventaja competitiva sostenible y escalar su crecimiento de manera eficiente.



## PARTE II:

### CAPITULO 1: Sistemas de Información

#### 1.1. Fuentes de información internas y externas

Para la elaboración de algunos análisis e indicadores para el desempeño comercial de Carnivery es importante recolectar información tanto de fuentes internas y externas como son:

- **Fuentes Internas**

La información histórica que se mantiene en la compañía dentro de los aplicativos disponibles nos permitirá definir tendencias de ventas por temporada, frecuencia promedio de compra, tasa de recompra, ticket promedio, entre otros. A continuación detallo los aplicativos de donde se obtendrá estas fuentes.

Zoho: el sistema utilizado dentro de la compañía nos permitirá conocer datos históricos de ventas y clientes captados.

- Base de datos total de clientes activos dentro del último año
- Base de ventas con información a detalle

Woocommerce: dentro de la tienda online es importante recolectar información acerca de:

- Tasa de abandono del carrito: El porcentaje de usuarios que inician un proceso de compra pero no lo finalizan, una métrica clave para la optimización de los embudos.
- Usuarios nuevos vs. usuarios regulares: Ayuda a entender si la tienda atrae nuevas audiencias y si los clientes regresan y en que zonas están ubicados, importante que la página solicite como dato obligatorio dirección de domicilio.
- Datos demográficos de consumidores y tendencias de consumo: Estos datos demográficos permiten segmentar y comprender mejor al público objetivo para personalizar las estrategias.
- Mapa de calor: un análisis de los productos que son revisados a detalle ingresando a la fotografía y descripción de dicha técnica nos permite entender que productos están atrayendo más al consumidor y disparar campañas de marketing y promociones sobre los mismos.

Facebook: información relevante acerca de comentarios, likes en la página sobre productos y recolectar resultados de trivias realizadas en redes sociales acerca de las preferencias del consumidor.

- **Fuentes Externas**


Dentro del Ecuador existe información referencial de datos demográficos y tendencias de precios de algunas proteínas como son:

- INEC: datos de la población por edad y género para definir con ello el público objetivo.
- MAGAP: información semanal de precios de cerdo, pollo pie que pueden servir de referencia para entender el precio en el mercado de los productos procesados y nos permita reconfigurar los volúmenes de venta en mercados sobre ofertados y también el ajuste en compra de materia prima y aplicación de promociones para no perder participación en el mercado.
- Compra de base de datos: En Ecuador es permitido la compra de bases de datos siempre que las mismas cumplan lo establecido en el artículo 66, numeral 19 de la Constitución de la República, que establece “el derecho a la protección de datos de carácter personal, que incluye el acceso y la decisión sobre información y datos de este carácter, así como su correspondiente protección. La recolección, archivo, procesamiento, distribución o difusión de estos datos o información requerirán la autorización del titular o el mandato de la ley”. Esta opción nos permitiría recolectar datos acerca de personas en un sector específico y el nivel socioeconómico para lanzar nuestras campañas de socialización digitales.

A continuación, se resumen las fuentes internas y externas identificadas y su clasificación:

**TABLA 4: CLASIFICACIÓN DE FUENTES DE DATOS**

Fuentes	Origen	Datos	Estructurados	No Estructurados	Semi estructurados
Internas	Zoho	Base de datos total de clientes activos dentro del último año	X		
		Base de ventas con información a detalle	X		
	Woocommerce	Tasa de abandono del carrito			X
		Usuarios nuevos vs. usuarios regulares			X
		Datos demográficos de consumidores y tendencias de consumo	X		
		Mapa de calor			X
	Facebook	Interacciones y comentarios	X	X	X
Externas	INEC	Información demográfica		X	
	MAGAP	Precios referenciales de animales en pie		X	
	Proveedor externo	Marketing de base de datos (compra)	X		

<b>Reporte:</b>	Precios en Agroindustria
<b>Descripción:</b>	Información referencial de precios de productos agropecuarios han sido transformados bajo algún proceso agroindustrial.
<b>Período de datos:</b>	Desde Noviembre 2011
Seleccione el rango de fechas entre las cuales desea consultar los precios	
Desde:	01/09/2025
Hasta:	09/09/2025
Provincia:	--TODAS--
Seleccione la localidad de donde desea obtener los datos	
<input type="radio"/> Balanceados <input type="radio"/> Centros de Acopio Grano <input type="radio"/> Centros de Acopio Cacao <input type="radio"/> Centros de Acopio Café <input type="radio"/> Centros de Acopio e Industrias Artesanales de Lácteos	<input checked="" type="radio"/> Centros de Faenamiento de Pollo <input type="radio"/> Extractoras de Aceite <input type="radio"/> Ingenios Azucareros <input type="radio"/> Molinos Empresariales
<input type="radio"/> Industrias Lácteas <input type="radio"/> Piladoras <input type="radio"/> Planteles Avícolas <input type="radio"/> Productores de leche	
 <b>REPORTE</b>	

**GRÁFICO NO. 5: CLASIFICACIÓN DE FUENTES INTERNAS Y EXTERNAS DE DATOS**

## 1.2. Relación de los datos con la propuesta de BI

### 1.2.1. Vinculación de las fuentes de datos con el modelo de BI planteado

La propuesta de BI de *Carnivery* se fundamenta en integrar múltiples fuentes de datos —tanto internas como externas— para alimentar un sistema inteligente de análisis que permita mejorar la toma de decisiones en tiempo real. Cada fuente aporta datos estructurados, no estructurados o semi-estructurados que se procesan mediante técnicas de Big Data para generar valor en áreas clave como logística, marketing, ventas y atención al cliente.

**TABLA 5: CLASIFICACIÓN DE FUENTES Y TIPOS DE DATOS EN BI**

Fuente	Origen	Tipo de dato	Utilidad dentro del modelo BI
<b>Zoho CRM</b>	Interna	Estructurado	Segmentación de clientes, análisis de comportamiento de compra, historial detallado de ventas.
<b>WooCommerce</b>	Interna	Semi-estructurado	Seguimiento de carritos abandonados, nuevos vs. recurrentes, tendencias de navegación y mapas de calor.
<b>Facebook</b>	Interna/Externa	Mixto	Sentiment analysis, campañas personalizadas, comportamiento en redes sociales.

<b>INEC</b>	Externa	Estructurado	Enriquecimiento del perfil demográfico de clientes para segmentaciones más precisas.
<b>MAGAP</b>	Externa	Estructurado	Fluctuación de precios de animales en pie, planificación de compras y ajustes de precios.
<b>Base de datos comprada (proveedor externo)</b>	Externa	Estructurado	Captación de nuevos leads, diseño de audiencias lookalike, expansión de mercado.

Estas fuentes permiten un flujo continuo de datos hacia el modelo de BI, que a su vez se soporta en tres capas: adquisición, análisis y visualización. Este flujo habilita cuadros de mando en tiempo real que fortalecen la capacidad de Carnivery para actuar con rapidez y precisión ante cualquier cambio operativo o comercial.

### 1.2.2. Relación de las fuentes con los KPIs del proyecto

Los tres KPIs estratégicos definidos para Carnivery son: ticket promedio, frecuencia de compra y tasa de recompra. Las fuentes de datos alimentan directa o indirectamente a cada uno:

**TABLA 6: RELACIÓN ENTRE KPI Y FUENTES DE IMPACTO**

<b>KPI</b>	<b>Fuentes relacionadas</b>	<b>Forma en que se impacta</b>
<b>Ticket promedio</b>	Zoho (ventas), WooCommerce (productos comprados), MAGAP	Identifica patrones de gasto, precios óptimos, bundles más rentables.
<b>Frecuencia de compra</b>	WooCommerce (histórico por usuario), Facebook (engagement), Zoho	Define campañas de retargeting, recordatorios y suscripciones automáticas.
<b>Tasa de recompra</b>	WooCommerce, Facebook (feedback), Zoho (segmentación)	Detecta clientes fieles, mide satisfacción postventa, activa promociones personalizadas.

El uso de modelos predictivos basados en estos datos permite anticipar comportamientos de compra, optimizar catálogos y realizar promociones cruzadas inteligentes, lo que mejora los KPIs de manera sostenible.

### 1.2.3. Alimentación de dashboards y apoyo a la toma de decisiones

El sistema de BI transforma los datos en información visual y accionable mediante dashboards diseñados por área funcional:

#### a) Dashboard de Ventas y Marketing

- Métricas clave: ticket promedio, conversión por campaña, clics vs. compras.
- Fuentes principales: Zoho, WooCommerce, Facebook.
- Aplicación: campañas personalizadas, diseño de catálogos dinámicos, análisis de tendencias de consumo por segmento.

#### b) Dashboard de Logística

- Métricas clave: entregas a tiempo, costos de transporte, stock crítico.
- Fuentes principales: WooCommerce, MAGAP.
- Aplicación: rutas optimizadas, control de inventario predictivo, reducción de quiebres.

#### c) Dashboard de Atención al Cliente

- Métricas clave: satisfacción del cliente, tiempos de respuesta, NPS.
- Fuentes principales: Facebook (comentarios), Zoho (historial de tickets).
- Aplicación: mejoras del servicio, respuestas automatizadas, gestión proactiva de quejas.

#### d) Dashboard Financiero

- Métricas clave: rentabilidad por cliente, margen por producto, costo por adquisición.
- Fuentes principales: Zoho, base de datos externa.
- Aplicación: evaluación de ROI de campañas, ajustes en precios y políticas de fidelización.

### Conclusión

La integración inteligente de fuentes de datos con el modelo de BI de Carnivery permite un monitoreo dinámico y en tiempo real de los indicadores clave del negocio. Cada dato capturado sea estructurado, semi o no estructurado contribuye a generar inteligencia de negocio que mejora la rentabilidad, la relación con el cliente y la eficiencia operativa. La personalización de dashboards por área facilita

la toma de decisiones informadas y ágiles, consolidando a Carnivery como una empresa orientada al dato y con visión estratégica.

### 1.3. Casos de usos y aplicabilidad

A partir de las fuentes internas (Zoho, WooCommerce, Facebook) y externas (INEC, MAGAP, compras de bases de datos) identificadas por el equipo, se diseñan casos de uso que combinan datos estructurados, semiestructurados y no estructurados para mejorar los indicadores clave de Carnivery: ticket promedio, frecuencia de compra y tasa de recompra. La siguiente sección sustituye la tabla por una descripción detallada en formato de listas.

#### 1.3.1. Segmentación RFM y campañas de recompra

##### Datos utilizados:

- Historial de ventas y clientes de Zoho (recencia, frecuencia y valor de las compras).
- Métricas de comportamiento de WooCommerce, como tasa de abandono del carrito y distinción entre usuarios nuevos y recurrentes.
- Comentarios e interacciones en Facebook.
- Variables demográficas del INEC para enriquecer la segmentación.

**Problema o necesidad:** identificar grupos de clientes para campañas personalizadas que incentiven la recompra. El enfoque RFM permite clasificar a los clientes según frecuencia y valor de las compras, lo que ayuda a detectar clientes leales y aquellos en riesgo de abandono

**Indicador que mejora:** la segmentación RFM aumenta la tasa de recompra y la frecuencia de compra al focalizar esfuerzos en clientes de alto valor y diseñar campañas de retención más efectivas

#### 1.3.2. Análisis del abandono del carrito y recuperación de ventas

##### Datos utilizados:

- Registros de abandono del carrito y pasos del embudo de compra en WooCommerce.
- Diferenciación de usuarios nuevos y recurrentes.
- Feedback en redes sociales y bases externas para comprender motivos de abandono.

**Problema o necesidad:** reducir el porcentaje de usuarios que inician el proceso de compra, pero no lo finalizan. La analítica de big data permite detectar puntos de fricción y automatizar mensajes o incentivos para retomar la compra

**Indicador que mejora:** al optimizar el embudo de ventas y ofrecer incentivos personalizados, se incrementa el ticket promedio y la tasa de conversión en el carrito.

### 1.3.3. Predicción de demanda e inventarios

#### Datos utilizados:

- Historial de ventas y clientes de Zoho.
- Tendencias de consumo y mapas de calor de WooCommerce.
- Datos de precios de MAGAP y variables demográficas de INEC.

**Problema o necesidad:** anticipar la demanda para ajustar la planificación de compras y evitar sobrestock o roturas de inventario. Las técnicas de big data permiten combinar datos históricos con variables externas para generar modelos de predicción

**Indicador que mejora:** mejora la rotación de inventario y el ticket promedio, ya que garantiza la disponibilidad de los productos adecuados al momento correcto.

### 1.3.4. Optimización de logística y experiencia de entrega

#### Datos utilizados:

- Información de pedidos, direcciones de entrega y tiempos de procesamiento de WooCommerce
- Datos de ubicación en tiempo real (IoT/GPS) y condiciones de tráfico.
- Feedback de clientes en redes sociales.

**Problema o necesidad:** mejorar la eficiencia en rutas de entrega y reducir los tiempos de distribución. El análisis en tiempo real de datos de tráfico y condiciones climáticas permite optimizar las rutas, reducir costos logísticos y ofrecer visibilidad al cliente

**Indicador que mejora:** aumenta la frecuencia de compra y la satisfacción del cliente, ya que una mejor experiencia de entrega fomenta la recompra.

### 1.3.5. Personalización de ofertas y cross-selling

**Datos utilizados:**

- Historial de compras y comportamiento de navegación (Zoho y WooCommerce).
- Segmentación RFM y datos demográficos.
- Interacciones y preferencias recogidas en redes sociales.

**Problema o necesidad:** ofrecer recomendaciones de productos complementarios o promociones personalizadas que incrementen el valor de la compra. Big data permite generar recomendaciones basadas en patrones de consumo y perfiles de cliente

**Indicador que mejora:** incrementa el ticket promedio y la tasa de recompra al ofrecer productos relevantes para cada cliente

#### 1.4. Justificación del Impacto del Big Data

El Big Data transforma el proyecto de Carnivery al permitir la captura, almacenamiento y análisis de volúmenes masivos de datos de diversas fuentes a alta velocidad (las “4 Vs”: volumen, velocidad, variedad y veracidad). La literatura señala que el uso de Big Data influye en tres grandes áreas de decisión: eficiencia operativa, personalización y predicción. Su incorporación tendrá repercusiones directas en los procesos de negocio y en los departamentos clave.

- **Impacto en los procesos de negocio**

**Optimización de operaciones:** El análisis de datos permite automatizar tareas repetitivas e identificar cuellos de botella. En logística, Big Data habilita la optimización de rutas, la gestión eficiente del inventario y la reducción de costos de transporte. En Carnivery, esto se traduce en entregas más rápidas y menores gastos logísticos.

**Agilidad y precisión en la toma de decisiones:** Los sistemas de Big Data proporcionan acceso en tiempo real a información crítica, lo que mejora la velocidad y la precisión de las decisiones. Esto reduce la dependencia de intuiciones y permite responder a tendencias del mercado con rapidez, evitando análisis con datos desactualizados.

**Predicción y gestión del riesgo:** Modelos predictivos basados en datos históricos y variables externas permiten anticipar la demanda, ajustar compras y evitar sobrestock o quiebres. Asimismo, ayudan a detectar posibles riesgos o fraudes en las operaciones.



Sostenibilidad y eficiencia energética: El uso de Big Data en logística posibilita rutas más eficientes que reducen consumo de combustible y huella de carbono, aportando a los objetivos de sostenibilidad.

- **Impacto en departamentos clave**

Departamento comercial y marketing: El Big Data sustenta estrategias de personalización y segmentación. Permite identificar clientes de alto valor y diseñar campañas específicas, optimizando el presupuesto de marketing y aumentando la efectividad de promociones. La disponibilidad de datos en tiempo real facilita un marketing ágil basado en comportamiento y tendencias.

Departamento de logística: La integración de datos de rutas, inventarios y demanda optimiza la cadena de suministro. Las empresas logísticas que utilizan Big Data obtienen visibilidad en tiempo real, predicen la demanda y reducen costos operativos.

Atención al cliente: El análisis de comentarios y feedback de redes sociales permite detectar áreas de mejora y ofrecer respuestas personalizadas. Los sistemas de Big Data mejoran la calidad del servicio al anticipar necesidades y gestionar quejas de forma proactiva, lo que impacta positivamente la satisfacción y la lealtad del cliente

- **Beneficios para la empresa**

Incorporar Big Data en el modelo de BI de Carnivery otorga eficiencia, personalización, predicción y agilidad operativa. Se mejora la competitividad al optimizar procesos, reducir costos y ofrecer experiencias personalizadas que fortalecen la relación con el cliente. Además, la capacidad de anticipar tendencias de consumo y demandas emergentes brinda una ventaja estratégica y mejora la rentabilidad.

## **Conclusión**

Los casos de uso propuestos demuestran cómo las fuentes de datos existentes pueden transformarse en iniciativas de alto impacto mediante técnicas de Big Data. Estas iniciativas no solo mejoran los KPIs actuales (ticket promedio, frecuencia de compra y tasa de recompra), sino que también preparan a Carnivery para un entorno digital competitivo donde la información en tiempo real y la personalización son clave. La inversión en infraestructura y talento para Big Data es esencial para implementar estos casos de uso y maximizar los beneficios de la analítica avanzada.

## CAPITULO 2: Creación del Área de BI

### 2.1. Analizar las alternativas de diferentes proveedores propuestos

El ecosistema de Big Data evoluciona hacia plataformas que combinan almacenamiento, procesamiento y analítica avanzada. De los proveedores analizados, se han considerado Google BigQuery, Amazon Redshift, Azure Synapse Analytics e IBM Cloud Pak for Data. La tabla siguiente sintetiza sus características principales:

**TABLA 7: EVALUACIÓN DE PROVEEDORES Y CAPACIDADES TECNOLÓGICAS**

Proveedor	Escalabilidad y rendimiento	Integración de IA/ML	Gestión de cómputo y almacenamiento	Integración con el ecosistema	Costos y esfuerzo operativo	Fuentes
<b>Google BigQuery (Google Cloud)</b>	Es un almacén de datos empresarial totalmente administrado y <i>serverless</i> que permite análisis de gran escala mediante SQL; decupla el cómputo y el almacenamiento para escalar de forma independiente. Utiliza almacenamiento columnar motor MPP ( <i>Dremel</i> ) y puede consultar petabytes de datos en minutos.	Integra <i>BigQuery ML</i> y capacidades de inteligencia artificial; las nuevas funciones "Gemini" ayudan a explorar datos y generar consultas.	El modelo <i>serverless</i> gestiona automáticamente los recursos y separa cómputo y almacenamiento, lo que permite pagar sólo por el uso.	Ofrece integración nativa con Google Analytics y otros servicios de Google, lo que lo hace atractivo para empresas de marketing digital.	Al ser <i>serverless</i> , elimina la administración de infraestructura y su coste se basa en el volumen de consultas.	(Reynolds, 2025).
<b>Amazon Redshift (AWS)</b>	Almacén de datos totalmente administrado, de escala en petabytes, diseñado para análisis y reportes de alto rendimiento. Usa procesamiento masivamente paralelo (MPP) y almacenamiento columnar. Existen dos modelos: <i>cluster</i> provisionados y Redshift Serverless, el segundo ofrece escalado automático.	Se integra con AWS SageMaker para crear y entrenar modelos de aprendizaje automático desde SQL.	Los nodos RA3 y Redshift Spectrum permiten separar almacenamiento y cómputo y consultar datos en S3 sin cargarlos al clúster.	Está profundamente integrado en el ecosistema AWS: se conecta con S3, Glue, Kinesis, IAM, CloudWatch y otros servicios.	El cliente elige entre pago por uso ( <i>serverless</i> ) o <i>cluster</i> fijos; los costes varían según el consumo y la capacidad reservada.	(Reynolds, 2025).
<b>Azure Synapse Analytics (Microsoft)</b>	Plataforma analítica "sin límites" que combina el antiguo Azure SQL Data Warehouse con motores de Big Data como Apache Spark y Data Explorer. Emplea procesamiento masivamente paralelo y permite consultas SQL en paralelo y escala de forma horizontal.	Integra Synapse Pipelines para ETL/ELT, Synapse Studio como entorno unificado y conexiones nativas a Power BI y Azure Machine Learning.	Separa cómputo y almacenamiento, ofreciendo grupos SQL dedicados (de pago fijo) y grupos SQL <i>serverless</i> con pago por consulta.	Se integra estrechamente con servicios de Azure como Data Lake Storage Gen2, Entra ID y Power BI.	El modelo de pago es flexible (dedicado vs <i>serverless</i> ) y permite pausar los recursos para optimizar el coste.	(Reynolds, 2025).
<b>IBM Cloud Pak for Data (IBM)</b>	Plataforma de datos que agrupa servicios de datos e inteligencia artificial en una arquitectura abierta. Facilita la gestión, gobernanza y análisis en un entorno unificado.	Integrada con IBM Watson; acelera el despliegue de IA y ayuda a descubrir patrones para obtener información valiosa.	Diseñada para ejecutarse en cualquier nube y ofrece una arquitectura extensible; soporta análisis en tiempo real y automatiza el ciclo de vida de IA.	Su arquitectura abierta permite combinar servicios IBM y de terceros.	El modelo de licenciamiento varía según el paquete; la unificación de servicios reduce la complejidad operativa.	(Belova, 2024).

**Análisis comparativo:** BigQuery destaca por su modelo *serverless* y su integración con herramientas de marketing digital; Redshift sobresale en ambientes que ya utilizan AWS y ofrece la opción de clústeres dedicados o *serverless*; Synapse brinda un servicio unificado de almacenamiento y análisis que combina SQL y Big Data con integración profunda a Power BI; Cloud Pak for Data se orienta a empresas que buscan una solución basada en IA y unificada que pueda ejecutarse en cualquier nube.

### 2.2. Hacer una selección de nuestra alternativa mejor de proveedor

Criterios de selección

- Integración con el ecosistema actual de Carnivery. La compañía usa Zoho One, WooCommerce y Facebook como principales fuentes de datos. Su equipo se apoya en herramientas de visualización como Power BI y Looker Studio. Por tanto, es fundamental que el proveedor se integre con fuentes heterogéneas y con herramientas de marketing digital.

- Escalabilidad y modelo de costes. Al tratarse de un e commerce en crecimiento, el proveedor debe gestionar volúmenes crecientes sin inversión inicial elevada y permitir pagar por consumo.
- Capacidades de analítica avanzada. Se requiere soporte para machine learning, segmentación y predicción para ejecutar casos de uso como segmentación RFM y predicción de demanda.
- Simplicidad operativa. Debido a que Carnivery no dispone de un equipo amplio de TI, se valora un servicio serverless que minimice la administración de infraestructuras.

### **Proveedor seleccionado**

Google BigQuery es la alternativa mejor alineada a estas necesidades. Su modelo serverless reduce la carga operativa y permite pagar por consultas ejecutadas. La separación de cómputo y almacenamiento ofrece elasticidad para crecer según la demanda (Reynolds, 2025). Además, BigQuery incorpora nativamente capacidades de aprendizaje automático y análisis geoespacial, lo que facilita el desarrollo de modelos predictivos y segmentación avanzada sin salir de la plataforma. La integración con Google Analytics y otras herramientas de marketing digital apoya a las áreas de ventas y mercadeo de Carnivery (Reynolds, 2025). Finalmente, su compatibilidad con formatos abiertos y su capacidad para consultar datos en data lakes (como archivos en Google Cloud Storage) brindan flexibilidad a futuro (Reynolds, 2025).

### **2.3. Definir nuestro modelo de cloud computing: SaaS, PaaS o IaaS**

Las tres capas de servicios en la nube se diferencian por el nivel de control y responsabilidad. En IaaS (Infrastructure as a Service) el proveedor suministra recursos de computación y almacenamiento y el cliente gestiona el sistema operativo y las bases de datos; PaaS (Platform as a Service) ofrece un entorno de desarrollo o análisis en el que el proveedor administra la infraestructura y el cliente se centra en el desarrollo de soluciones; SaaS (Software as a Service) suministra aplicaciones completas accesibles por web.

Para Carnivery se propone un modelo mixto basado en PaaS y SaaS:

- PaaS para el almacén de datos y procesamiento. BigQuery se consume como una plataforma de análisis serverless (PaaS). Permite ejecutar SQL, construir modelos de machine learning y gestionar petabytes de datos sin administrar máquinas ni bases de datos. La capa de ETL/ELT puede implementarse con servicios como Cloud Dataflow o Cloud Composer, que también se ofrecen como PaaS.
- SaaS para la visualización y colaboración. Herramientas de visualización como Looker Studio o Power BI (en modalidad en línea) proporcionan

dashboards interactivos y colaborativos, alineados con las necesidades de los departamentos comercial, logística y atención al cliente. Estas aplicaciones se consumen como SaaS y permiten compartir reportes sin gestionar infraestructura.

- IaaS mínimo. La infraestructura subyacente se delega al proveedor, evitando la necesidad de mantener servidores propios. Sólo se recurriría a IaaS para almacenar archivos semi estructurados en un data lake (como Google Cloud Storage), cuyo acceso se controla a través de la plataforma.

Este enfoque híbrido equilibra la flexibilidad de PaaS (para desarrollar modelos y procesos personalizados) con la simplicidad de SaaS (para visualización y colaboración). Al apoyarse en servicios gestionados, se reduce la complejidad operativa y se favorece la escalabilidad.

## **2.4. Flujo de la integración de nuestra arquitectura Big Data con el data warehouse o estructura de BI alternativa**

La integración debe garantizar que todas las fuentes alimenten un flujo continuo de datos hacia el modelo de BI. La propuesta se basa en tres capas identificadas en la entrega previa: adquisición, análisis y visualización, permitiendo cuadros de mando en tiempo real. A continuación, se describe cada etapa y su conexión con el almacén de datos:

### **2.4.1 Captura y adquisición de datos**

- Fuentes internas: Zoho CRM/Inventory, WooCommerce y Facebook proporcionan datos de clientes, ventas, abandonos de carrito e interacciones sociales. Estos datos se extraen mediante APIs o conectores nativos y se cargan en un área de staging en un data lake (Google Cloud Storage), donde se conserva tanto información estructurada como semiestructurada (logs de compra, comentarios de redes sociales).
- Fuentes externas: INEC (demografía), MAGAP (precios referenciales) y bases de datos de marketing compradas. Estos conjuntos se ingestan de forma periódica para enriquecer las variables demográficas y socioeconómicas.

### **2.4.2 Procesamiento y almacenamiento**

- ETL/ELT: Se emplea Cloud Dataflow o Cloud Composer para orquestar flujos de ingestión, limpieza y transformación. El proceso corrige valores nulos, consolida formatos, detecta duplicados y genera tablas normalizadas.
- Data warehouse en BigQuery: Después del procesamiento, los datos se cargan en BigQuery. Este almacén soporta consultas complejas mediante SQL estándar y almacena la información en formato columnar, optimizando el acceso a grandes volúmenes de datos (Reynolds, 2025).. BigQuery permite

integrar datos directamente desde el data lake (sin copiarlos), utilizando formatos abiertos y reduciendo la duplicación.

- Modelos analíticos y ML: Los analistas pueden crear vistas y modelos de machine learning dentro de BigQuery para segmentar clientes (modelo RFM), predecir la demanda y calcular el Customer Lifetime Value utilizando BigQuery ML.

### **2.4.3 Explotación y visualización**

- Dashboards: La capa de visualización utiliza herramientas SaaS como Looker Studio o Power BI para consultar BigQuery y generar dashboards por área funcional. En la entrega anterior se plantean tableros de ventas y marketing, logística, atención al cliente y financiero, los cuales se nutren de los datos del data warehouse.
- Toma de decisiones en tiempo real: El flujo continuo de datos hacia el modelo de BI permite cuadros de mando dinámicos que facilitan reaccionar a cambios operativos y comerciales. Se pueden definir alertas para variaciones en indicadores como carritos abandonados, niveles de inventario o tendencia de precios.
- Gobernanza y seguridad: El acceso se gestiona mediante Google Identity and Access Management y se implementan políticas de calidad y retención. Se establecen procesos de respaldo automático y versiones históricas para garantizar la trazabilidad, superando la carencia actual de un data warehouse formal.

### **2.4.4 Beneficios esperados**

Implementar este flujo integral permitirá centralizar la información, automatizar los reportes y generar dashboards interactivos; ello posibilitará decisiones basadas en datos y mejorará la capacidad de respuesta ante variaciones del mercado digital. Los modelos predictivos permitirán anticipar comportamientos de compra, optimizar catálogos y lanzar promociones cruzadas que incrementen la tasa de recompra. Además, la combinación de Big Data y ML incrementa la eficiencia operativa y la personalización, como indican las investigaciones sobre el impacto del Big Data en logística y negocios (Ryzhokhin, A. 2024).

## CAPITULO 3: FRAMEWORKS APLICABLES

En el contexto de Carnivery, un e-commerce ecuatoriano especializado en productos cárnicos, la adopción de tecnologías de Big Data e Inteligencia Artificial (IA) resulta importante para potenciar la personalización de ofertas, la fidelización de clientes y la toma de decisiones basada en datos. Un proyecto de tesis en esta línea requiere analizar críticamente qué frameworks de IA/Machine Learning (ML) y qué componentes del ecosistema de Big Data se adecuan mejor a las necesidades de Carnivery. A continuación, se presenta un análisis comparativo de los principales frameworks de IA (TensorFlow, PyTorch, Scikit-learn, entre otros) y una evaluación de las herramientas complementarias (lenguajes de programación, bases de datos, integración de fuentes como Google Analytics/WooCommerce, visualización de datos, etc.), para finalmente proponer una selección óptima. El enfoque combina rigurosidad académica con un sentido práctico-profesional, buscando sorprender tanto al cuerpo docente como a la dirección de la empresa.

Existen diversos frameworks populares en el ámbito de la IA/ML, cada uno con fortalezas particulares. Entre los más destacados se encuentran Scikit-learn, TensorFlow (y su API de alto nivel Keras) y PyTorch, los cuales cubren desde algoritmos tradicionales hasta deep learning. En la siguiente tabla se resumen sus características principales:

**TABLA 8: ANÁLISIS DE FRAMEWORKS: RENDIMIENTO Y CASOS DE USO**

Framework	Nivel de Abstracción & Facilidad	Enfoque & Capacidades	Rendimiento & Escalabilidad	Casos de Uso Recomendados
<b>Scikit-learn</b>	Alto nivel, API sencilla y amigable para principiantes. Incluye muchas utilidades para preprocesamiento y modelado "caja negra".	ML "clásico" (regresión, clasificación, clustering, etc.) sobre datos estructurados. No soporta entrenamiento en GPU ni redes neuronales profundas de forma nativa.	Excelente para conjuntos de datos pequeños a medianos; no distribuido (todo en un solo equipo). No ideal para deep learning avanzado.	Proyectos exploratorios o prototipos con volumen de datos moderado. Análisis estadístico, predicciones básicas y segmentación estándar.
<b>TensorFlow</b>	Medio a bajo nivel (requiere conocimientos de redes neuronales; aunque Keras lo hace más accesible). Documentación amplia; más complejo que scikit-learn.	Framework de <i>deep learning</i> integral desarrollado por Google. Soporta redes neuronales complejas (CNN, RNN, transformers) computación en GPU/TPU y <i>deploy</i> en producción. Keras [API oficial] simplifica la construcción de modelos.	Altamente optimizado para gran escala y cómputo distribuido. Permite entrenar en múltiples GPUs y clusters, y cuenta con herramientas como TensorBoard para visualización y TensorFlow Serving para despliegue.	Aplicaciones de <i>deep learning</i> en visión, NLP, recomendadores avanzados, etc., especialmente cuando se dispone de grandes volúmenes de datos o necesidad de escalar en producción.
<b>PyTorch</b>	Medio nivel, muy "pythónico" y flexible (define modelos de forma imperativa y orientada a objetos). Curva de aprendizaje manejable; preferido en entornos de investigación por su facilidad de depuración.	Framework de <i>deep learning</i> impulsado por Meta (Facebook). Enfatiza la construcción dinámica de grafos computacionales permitiendo modificaciones sobre la marcha. Excelente para personalizar algoritmos y experimentar con nuevas arquitecturas.	Rendimiento equiparable a TensorFlow en entrenamiento con GPU. Ofrece menor madurez en herramientas de despliegue aunque ha mejorado (TorchServe, ONNX). Comunidad activa; ideal para prototipos de investigación.	Proyectos de investigación y desarrollos a medida de modelos IA (NLP, visión, series de tiempo complejas) donde se requiera máxima flexibilidad. También útil en producción si el equipo cuenta con la expertise, aunque TensorFlow históricamente ha dominado ese ámbito.

**Análisis Comparativo:** Scikit-learn destaca por su simplicidad y accesibilidad, siendo apropiado para quienes se inician en ML o para soluciones basadas en algoritmos

estándar de aprendizaje automático (Ruczyński, Ł., 2025). Sin embargo, sus limitaciones surgen frente a problemas que involucran redes neuronales profundas o enormes conjuntos de datos – ámbitos en los que TensorFlow y PyTorch sobresalen. TensorFlow, respaldado por Google, se ha consolidado como un entorno robusto para deep learning a escala industrial, con un ecosistema que incluye soporte multi-lenguaje (APIs en Python, C++, Java, JavaScript, etc.) y herramientas avanzadas de despliegue (Pykes, 2022). Por su parte, PyTorch ha ganado popularidad en la academia e investigación por su enfoque más dinámico y código más limpio para desarrolladores Python (Pykes, 2022).

Según Diego Céspedes (entrevista 23 de septiembre de 2025) No existe un “ganador absoluto” entre TensorFlow y PyTorch; de hecho, una regla pragmática es elegir el que el equipo domine mejor. Generalmente, PyTorch se asocia a entornos de experimentación (donde la rapidez para probar nuevas ideas es necesaria), mientras que TensorFlow suele preferirse en entornos de producción por su madurez en herramientas de servificación de modelos. En cuanto a Keras, cabe señalar que históricamente fue un framework independiente de alto nivel, pero desde TensorFlow 2.x se integra como su API oficial, heredando todas las ventajas de TensorFlow. Keras facilita enormemente la implementación de redes neuronales (“Deep learning for humans” es su lema), reduciendo la complejidad de código y haciendo más accesible el desarrollo de IA avanzada. De hecho, frameworks como Keras y PyTorch han democratizado el uso del deep learning, acercándolo a problemas aplicados como el reconocimiento de imágenes sin requerir conocimientos extremadamente especializados.

En síntesis, Scikit-learn es ideal para tareas de análisis de datos tradicional (por ejemplo, modelos predictivos basados en regresión/clasificación sobre datos tabulares relativamente pequeños), mientras que TensorFlow/Keras o PyTorch son preferibles para construir modelos predictivos más sofisticados cuando se cuenta con grandes volúmenes de información o se abordan problemas complejos de visión por computador, procesamiento de lenguaje natural, recomendación con embeddings, etc. La decisión debe basarse tanto en los requerimientos técnicos (tipo de algoritmo, tamaño de datos, necesidad de escalar) como en la capacitación del equipo de desarrollo. Como indicó el experto Diego Céspedes en reciente entrevista, es fundamental no “sobre-ingeniar” la solución: conviene adoptar el framework más potente solo si el caso de uso lo amerita, evitando la complejidad innecesaria (Céspedes, entrevista, 2025). En el siguiente apartado se aplica esta lógica al caso Carnivery.

### **3.1. Selección del Framework Óptimo para Carnivery**

Tras la comparación anterior, es pertinente preguntarse: ¿Cuál de estos frameworks conviene adoptar en Carnivery? Dada la naturaleza del negocio (comercio



electrónico) y los objetivos planteados (p.ej., segmentación de clientes, predicción de churn o abandono, recomendación de productos, análisis de sentimientos en feedback, etc.), la solución podría requerir desde algoritmos de machine learning “clásicos” hasta técnicas de deep learning. A continuación, se justifica la selección propuesta:

### **3.1.1. Enfoque inicial con Scikit-learn y librerías afines:**

Para muchas de las necesidades inmediatas de Carnivery (análisis descriptivo de ventas, segmentación RFM de clientes, predicción básica de recurrencia de compra), Scikit-learn resulta apropiado por su rapidez de desarrollo y baja complejidad. Su facilidad de uso permite iterar rápidamente sobre modelos de regresión logística, árboles de decisión, random forests, clustering k-means, etc., obteniendo resultados accionables en poco tiempo (Ruczyński, Ł, 2025). Además, scikit-learn cuenta con utilidades para validación cruzada, pipelines de pre procesamiento y métricas, lo que agiliza el ciclo de modelado sin requerir una gran infraestructura. Dado que Carnivery dispone de datos transaccionales y de clientes de volumen relativamente manejable (miles de registros provenientes de la tienda online y CRM), estas técnicas pueden capturar patrones valiosos sin incurrir en la complejidad de una red neuronal profunda. Por ejemplo, un modelo de bosque aleatorio en scikit-learn podría bastar para identificar atributos de clientes con mayor probabilidad de recompra.

### **3.1.2. Escalabilidad a Deep Learning con TensorFlow/Keras:**

Ahora bien, a medida que Carnivery incremente el alcance de sus datos (por ejemplo, integrando eventos detallados de navegación web, historiales completos de interacciones o datos no estructurados como comentarios en redes sociales), podría ser necesario recurrir a modelos de deep learning para aprovechar toda la información. En ese escenario, se propone adoptar TensorFlow como framework base, haciendo uso de Keras para implementar las redes neuronales de forma eficiente. La elección por TensorFlow sobre PyTorch responde a criterios de producción e integración: Carnivery planea montar su arquitectura analítica en la nube de Google (como se detalla más adelante), donde TensorFlow se integra naturalmente con servicios como BigQuery ML y APIs de Google Cloud AI. De hecho, TensorFlow ofrece integración nativa con múltiples lenguajes (Python, Java, JavaScript) y soporte multiplataforma, lo que facilita desplegar modelos entrenados en distintos entornos. Asimismo, TensorFlow cuenta con TensorBoard para visualizar el entrenamiento y con un sistema de servificación maduro (TF Serving/TensorFlow Lite) para llevar modelos a producción en servidores o dispositivos móviles. Estas características resultan valiosas si, por ejemplo, Carnivery quisiera incorporar un modelo de recomendación en tiempo real dentro de su sitio web o una app móvil. Si bien PyTorch podría igualmente usarse (su



rendimiento de entrenamiento es similar y su uso en investigación está muy difundido), la balanza se inclina hacia TensorFlow en Carnivery por la disponibilidad de soporte corporativo (Google) y la posible sinergia con otras herramientas del ecosistema Google Analytics/Cloud que la empresa ya emplea. No obstante, cabe aclarar que la adopción de TensorFlow/Keras sería gradual y justificada por casos de uso específicos, por ejemplo, si se decide entrenar un modelo profundo de learning-to-rank para el motor de recomendaciones o una red neuronal recurrente para pronóstico de demanda, ámbitos donde las redes neuronales han demostrado beneficios.

En resumen, se propone una estrategia híbrida y escalable: comenzar resolviendo los problemas de business intelligence y analítica predictiva inmediata con herramientas ligeras (Python + scikit-learn, complementado con librerías como XGBoost/LightGBM para modelos de boosting en caso necesario), y estar preparados para dar el salto a deep learning con TensorFlow/Keras cuando la complejidad de los datos o las ambiciones analíticas lo requieran. Esta aproximación incremental equilibra bien la relación costo-beneficio: se obtienen “quick wins” analíticos a corto plazo, mientras se sienta una base tecnológica compatible con iniciativas de IA más avanzadas en el futuro cercano. Además, al centrar la solución en Python (común denominador de scikit-learn y TensorFlow), se simplifica la curva de aprendizaje para el equipo y la integración entre componentes.

### **3.2. Lenguajes de Programación en el Ecosistema Big Data**

La elección del lenguaje base es un pilar fundamental en la arquitectura de cualquier solución de Big Data e IA. En el caso de Carnivery, se ha determinado que Python será el lenguaje principal de desarrollo, dadas sus múltiples ventajas en ciencia de datos. Python es un lenguaje de alto nivel cuya sintaxis clara y extensa comunidad lo han posicionado como el favorito para analítica y ML (Full Scale, 2024). Su facilidad de uso permite que profesionales no necesariamente formados en ciencias de la computación (estadísticos, analistas de negocio, etc.) puedan construir análisis sofisticados con relativa rapidez. Asimismo, Python cuenta con un ecosistema rico de librerías: desde pandas para manipulación de datos tabulares, hasta NumPy/SciPy para cómputo numérico, pasando por las ya mencionadas scikit-learn, TensorFlow, PyTorch, etc. – esto hace posible abordar prácticamente cualquier tarea de datos con herramientas probadas. No es de extrañar que Python sea considerado “una herramienta muy eficiente para Big Data, en parte por la gran comunidad que desarrolla librerías de terceros”.

Ahora bien, es importante reconocer que Python no está exento de desventajas. Al ser un lenguaje interpretado, su velocidad de ejecución cruda es menor comparada con lenguajes compilados como C++ o Java. Esto significa que bucles muy intensivos u operaciones sobre enormes volúmenes de datos podrían ser más lentos en Python.

En la práctica, este inconveniente se mitiga porque muchas operaciones pesadas están delegadas a librerías optimizadas en C/C++ (por ejemplo, las operaciones matriciales de NumPy, o el backend de TensorFlow). Aun así, se suele emplear Python “para tareas de integración o donde no haya cálculos demasiado pesados”. En el diseño para Carnivery, Python orquestrará los flujos de datos y la lógica de negocio analítica, apoyándose en componentes de bajo nivel optimizados cuando sea necesario (p. ej., consultas SQL en la base de datos, rutinas numéricas en NumPy, etc.).

En comparación, R surgió como el lenguaje preferido por estadísticos y sigue siendo muy utilizado en entornos académicos y análisis exploratorio. R es potente para procedimientos estadísticos y visualización rápida, con paquetes como ggplot2, y tiene un largo historial en data mining. Sin embargo, R presenta una sintaxis y estilo más cercanos a las matemáticas que a la ingeniería de software, lo cual puede suponer una barrera para programadores formados en lenguajes tradicionales. En otras palabras, R está altamente optimizado para análisis estadístico (ofrece infinidad de librerías especializadas en pruebas, modelos lineales, series de tiempo, etc.), pero su uso en entornos de producción web/empresarial ha sido más limitado. Además, la comunidad de R si bien es muy activa, es menor en tamaño que la de Python. En Carnivery, Python ofrecerá mayor versatilidad para integrar la analítica con los sistemas operativos de la empresa (web, bases de datos, servicios REST), por lo que R no será el lenguaje central. No obstante, el equipo no descarta utilizar R de forma complementaria en análisis específicos donde brille – por ejemplo, R cuenta con implementaciones de vanguardia en métodos estadísticos y podría emplearse vía RStudio para exploraciones ad-hoc de datos demográficos o econométricos, si surgiera la necesidad. En resumen, R se considera una herramienta secundaria, útil para ciertos análisis de nicho, pero Python será el estándar dado su balance entre facilidad y capacidad de producción.

En el ecosistema Big Data de Carnivery también es relevante mencionar Java (y su primo cercano, Scala). Muchas de las grandes tecnologías de big data se han escrito en Java/Scala – por ejemplo, Hadoop y Apache Spark están construidos sobre la máquina virtual Java aprovechando su rendimiento y escalabilidad. Esto ha llevado a que Java conserve un lugar en desarrollos de back-end y procesamiento distribuido a gran escala. Sin embargo, desde la perspectiva de un equipo de análisis de datos, Java puro no suele ser la primera opción debido a su verbosidad y mayor tiempo de desarrollo para prototipado. En lugar de ello, se aprovecha su poder a través de herramientas que lo abstraen: por ejemplo, si Carnivery adoptara Spark para procesar grandes datasets, podría utilizar PySpark, que permite escribir las transformaciones en Python mientras el motor Spark (en Java/Scala por debajo) las ejecuta en cluster. De igual modo, existen librerías que permiten integrar código Java con Python (JPype, Py4J) o viceversa, lo que brinda flexibilidad. En definitiva,

Java/Scala serán lenguajes “invisibles” pero presentes en la arquitectura, especialmente si escalamos a soluciones como Spark o frameworks de streaming; no se prevé desarrollar módulos de IA manualmente en Java, pero sí apoyarse en su ecosistema cuando corresponda. El rol de Java se alinea más al desarrollo de aplicaciones empresariales y a soportar sistemas de alto desempeño, mientras que Python domina la capa analítica de modelado.

Finalmente, es imposible hablar de manejo de datos sin mencionar SQL. SQL (Structured Query Language) no es un lenguaje de propósito general como los anteriores, pero es fundamental en el arsenal de un científico de datos. Carnivery planea consolidar una data warehouse donde SQL será clave para consultar y transformar datos. Actualmente, la organización carece de esta cultura analítica basada en SQL – se identificó que no existe una data warehouse formal ni se emplean herramientas avanzadas de análisis como SQL o Power BI en el equipo actual, lo cual representa una brecha a subsanar.

El dominio de SQL permitirá extraer información de las bases de datos transaccionales (por ejemplo, consultas que unan tablas de ventas, clientes y productos para construir indicadores), crear views o tablas resumen para modelos, y en general, asegurará que la analítica se pueda escalar con eficiencia. Incluso, SQL está evolucionando para abarcar machine learning: proveedores cloud como Google Cloud ofrecen BigQuery ML, que permite entrenar modelos de regresión, clasificación o clustering directamente mediante consultas SQL en el almacén de datos. Esta funcionalidad podría ser aprovechada en Carnivery para modelos sencillos ejecutados 100% en la base de datos (por ejemplo, un modelo de regresión logística para propensión de compra, entrenado y aplicado dentro de BigQuery sin sacar datos).

En suma, SQL complementa a Python: mientras Python gestionará el flujo general y el modelado avanzado, SQL servirá para manipular los datos en su almacenamiento masivo de forma óptima y reproducible. Es por ello que se contempla capacitación del equipo en SQL y su uso intensivo en la capa de datos.

### **3.3. Tecnologías de Almacenamiento de Datos: Bases Relacionales vs. NoSQL**

Otra decisión medular es la selección de base de datos o sistema de almacenamiento que sustentará la data warehouse y los data marts analíticos de Carnivery. Actualmente, los datos de la empresa se hallan dispersos: por un lado, transacciones e información de productos/clientes residen en la base de datos MySQL de WooCommerce/WordPress y en el sistema ERP (Contífico); por otro lado, datos de interacción y marketing se capturan en herramientas de terceros como Google Analytics, redes sociales, Zoho CRM, etc. El proyecto busca integrar estas fuentes en

un repositorio unificado, por lo que debemos contemplar qué tecnología de base de datos emplear para centralizar y explotar la información.

La disyuntiva principal suele ser entre sistemas relacionales tradicionales (SQL databases) y sistemas NoSQL de nueva generación. Ambas tienen pros y contras para nuestro caso:

### **3.3.1. Bases de datos relacionales (SQL):**

Son las de arquitectura clásica (modelo entidad-relación), como PostgreSQL, MySQL, Microsoft SQL Server, Oracle, etc. Se caracterizan por garantizar consistencia y transaccionalidad (ACID), usar esquemas predefinidos y admitir el potente lenguaje SQL para consultas. Para Carnivery, que maneja datos estructurados (ventas, clientes, inventario) con relaciones definidas, un almacén relacional es una opción natural. PostgreSQL en particular destaca por ser open-source, robusto y extensible; podría desplegarse como data warehouse a pequeña escala para consolidar datos de WooCommerce (inicialmente en MySQL) junto con otros datos.

Una ventaja de PostgreSQL es su soporte de JSON y extensiones geoespaciales, aportando flexibilidad semi-estructurada sin dejar de ser relacional. Otra opción valorada es recurrir a una data warehouse en la nube: por ejemplo, Google BigQuery o Amazon Redshift, que internamente son sistemas columnar y distribuidos, pero se consultan con SQL.

De hecho, en la evaluación tecnológica se consideraron BigQuery y Redshift como candidatos para implementar la data warehouse de Carnivery, dado que ambas solucionan la escalabilidad de manera transparente. BigQuery tiene la ventaja de ser serverless y con integración nativa a Google Analytics, lo que la hace atractiva para casos de marketing digital como este. Usar una data warehouse relacional/columnar facilitará cruzar datos de diferentes fuentes mediante JOINS y consultas complejas, algo crucial para obtener vistas 360° del cliente.

Además, la mayoría de herramientas de BI y visualización se conectan cómodamente a bases SQL. Por todo ello, la columna vertebral de la arquitectura de Carnivery será un almacén de datos relacional (muy probablemente BigQuery en la nube, por su capacidad de escalar a volúmenes masivos y su conexión directa con Google Analytics).

### **3.3.2. Bases de datos NoSQL:**

Representan un paradigma distinto, orientado a escenarios donde las 3V de Big Data (volumen, velocidad, variedad) exceden las capacidades de los SQL tradicionales. Existen varios tipos (documentos, clave-valor, columnas, grafos). En el ecosistema de Big Data, MongoDB es una de las NoSQL más reconocidas, con un modelo orientado a documentos JSON que la hace muy flexible. A diferencia del rígido

esquema relacional, MongoDB almacena cada registro como un documento (similar a un objeto JSON), permitiendo que diferentes documentos en una colección tengan estructuras distintas. Esta flexibilidad resulta útil cuando los datos provienen de fuentes heterogéneas o cuando el esquema puede cambiar con frecuencia.

En Carnivery podríamos considerar MongoDB para almacenar, por ejemplo, datos semiestructurados como eventos de navegación web o clics, donde no queremos descartar atributos poco comunes. Otra fortaleza de MongoDB es su fácil escalado horizontal: mediante replicación y sharding (particionamiento) puede distribuir datos en múltiples nodos, permitiendo manejar crecimientos altos de volumen.

Por eso se la recomienda en entornos que requieren escalabilidad elástica. Sin embargo, NoSQL también tiene limitaciones: suelen sacrificar las garantías ACID plenas (adoptando consistencia eventual), y las consultas complejas o agregaciones avanzadas pueden no ser tan sencillas como con SQL estándar (aunque MongoDB ofrece su propio lenguaje de agregación). En Carnivery, el uso de NoSQL se justifica si enfrentamos datos muy diversos o en cantidades realmente masivas.

Por ejemplo, si se almacenan millones de registros de actividad de usuarios (logs web, interacciones en redes), una base NoSQL como MongoDB o un almacén de big data como Apache Cassandra podría complementar a la data warehouse relacional. Cassandra, en particular, es una base de datos NoSQL de tipo column family usada en escenarios de escrituras a altísima velocidad (inicialmente desarrollada por Facebook).

Otra herramienta interesante es Elasticsearch, un motor de búsqueda y analítica en tiempo real que indexa datos semiestructurados y permite consultas de texto completo con gran rapidez. Si Carnivery quisiera habilitar búsquedas eficientes (por ejemplo, de comentarios de clientes, o de productos por descripción) y analítica de logs, Elastic sería una pieza a considerar en la arquitectura.

En conclusión, la estrategia de Carnivery combinará lo mejor de ambos mundos: se establecerá una data warehouse relacional principal (SQL) para las entidades centrales del negocio (ventas, clientes, productos, métricas consolidadas), asegurando consistencia y facilidad de análisis con SQL. Encima de este, se construirán data marts temáticos para distintas áreas (marketing, operaciones, gerencia) que alimentarán dashboards.

Paralelamente, se mantendrá la puerta abierta al uso de bases NoSQL especializadas para datos complementarios: por ejemplo, si se recogen grandes volúmenes de datos de comportamiento de usuarios en la web, estos podrían almacenarse en una colección MongoDB o en Google Cloud Bigtable, desde donde se podrían extraer insights que luego se integren con los datos principales. Lo crucial es que haya mecanismos de integración entre ambos almacenes (por ejemplo, cargar resúmenes

de datos NoSQL al warehouse SQL regularmente, o usar federación de consultas si la tecnología lo permite).

Esta arquitectura híbrida garantiza que Carnivery pueda escalar en volumen y variedad de datos sin comprometer la facilidad de explotación analítica. Como menciona Pérez Galán (2019), la amplia disponibilidad de datos hoy obliga al uso de tecnologías Big Data para un aprovechamiento óptimo, combinando distintas herramientas especializadas la clave está en elegir la herramienta adecuada para cada tipo de datos.

### **3.4. Integración de Fuentes de Datos y Herramientas de Soporte**

Además de frameworks de ML, lenguajes y bases de datos, el ecosistema analítico de Carnivery requiere integrar múltiples herramientas de soporte que garanticen el flujo de datos desde la captura hasta la visualización de insights. Entre los elementos críticos identificados están: la conexión con fuentes externas (p. ej. Google Analytics y la plataforma e-commerce), los procesos ETL/ELT, y las plataformas de visualización de datos. Se abordan a continuación:

#### **3.4.1. Integración con Google Analytics 4 (GA4):**

Carnivery obtiene abundante información sobre el comportamiento de usuarios en su sitio web mediante Google Analytics. Para explotar plenamente estos datos (clics, páginas vistas, embudos de conversión, etc.), se habilitará la exportación nativa de GA4 hacia BigQuery. GA4 ofrece una integración gratuita con BigQuery que exporta todos los eventos crudos diariamente. Esto permite combinar los datos analíticos web con datos transaccionales propios en la data warehouse, utilizando SQL para lograr reportes avanzados sin las limitaciones de muestreo de GA (Google Cloud, 2023).

Por ejemplo, en BigQuery podríamos unir los eventos de GA (como “añadir al carrito”) con la base de pedidos reales de Carnivery para calcular conversiones efectivas o para crear segmentos de clientes según su comportamiento de navegación. Asimismo, BigQuery abre la puerta a construir modelos de ML simples sobre los datos de GA, como segmentar audiencia por propensión de compra o detectar anomalías en tráfico.

La integración GA4-BigQuery se configura una sola vez y luego los datos fluyen automáticamente cada día, resolviendo el problema identificado de extracción manual desde GA. Con esto, Carnivery contará con datos de analítica digital sin procesar en su poder, fomentando análisis personalizados (ejemplo: correlacionar campañas de marketing digital con ventas efectivas, más allá de lo que GA provee por defecto).



### 3.4.2. Integración con WooCommerce y sistemas operativos:

La tienda en línea de Carnivery funciona sobre WooCommerce (WordPress), que a su vez utiliza MySQL como base de datos. Adicionalmente, se tiene Zoho CRM conectado a WooCommerce para gestionar clientes y Contífico como ERP para facturación. Actualmente, estos sistemas están parcialmente integrados mediante conectores nativos, pero no existe un repositorio analítico consolidado. El plan es implementar procesos ETL (Extracción, Transformación y Carga) que tomen datos de estas fuentes y los alojen en la data warehouse.

Por ejemplo, mediante scripts Python y conectores de base de datos (usando bibliotecas como pymysql o psycopg2), se extraerán periódicamente las tablas clave de WooCommerce (órdenes, detalles de órdenes, clientes, productos) y se volcarán al almacén de datos. Lo mismo aplicará para datos de Zoho CRM (que ofrece APIs para exportar contactos, leads, etc.) y de Contífico (posiblemente mediante reportes CSV o integración directa si dispone de API).

Dado que, según el diagnóstico, actualmente no hay flujos automatizados de carga, se pondrá mucho énfasis en calendarizar estas cargas (p.ej., diarias para transacciones, semanales para datos menos volátiles) y en establecer mecanismos de control de calidad de datos. Se podrían incorporar herramientas Open Source como Apache NiFi o simplemente orquestar con cronjobs y scripts Python, dependiendo de la complejidad. Lo importante es lograr una alimentación continua y confiable de la información hacia la plataforma de BI.

En cuanto a WooCommerce específicamente, existe una REST API robusta que permite extraer información de pedidos, productos y clientes con autenticación segura. Se puede emplear la biblioteca Python woocommerce-api para simplificar esta integración. Así, cada nueva venta registrada en WooCommerce podría reflejarse en la data warehouse casi en tiempo real o con mínima latencia. Esto abonará al objetivo de Carnivery de tener una analítica cercana al tiempo real, al menos en lo referente a ventas y stock.

### 3.4.3. Procesamiento de datos y Big Data:

En la fase inicial, el volumen de datos de Carnivery será manejable con herramientas tradicionales (unas decenas de miles de registros de ventas y navegación). Sin embargo, es prudente contemplar la adopción de frameworks de Big Data processing si los datos crecieran exponencialmente. En este sentido, Apache Spark se vislumbra como un aliado potencial. Spark es un motor de procesamiento en memoria muy veloz y capaz de distribuir tareas en un clúster.

Además, permite escribir las transformaciones en Python (PySpark) beneficiándose de su motor escalable escrito en Scala/Java. Si Carnivery llegara a manejar datasets de órdenes históricas, clics o interacciones del orden de millones de registros, se podría incorporar Spark para realizar cálculos intensivos (por ejemplo, recomputar clusters de clientes o entrenar modelos sobre un cluster) de forma eficiente. Spark se integraría bien con la data lake/warehouse, leyendo datos desde BigQuery (usando el connector Spark-BigQuery) o desde almacenamiento en nube (Google Cloud Storage o HDFS).

Otra herramienta emergente a monitorear es Apache Flink para streaming de datos en tiempo real, pero probablemente en Carnivery un enfoque de micro-batch con Spark Structured Streaming (o incluso simplemente cargas frecuentes) será suficiente para las necesidades de negocio.

En resumen, la arquitectura se diseñará Big Data-ready, aunque sin incurrir en la complejidad hasta que sea necesario: esto implica favorecer soluciones escalables (BigQuery, Spark) pero arrancando con implementaciones simples (PostgreSQL local, pandas) y migrando a la tecnología de mayor escala conforme los datos lo demanden.

#### **3.4.4. Plataformas de visualización y BI:**

El output del proceso analítico son los insights que deben llegar a los tomadores de decisión de Carnivery (CEO, áreas de marketing, operaciones). Para ello, se definió la implementación de dashboards interactivos mediante herramientas de visualización.

Tras evaluar opciones, se han considerado Microsoft Power BI y Google Looker Studio (antes Google Data Studio) como las principales alternativas, por su facilidad de uso y compatibilidad con nuestras fuentes. Power BI es una solución empresarial ampliamente utilizada que permite crear cuadros de mando ricos y compartirlos con usuarios a través de su servicio en la nube.

Por su parte, Looker Studio (de Google) ofrece una interfaz 100% web, gratuita, ideal para integrar con BigQuery y Google Analytics de forma nativa. De hecho, dado que Carnivery ya opera en ecosistema Google, Looker Studio resulta atractivo para visualizar los datos integrados de BigQuery/GA4 (Google Cloud, 2023). Ambas herramientas tienen conectores hacia múltiples fuentes (SQL databases, hojas de cálculo, APIs), por lo que se podrían usar complementariamente incluso.

En la práctica, se podría adoptar Power BI para análisis internos más complejos y detallados (por ejemplo, un dashboard financiero avanzado con funciones DAX), mientras que Looker Studio podría emplearse para reportes ligeros de marketing compartidos con socios externos o para incrustar en la intranet. Durante la



ejecución del proyecto, se sugiere hacer prototipos en ambas plataformas y evaluar cuál se ajusta mejor al flujo de trabajo de Carnivery. Lo importante es que con cualquiera de ellas se lograría el objetivo de democratizar el acceso a la información en la empresa, permitiendo a personal no técnico interactuar con los datos en tiempo real. Por ejemplo, un dashboard podría mostrar indicadores clave como ventas diarias por categoría, ticket promedio, tasa de carritos abandonados, etc., y actualizarse automáticamente con cada carga de datos. Esto eliminaría la dependencia actual de reportes manuales en Excel y aumentaría la agilidad en la toma de decisiones.

Adicionalmente, para análisis ad hoc exploratorios, el equipo de datos seguirá usando Jupyter Notebooks en Python, generando visualizaciones con librerías como Matplotlib, Seaborn o Plotly según convenga, antes de incorporar hallazgos consolidados a los dashboards oficiales.

**Otros componentes de soporte:** Cabe mencionar brevemente otras herramientas que complementarán el ecosistema: (a) Sistemas de control de versiones como Git, para mantener el código de ETL, análisis y modelos bajo buena gestión (por ejemplo, en un repositorio GitHub o GitLab privado del proyecto); (b) Entornos de desarrollo interactivo como JupyterLab o RStudio (este último si se utilizan scripts en R), que facilitan la colaboración y documentación del análisis; (c) Herramientas de automatización y orquestación, como Apache Airflow o simplemente tareas programadas, que aseguren que los flujos de datos se ejecuten en orden y con dependencias definidas (por ejemplo, primero extraer datos de WooCommerce, luego actualizarlos en BigQuery, luego refrescar dashboards); (d) Consideraciones de seguridad y gobernanza de datos, implementando roles de acceso a la base de datos, anonimización de datos sensibles de clientes, y manteniendo un data catalog que describa los campos disponibles para análisis.

## Conclusiones

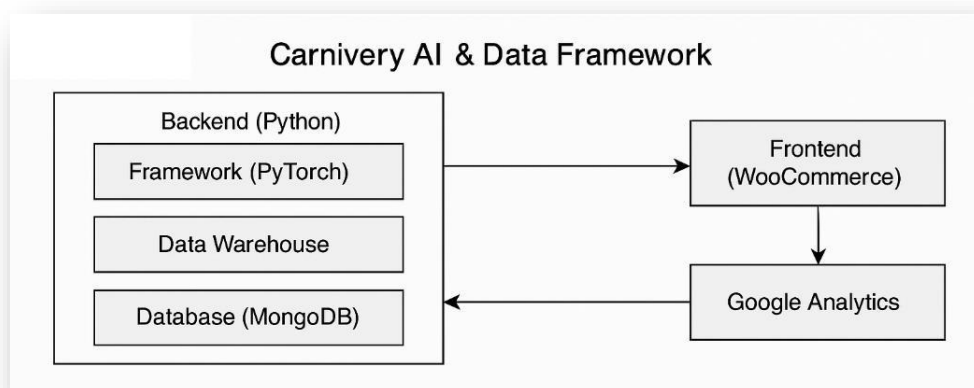
Integrar un ecosistema de Big Data e IA exitoso en Carnivery requiere un equilibrio entre adoptar tecnologías vanguardistas y ajustarlas a la realidad operativa de la empresa. Tras un análisis exhaustivo, se recomienda cimentar la plataforma analítica sobre Python como lenguaje unificador, capitalizando su facilidad de desarrollo y vasta comunidad. En cuanto a frameworks de IA/ML, la estrategia propuesta combina lo mejor de dos mundos: aprovechar Scikit-learn y algoritmos tradicionales para lograr valor inmediato en problemas de negocio actuales, y paralelamente preparar el terreno para incorporar frameworks de deep learning (TensorFlow/Keras) cuando la complejidad de los datos y los casos de uso lo demanden (p. ej. motores de recomendación personalizados, predicciones más precisas mediante redes neuronales). Esta elección dual está orientada por la

necesidad de resultados tempranos y concretos, sin perder de vista la escalabilidad futura hacia IA más sofisticada.

En el ecosistema de soporte, se opta por un modelo híbrido de almacenamiento de datos: una data warehouse relacional (con BigQuery/PostgreSQL) como repositorio central de información consolidada, complementado con tecnologías NoSQL bajo demanda para datos semiestructurados o de altísimo volumen (MongoDB, Elastic, etc.). Se enfatiza la integración continua de fuentes clave: la conexión nativa de Google Analytics 4 con BigQuery asegurará una rica fuente de datos de comportamiento de usuarios, mientras que conectores ETL extraerán datos de WooCommerce, Zoho y Contífico hacia la misma plataforma unificada. Sobre esta base, florecerá la capa de Business Intelligence, con dashboards interactivos en Power BI/Looker Studio que democratizarán el acceso a insights en Carnivery.

El resultado esperado es una organización data-driven, donde decisiones de marketing, operaciones y estrategia se fundamenten en análisis rigurosos: por ejemplo, segmentaciones avanzadas que incrementen la tasa de recompra y reduzcan la fuga de clientes, pronósticos que optimicen el inventario, y ofertas personalizadas que eleven el ticket promedio.

En definitiva, la combinación de herramientas seleccionada – frameworks de IA modernos, lenguajes apropiados, bases de datos escalables y plataformas de visualización intuitivas – proporcionará a Carnivery una ventaja competitiva basada en datos. No solo se atenderán las necesidades analíticas actuales, sino que la empresa quedará preparada para incorporar innovaciones de inteligencia artificial conforme estas surjan. Esta visión integral y crítica de las tecnologías garantiza que la tesis no solo cumpla con el rigor académico, sino que también aporte un plan tecnológico realista y de alto impacto para la empresa, sorprendiendo gratamente tanto al comité evaluador como a la dirección de Carnivery.



**GRÁFICO NO. 6: CARNIVERY IA & DATA FRAMEWORK**

## **CAPITULO 4 MODELO DE PROPENSIÓN DE COMPRA CON MACHINE LEARNING (NON REAL TIME)**

En los últimos años se ha generado una cantidad de datos sin precedentes, pero solo una fracción mínima es analizada, pese al potencial disruptivo de la información para optimizar campañas, mejorar la competitividad y aumentar ingresos. Carnivery, una empresa de e-commerce cárnico, busca aprovechar este potencial mediante dos casos de uso de ciencia de datos y machine Learning.

### **4.1. Objetivo**

Para el presente caso, se desarrollan actividades de marketing offline o batch, es decir, no requiere ejecutarse en milisegundos para el usuario final, sino que funciona en ciclos (por ejemplo, semanalmente) para informar las decisiones de campañas de mercadotecnia. El proyecto consiste en construir un modelo de propensión de compra (propensity to buy) que, a partir de patrones históricos de comportamiento de los clientes, pueda predecir qué usuarios tienen mayor probabilidad de realizar una compra en el futuro cercano. El objetivo de negocio de este modelo es optimizar las campañas semanales de marketing, identificando a qué clientes se debe dirigir una promoción o incentivo cada semana para maximizar la conversión y el retorno de la campaña.

En términos sencillos, el modelo asignará a cada cliente (o cada prospecto) un puntaje de propensión o probabilidad de compra. Por ejemplo, podría predecir que el cliente A tiene un 80% de probabilidad de comprar esta semana (quizá porque ha estado navegando productos y hace tiempo que no compra, o porque pertenece a un segmento altamente activo), mientras que el cliente B tiene solo un 20%. Con esta información, el equipo de marketing puede segmentar la audiencia de sus campañas, focalizando los esfuerzos (como emails promocionales, anuncios personalizados o cupones) en aquellos clientes con mayor propensión de conversión, y evitando saturar o gastar recursos en quienes probablemente no responderán. Esto se traduce en campañas más eficientes, personalizadas y rentables.

El modelo de propensión se alinea con una tendencia clave del marketing moderno: pasar de campañas masivas iguales para todos, a campañas dirigidas por datos donde cada cliente recibe el mensaje adecuado según su probabilidad de responder. Chaffey & Ellis-Chadwick (2016) señalan que el análisis de patrones de compra y preferencias de los clientes permite diseñar campañas promocionales dirigidas, aumentando su efectividad. Asimismo, desde la perspectiva de Carnivery, se busca mejorar la tasa de conversión de campañas y la tasa de respuesta (por ejemplo, porcentaje de clics en un correo o de clientes que usan un cupón enviado) al utilizar este modelo predictivo.

En suma, el objetivo es mejorar la toma de decisiones de marketing con analítica predictiva: saber quién tiene mayor propensión a comprar y cuándo, para así anticiparse y dirigirle estímulos que lo conviertan en comprador efectivo.

Un objetivo secundario, pero importante, es incrementar el ROI de marketing. Al centrar los recursos (descuentos, comunicaciones, etc.) en clientes de alta propensión, se espera que cada dólar invertido en la campaña genere más retorno que en el esquema tradicional. También puede ayudar a reactivar clientes inactivos: el modelo podría detectar patrones de clientes que están al borde de la deserción (churn) y que, con el incentivo correcto en el momento justo, podrían volver a comprar.

Así, podríamos generar iniciativas de retención proactiva. En definitiva, el modelo de propensión de compra servirá como un “filtro inteligente” para que Carnivery enfoque sus esfuerzos semanales en los clientes correctos con el mensaje correcto, elevando la eficacia de sus estrategias de marketing relacional (Kotler & Keller, 2016).

#### **4.2. Fuentes de datos utilizadas**

El desarrollo de un modelo de propensión de compra requiere recopilar y unificar diversos datos históricos y características de los clientes, que en conjunto permitan predecir el comportamiento futuro. Las fuentes de datos principales para este caso de uso incluyen:

- **Datos de comportamiento de navegación web:** Al igual que en el caso anterior, los datos de navegación (páginas vistas, productos buscados, tiempo en el sitio, carritos abandonados, etc.) son insumos valiosos. Sin embargo, en este caso histórico, se analizarán patrones acumulados en periodos largos (semanas o meses) en lugar de reaccionar en tiempo real. Por ejemplo, cuántas visitas realizó un cliente en el último mes, o si miró ciertos productos repetidamente sin comprar. Google Analytics 4 u otras herramientas analíticas proveen estos datos agregados por usuario – incluso GA4 ofrece métricas como “probabilidad de compra” nativas, pero entrenar un modelo propio permite mayor personalización. Carnivery podría exportar las colecciones de eventos de GA (via BigQuery) o los logs de servidor a un entorno de análisis (Hadoop/Spark) para extraer características de navegación de cada cliente.
- **Historial de conversiones y transacciones:** Los datos de ventas pasadas de cada cliente son fundamentales. Incluye: fecha y hora de compras anteriores, frecuencia de compra (ej. número de pedidos en último año), recencia de la última compra (cuántos días desde la última compra), valor monetario total gastado por el cliente, categorías de productos comprados, entre otros. Estos

indicadores suelen ser muy predictivos: por ejemplo, en retail es común la heurística RFM mencionada (Recency, Frequency, Monetary) para segmentar clientes. Un cliente con alta frecuencia de compra, compras muy recientes y alto gasto acumulado claramente tiene mayor propensión a seguir comprando (CleverTap, 2025). El modelo cuantificará estas relaciones de manera más precisa. Los datos provienen del sistema transaccional/ERP/CRM de Carnivery. Si existieran compras offline (tienda física), también deberían integrarse para tener la visión 360° del cliente.

- Datos de interacción con campañas previas: Información de marketing relacional histórico, por ejemplo: aperturas y clics de emails promocionales enviados, uso de códigos de descuento o cupones anteriores, si el cliente responde a anuncios en redes sociales (medido por pixel de Facebook, etc.). Esto ayuda a perfilar qué clientes son más sensibles o responsivos a las campañas. Por ejemplo, un cliente que consistentemente abre los boletines y ha usado cupones antes, podría tener alta propensión si se le envía el mensaje adecuado. Estos datos suelen residir en plataformas de automatización de marketing o en el CRM. Integrarlos mejora el poder predictivo del modelo (Delen et al., 2016 muestran que combinar múltiples técnicas y fuentes en segmentación mejora la identificación de clientes de alto potencial).
- Características demográficas o de segmento: Cualquier dato estático del cliente (residencia, edad si se conoce, segmento B2B vs B2C, etc.) o segmentación previa (como los segmentos que Carnivery haya definido: por ejemplo “clientes mayoristas” vs “clientes minoristas”, o segmentos por canal de adquisición). A veces, ciertos segmentos presentan propensiones diferentes (ej.: jóvenes tecnológicos responden más a campañas digitales que adultos mayores). Incluir variables categóricas de este tipo en el modelo puede ayudar a capturar patrones diferenciados por subgrupos. Si Carnivery cuenta con una clasificación de clientes VIP o frecuentes, se puede usar como insumo para calibrar el modelo.
- Datos externos y estacionales: Aunque no imprescindibles, se podrían incorporar variables macro como estacionalidad (mes del año, semana festiva o no) que afectan la probabilidad general de compra de todos los clientes. Por ejemplo, semanas cercanas a feriados nacionales podrían elevar la probabilidad de compra de ciertos productos (carne para parrilladas en fiestas, etc.). También, si la empresa tiene presencia en redes sociales, métricas como engagement o menciones de la marca podrían correlacionar con aumento de propensión en ciertos clientes (por ejemplo, quien sigue a Carnivery en Instagram quizás esté más propenso a comprar nuevas ofertas promocionadas allí).

Para manejar estas fuentes, Carnivery podría utilizar una data lake o data warehouse donde consolidar los datos históricos de clientes. Dado que se trata de procesamiento por lotes (no real time), se pueden explotar frameworks como Apache Spark en modo batch para procesar grandes volúmenes históricos de navegación y transacciones. Spark es un framework de procesamiento distribuido en memoria muy versátil y compatible con Python, Scala, SQL, etc., utilizado ampliamente por su velocidad y capacidad de escalar análisis. Con Spark o similar, se pueden generar las features o variables de entrada del modelo para cada cliente (por ejemplo: “número de visitas último 30 días”, “días desde última compra”, “importe total último trimestre”, “respondió campaña X (0/1)”, etc.). Estas tablas de características luego se usarían para entrenar un modelo de machine learning. El modelado en sí podría hacerse en Python usando librerías como scikit-learn, XGBoost o incluso TensorFlow/Keras para técnicas más avanzadas; Python es un lenguaje estándar en ciencia de datos por su amplio ecosistema de librerías analíticas (Pérez Galán, 2019a).

En cuanto a las plataformas de datos, la arquitectura de BI de Carnivery posiblemente integre tanto datos internos (ventas, ERP) como externos (GA, redes sociales). Este modelo de propensión sería un claro caso de uso de Big Data integrado: tomar datos de múltiples fuentes, transformarlos (ETL) para obtener información por cliente, y aplicar analítica predictiva. Herramientas de orquestación ETL podrían ser útiles (p.ej., AWS Glue, Azure Data Factory, o simplemente jobs programados de Spark/SQL). También Google BigQuery es una opción: dado que GA4 exporta allí, se puede aprovechar BigQuery ML para entrenar modelos de propensión directamente sobre los datos de Analytics (Google Cloud, 2023). En resumen, se utilizarán datos históricos multivariados (navegación, compras, interacciones, segmentos) almacenados en un entorno consolidado, y tecnologías de procesamiento batch como Spark/SQL para preparar los datos, con herramientas de machine learning supervisado para entrenar el modelo de propensión.

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

El modelo de propensión de compra tiene su aplicación directa en el área de Marketing de Carnivery, específicamente en las funciones de marketing digital, CRM y ventas que gestionan las campañas semanales. A diferencia del caso de uso anterior (integrado en la web en tiempo real), aquí el modelo opera detrás de escena para asistir en la planificación de acciones promocionales. Detallamos los departamentos involucrados y cómo participan:

- Departamento de Marketing (Inteligencia de Clientes / CRM): Es el usuario principal de las salidas del modelo. Este departamento planifica las campañas semanales (por ejemplo, envíos de newsletters con ofertas, campañas de SMS, publicidad segmentada en redes sociales). Con el modelo de propensión,

marketing puede decidir a quién dirigir cada campaña. Por ejemplo, si se va a lanzar una campaña de email con un cupón de descuento, en lugar de enviarlo a toda la base de datos (lo que puede ser costoso y hasta contraproducente), podrían enviarlo solo al 30% de clientes con mayor propensión a compra según el modelo esta semana. El equipo de CRM/marketing relacional define también umbrales y estrategias: ¿enfocamos en el top 20% propenso para un trato VIP, o hacemos varios segmentos (alto, medio, bajo) con diferentes abordajes? Marketing debe interpretar los scores del modelo y combinarlas con su conocimiento del negocio para diseñar tácticas efectivas. Este proceso requiere capacitación analítica en el equipo, para confiar en las recomendaciones del modelo e integrarlas en su toma de decisiones. En Carnivery, previamente identificaron la necesidad de segmentación avanzada para diseñar ofertas específicas a cada segmento, y este modelo provee justamente esa capacidad predictiva y segmentadora.

- Departamento Comercial/Ventas: En algunas organizaciones, marketing y ventas están separados; marketing genera demanda y ventas la concreta (por ejemplo, en B2B o ventas corporativas). En Carnivery, al ser e-commerce B2C principalmente, el “cierre” de la venta es online. Aun así, el área comercial podría involucrarse en definir qué productos o categorías empujar en las campañas según los segmentos identificados. Por ejemplo, si el modelo revela que un segmento de clientes con alta propensión no ha comprado ciertos cortes premium, ventas podría coordinar con marketing para incluir ofertas de esos productos en la campaña, aprovechando la predisposición detectada.
- Departamento de Data Science/BI: Este equipo (que puede ser parte de TI o de Marketing, según la estructura) se encarga de desarrollar y mantener el modelo de propensión. Implica actividades de data engineering (consolidar datos de diversas fuentes en un dataset de entrenamiento), data science (probar algoritmos de clasificación como regresión logística, árboles de decisión, random forest, XGBoost, redes neuronales, etc., y elegir el de mejor desempeño) y evaluación continua. También configura la periodicidad de retraining: es razonable que el modelo se reentrene quizá mensualmente con datos recientes, para capturar cambios en patrones (p. ej., un cambio de temporada puede alterar comportamientos). Este equipo colabora estrechamente con Marketing para comprender las necesidades: por ejemplo, si Marketing quiere una predicción “¿comprará en los próximos 7 días?” o “¿comprará en este mes?”, la definición del objetivo (variable dependiente) del modelo debe ajustarse. Además, deben presentarse los resultados de manera utilizable: típicamente el output será una lista de clientes con sus puntajes de propensión, cargada en el CRM o



herramienta de campaña. El área de BI podría implementar dashboards donde Marketing vea la distribución de propensiones e incluso simule resultados de campañas dirigidas. La interacción entre BI y marketing es un claro ejemplo de alineación requerida para que la analítica genere valor (Provost & Fawcett, 2013, destacan la necesidad de traducir resultados de modelos en acciones de negocio).

- Tecnología (Infraestructura de Datos): El equipo de tecnología provee el entorno para la ejecución del modelo. Si se usan herramientas big data (Spark, BigQuery, etc.), se necesita personal de sistemas que administre esos sistemas, garantice la calidad y disponibilidad de datos, y automatice los flujos (por ejemplo, programar un job semanal que calcule las propensiones antes del inicio de la semana de campaña). Si Carnivery subcontrata parte de la solución (por ejemplo, usando un servicio de la nube o consultoría externa para desarrollar el modelo), igualmente el departamento de TI interno debe asegurarse de la integración con sistemas existentes (como el CRM, donde se ejecutarán las campañas). Adicionalmente, TI se encargará de la seguridad y privacidad de los datos de clientes durante el procesamiento, cumpliendo normativas (p.ej., GDPR si aplica).
- Departamento Financiero (Control de gestión): Aunque no involucrado directamente en la creación o uso diario del modelo, finanzas puede participar validando los resultados en términos de ROI. Por ejemplo, midiendo si las campañas dirigidas con propensión efectivamente reducen el Coste por Adquisición (CPA) o aumentan el Retorno sobre la Inversión publicitaria. Pueden incluir los indicadores de eficacia de las campañas en los informes periódicos. Esto cierra el ciclo, asegurando que el proyecto de analítica está alineado con objetivos financieros de la empresa.

En resumen, Marketing/CRM es el dueño funcional del modelo de propensión, usando sus salidas para focalizar las campañas, mientras Data Science/BI y TI son facilitadores técnicos que desarrollan la solución y garantizan su funcionamiento. La colaboración es vital: si marketing define mal el problema, el modelo puede optimizar para una métrica incorrecta; si BI construye un modelo complejo sin explicarlo bien, marketing podría desconfiar en aplicarlo.

La clave es una integración fluida de este modelo en los procesos de negocio semanales de Carnivery: que se vuelva rutina revisar las predicciones al planificar cada campaña, y no un análisis académico aislado. De esta forma, el departamento de marketing pasa de decisiones basadas solo en intuición, a decisiones basadas en



datos y predicciones, alineando la organización con la filosofía de competir analíticamente (Davenport & Harris, 2017).

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

La adopción de un modelo de propensión de compra en las campañas de marketing de Carnivery conlleva expectativas claras de mejora en diversos aspectos de las actividades comerciales y de marketing:

- Mayor tasa de conversión en campañas: En lugar de enviar promociones a una masa indiferenciada (donde solo un pequeño porcentaje responde), el modelo permite enfocarse en aquellos con mayor probabilidad de comprar, lo cual aumenta la tasa de éxito. Se espera, por tanto, que las campañas semanales (por ejemplo, un emailing con oferta de fin de semana) vean más conversiones por destinatario. Si antes una campaña lograba una conversión del 2% sobre todos los contactados, dirigirse solo al público con propensión alta podría elevar esa conversión efectiva a, digamos, 5% o más. En otras palabras, habrá menos “disparos al aire” y más aciertos. Esto mejora la eficiencia y también la relación con el cliente, pues este recibe menos comunicaciones irrelevantes. Chaffey & Ellis-Chadwick (2016) enfatizan que, al alinear ofertas con preferencias del cliente, se mejora la respuesta y por ende la conversión de las campañas. La mejora esperada es entonces campañas más efectivas, medido en un aumento de la tasa de respuesta o conversión específica de cada acción de marketing.
- Incremento del Retorno de Inversión (ROI) en Marketing: Al mejorar la tasa de éxito y optimizar el uso de recursos, el ROI de las campañas debe incrementarse. Menos gasto inútil (ej. no enviar catálogos impresos o cupones a quienes seguramente no comprarán, ahorrando costos) y más ingresos generados por los clientes realmente interesados. Según reportes de la industria, las empresas que implementan la modelización de propensión han visto en promedio un incremento del 15% en el ROI de sus esfuerzos de marketing (SuperAGI, 2025). Esto ocurre porque se reduce el “derroche” de esfuerzos en clientes de baja probabilidad y se maximizan los resultados en los de alta probabilidad. En Carnivery, podríamos medir el ROI comparando, por ejemplo, la ganancia por ventas adicionales de una campaña menos el costo de la campaña (descuentos ofrecidos, costo de plataforma de emailing, etc.), y ver cómo evoluciona tras usar el modelo. La expectativa es una mejora porcentual de dos dígitos en ROI. Esto no solo significa más ingresos, sino una mejor eficiencia en la asignación del presupuesto de marketing.

- Aumento en ventas y reducción de churn a largo plazo: Un modelo de propensión bien utilizado no solo genera compras inmediatas, sino que contribuye a mayores ventas totales al reactivar clientes que de otro modo se perderían. Por ejemplo, identificar a un cliente dormido, pero con cierto interés latente (quizá navegó el sitio recientemente sin comprar) y dirigirle una oferta personalizada puede convertirlo antes de que churnee definitivamente. Estudios señalan que adoptar técnicas avanzadas de segmentación y propensión puede elevar las ventas hasta un 25% y reducir la tasa de abandono de clientes en alrededor de 30% en promedio. En Carnivery, la mejora en ventas atribuible a las campañas segmentadas podría medirse como el uplift: ventas obtenidas de segmentación vs ventas que habría con campaña tradicional. Asimismo, se esperaría ver una mejora en la retención: clientes que gracias a recibir el mensaje adecuado vuelven a comprar en vez de irse con la competencia. Un indicador a vigilar es la tasa de re-compra en periodos posteriores a la implementación del modelo, la cual debería subir. Retener clientes tiene un efecto multiplicador a largo plazo (Reichheld & Sasser, 1990 argumentan que incluso pequeños aumentos en retención generan incrementos fuertes en la rentabilidad).
- Mejor focalización y personalización de la comunicación: Aunque intangible, una mejora importante es que Carnivery estará comunicándose de forma más personalizada y relevante. Los clientes notarán que reciben ofertas que calzan con sus intereses o momento de compra. Esto puede elevar la satisfacción y la percepción de la marca. En lugar de “spam”, la comunicación será percibida como útil. Esa mejora en la relación cliente-marca se refleja en indicadores como menor tasa de bajas de suscripción a emails, mayor engagement en redes sociales, etc. Además, el departamento de marketing mejora su aprendizaje sobre el cliente: al ver qué factores predicen la compra, entienden mejor qué motiva a distintos segmentos, pudiendo afinar no solo a quién contactan sino también el mensaje. Por ejemplo, si el modelo indica que cierto segmento compra con probabilidad alta cuando ofrece envío gratis, marketing puede enfocar ese beneficio en dicho segmento.
- Toma de decisiones basada en datos (Data-driven marketing): Una mejora esperada, aunque interna, es la madurez analítica del área de marketing. Al confiar en un modelo cuantitativo para planificar campañas, se establece un ciclo virtuoso de experimentar, medir y optimizar. Esto profesionaliza la función de marketing: se pasa de decisiones subjetivas (“enviemos a todos porque sí”) a decisiones fundamentadas (“enviemos a quienes el modelo indica y midamos resultados”). Con el tiempo, esta cultura de decisiones basadas en datos mejora

no solo las campañas, sino otras áreas (pricing, surtido de productos, etc., podrían en el futuro beneficiarse de modelos similares). Davenport (2014) menciona que las empresas que compiten en analítica logran ventajas sostenibles al institucionalizar el uso de datos en la toma de decisiones.

En resumen, la mejora esperada es una ejecución de marketing más inteligente, eficaz y rentable. Carnivery debería ver más ventas por campaña, a menor costo, fortaleciendo la relación con sus clientes. Los KPI a monitorear para evidenciar estas mejoras incluyen: tasa de apertura y clic de emails, tasa de conversión por campaña, costo por adquisición (CPA), ROI de marketing, retención de clientes, valor de vida del cliente (CLV), entre otros.

Una mejora consistente en estos indicadores tras implementar el modelo confirmaría el éxito del proyecto. Por ejemplo, si el ROI de las campañas semanales sube de 200% a 230%, o si el CPA baja de \$10 a \$8 por cliente convertido, serían logros concretos atribuibles al uso del modelo de propensión. Además, como punto de referencia, más del 70% de los marketers planean aumentar el uso de segmentación y IA en los próximos años, de modo que Carnivery se posicionaría a la vanguardia de esta tendencia, obteniendo antes los beneficios asociados.

#### **4.5. Cuantificación de la mejora (Business Case)**

Para cuantificar el impacto potencial del modelo de propensión en términos de un caso de negocio, utilizaremos un ejemplo numérico simplificado, comparando el escenario tradicional vs. el escenario con modelo para una campaña semanal típica de Carnivery (por ejemplo, un envío de correo promocional):

Supuestos iniciales (escenario sin modelo):

- Carnivery cuenta con 50.000 clientes en su base de datos a los que podría dirigir una campaña (por simplicidad, consideramos todos vía email).
- Sin usar modelo, Marketing envía una promoción estándar (ej: 10% de descuento en ciertos productos) a todos los 50.000 clientes cada semana.
- Históricamente, la tasa de conversión de esas campañas masivas es del 2% (es decir, de 50.000 destinatarios, unos 1.000 terminan haciendo una compra motivada por la campaña, directa o indirectamente).
- El gasto promedio por compra resultante de la campaña es \$40 (quizás menor que el AOV general porque suelen comprar artículos en oferta).
- Por tanto, ventas generadas por campaña  $\approx 1.000 \text{ compras} * \$40 = \$40.000$ .
- Supongamos que el costo de la campaña (plataforma emailing, diseño, etc.) es \$500, y el costo de los descuentos otorgados (ese 10%) equivale a \$4.000 (es decir, 10% de \$40.000 en ventas). Total costos = \$4.500.
- El ROI de la campaña se calcularía como  $(\text{Ingresos adicionales} - \text{Costos}) / \text{Costos}$ . En este caso, ingresos adicionales \$40.000, costos \$4.500,  $\text{ROI} \approx$

788% (bastante alto, pero recordemos que no consideramos márgenes; es un ROI directo de marketing).

Ahora veamos el escenario con modelo de propensión:

- El modelo analiza a los 50.000 clientes y determina, por ejemplo, que 15.000 clientes tienen alta propensión (digamos “probabilidad de compra > 30% en próximas semanas”), 20.000 media (5-30%) y 15.000 baja (<5%). Marketing decide enfocar la campaña solo en el segmento de alta propensión (15k) y medio (20k), excluyendo el de baja probabilidad para ahorrar costos y evitar desgaste. Es decir, se envía la promoción a 35.000 clientes en total en lugar de 50k.
- Gracias a la focalización, la tasa de conversión mejora porque los receptores eran más proclives. Supongamos que el segmento de alta propensión convierte al 5% (de 15k, 750 compras) y el medio al 2% (de 20k, 400 compras). En total, tendríamos 1.150 compras por campaña. A primera vista, es un 15% más que las 1.000 compras del escenario sin modelo.
- Además de más conversiones, es posible que al personalizar mejor el mensaje para cada segmento (por ejemplo, ofertando productos que sabemos que les interesan según su historial), el gasto promedio por compra aumente ligeramente. Pongamos que sube a \$45 por compra (clientes más receptivos podrían aprovechar más la oferta). Entonces las ventas generadas serían  $1.150 * \$45 = \$51.750$ .
- Costos de la campaña con modelo: Se envió a 35k en lugar de 50k emails, reduciendo algunos costos (quizá ahorros marginales en la plataforma de emailing, digamos \$50 menos, quedando \$450). El costo de descuentos es 10% de  $\$51.750 = \$5.175$ . Sumando, costos  $\approx \$5.625$ .
- ROI con modelo:  $(\$51.750 - \$5.625) / \$5.625 = 820\%$  aproximadamente. Esto representa una mejora de ROI respecto al 788% anterior. Más importante aún, la ganancia neta (ingresos menos costos de campaña) pasó de \$35.500 a \$46.125, que es \$10.625 adicionales de beneficio por campaña gracias al modelo. En una base semanal, eso anualizado sería  $\sim \$552k$  más de ganancia al año, solo optimizando los envíos.

- Ahorro y beneficio en clientes no contactados: Además, al no bombardear a 15.000 clientes de baja propensión, evitamos potencialmente que se saturen o molesten por comunicaciones irrelevantes. Eso puede repercutir en una menor tasa de opt-out (bajas de suscripción) y mejor preservación de la base de datos a largo plazo, aunque este beneficio es cualitativo. Si asumimos que sin ser contactados algunos igual hubieran comprado (digamos 100 compraban espontáneamente), pudimos omitirlos sin impacto significativo, y en todo caso podríamos en futuros análisis dirigirles campañas específicas para reactivación cuando sus probabilidades mejoren.
- Segmento reactivo (bonus): Imaginemos que entre los de alta propensión había un subgrupo de 5.000 clientes “VIP” con probabilidad altísima (>50%) a los que se les envió, además del email, un SMS personalizado. Este microsegmento convirtió al 10% (500 compras). Estas son tácticas posibles una vez tienes identificados a los mejores clientes. El costo del SMS es mayor, pero si el CLV de esos VIP es alto, vale la pena. La utilidad del modelo es que permite diferenciar esfuerzos por segmento, gastando más donde habrá más retorno.

Los números anteriores ilustran cómo, con el modelo de propensión, Carnivery podría ganar sustancialmente más con cada campaña y gastar un poco menos, logrando un marketing más rentable. En el ejemplo, un +15% en conversiones y +~28% en ingresos por campaña, con -30% en volumen de contactos, es un resultado muy positivo. Incluso si nuestras suposiciones son optimistas, la dirección de la mejora es consistente con casos reales: según datos publicados, compañías que implementaron modelos de propensión reportaron incrementos de ventas del orden de 20-25% y reducciones importantes en gastos innecesario.

Además, un aumento promedio del ROI de marketing del 15% se ha observado en diversas industrias tras adoptar la analítica predictiva (SuperAGI, 2025).

Otro aspecto cuantitativo es la reducción del churn. Supongamos que mensualmente Carnivery perdía 200 clientes (no volvían a comprar en un periodo prolongado). Aplicando modelos predictivos, se logra retener un 10% de esos que se iban (20 clientes salvados al mes) mediante ofertas proactivas. Si cada cliente tiene un CLV futuro estimado de \$300, estamos reteniendo \$6.000 de valor al mes que antes se perdía (~\$72k al año). Esto sumado al efecto de campañas más productivas puede generar un impacto anual de cientos de miles de dólares.

Por último, consideremos el costo de implementar el modelo. Puede implicar contratar a un científico de datos, infraestructura de datos y software. Estimemos un costo anual de \$50.000 (entre personal y tecnología). Aun así, con las mejoras

cuantificadas (ejemplos: +\$10k por campaña x 52 semanas = +\$520k, +\$72k por retención = \$592k total beneficio anual posible), el retorno sobre esa inversión es altísimo (casi 10 veces). Incluso con escenarios más conservadores, el modelo se pagaría por sí solo rápidamente.

En conclusión, el business case para el modelo de propensión muestra beneficios tangibles: más ingresos por campaña, mayor eficiencia en gasto de marketing y valor retenido de clientes, que en conjunto superan ampliamente el costo de desarrollo. No menos importante, Carnivery ganará agilidad y precisión en su estrategia comercial, pudiendo responder semana a semana con acciones basadas en datos. Este enfoque analítico posicionará a la empresa para seguir optimizando otras áreas con modelos similares (por ejemplo, propensión a churn, recomendación de productos en campañas, optimización de precios, etc.), conformando un ecosistema de Big Data que genera ventajas competitivas. En la economía digital actual, compañías que aprovechan datos y modelos predictivos en marketing logran marketing más ágil y efectivo, y Carnivery estaría dando un paso firme en esa dirección con este caso de uso.

## **CAPITULO 5 MODELO DE RECOMENDACIÓN PERSONALIZADA EN TIEMPO REAL**

### **5.1. Objetivo del modelo**

El presente modelo es un modelo de aprendizaje automático en tiempo real que identifique al usuario cuando ingresa al e-commerce y le sugiera de inmediato dos posibilidades: repetir su pedido anterior o recibir una recomendación personalizada de productos, basada en su historial reciente de navegación. El objetivo principal de este modelo es mejorar la experiencia y satisfacción del cliente a través de la personalización, aumentando así la probabilidad de conversión (es decir, que el visitante realice una compra) y fomentando la fidelización. En el contexto de Carnivery, esto significa que, al reconocer a un cliente recurrente, el sistema podría agilizar su proceso de compra recordándole sus productos habituales, o bien presentarles nuevos productos relevantes a sus preferencias actuales.

La literatura respalda que las recomendaciones personalizadas pueden incidir fuertemente en el desempeño de un e-commerce. Por ejemplo, Amazon atribuye aproximadamente un 35% de sus ingresos a su motor de recomendaciones (McKinsey, 2013, citado en Arsenault, 2022), evidenciando cómo una experiencia de compra adaptada a cada usuario puede impulsar las ventas. Asimismo, las recomendaciones pertinentes aumentan la probabilidad de compra al poner frente al cliente productos de su interés en el momento oportuno, transformando visitantes ocasionales en compradores leales (Chaffey & Ellis-Chadwick, 2016). En resumen, el objetivo del modelo es personalizar la interacción en tiempo real para cada usuario de Carnivery, con el fin de elevar la tasa de conversión y la retención de clientes a través de sugerencias oportunas y relevantes.

### **5.2. Fuentes de datos utilizadas**

Para entrenar y alimentar este modelo de recomendación en tiempo real, se requiere integrar múltiples fuentes de datos, tanto históricas como de streaming. Las principales fuentes de datos incluyen:

- **Historial de pedidos del cliente:** Los registros transaccionales del sistema de ventas de Carnivery (por ejemplo, en su base de datos de órdenes o ERP) proporcionan información sobre compras previas de cada usuario: qué productos compró, en qué cantidades y fechas, etc. Estos datos permiten identificar el “último pedido” de un cliente para ofrecerle una repetición rápida, y también extraer patrones de preferencia individual (p. ej., si suele comprar cortes de cerdo vs. mariscos, productos premium vs. estándar, etc.). Son datos estructurados internos que se actualizan con cada transacción.



- **Actividad de navegación web reciente:** Los eventos de navegación del usuario en el sitio (páginas de producto vistas, búsquedas realizadas, productos añadidos o eliminados del carrito, tiempo de permanencia, etc.) proveen señales en tiempo real sobre sus intereses actuales. Estas interacciones se pueden capturar mediante herramientas de analítica digital como Google Analytics 4, que registra el comportamiento del usuario en la web/app. Carnivery podría aprovechar la integración nativa entre GA4 y BigQuery de Google Cloud para exportar estos eventos casi en tiempo real al ecosistema de datos. Alternativamente, los propios logs del servidor web o eventos capturados por scripts personalizados (por ejemplo, utilizando Pixel o APIs de analítica) pueden alimentar un flujo de datos de navegación al sistema de recomendaciones. Esta fuente de datos es semiestructurada (p. ej., JSON) y de alta frecuencia, esencial para reflejar el contexto reciente del usuario.
- **Información de catálogo de productos:** Datos maestros de productos (categorías, precios, stock disponible, descripciones, etc.) almacenados en la base de datos de la tienda o en un sistema PIM (Product Information Management). Aunque indirectamente relacionados con el usuario, estos datos ayudan a filtrar y diversificar las recomendaciones (por ejemplo, evitando sugerir ítems fuera de stock, o relacionando productos complementarios).
- **Perfil y segmentación del cliente:** Cualquier información adicional del cliente almacenado en CRM o bases de datos de usuarios, como segmentación demográfica (ej. ubicación geográfica, tipo de cliente B2C/B2B), valor de vida del cliente estimado, o segmentos RFM (Recency, Frequency, Monetary). En Carnivery, podrían definirse segmentos según recencia de última compra, frecuencia de compra y valor monetario acumulado (enfoque RFM) para distinguir clientes “leales” de “ocasionales” (CleverTap, 2025). Estos datos permiten modular la lógica de recomendación (por ejemplo, ofreciendo repetir el último pedido solo a clientes frecuentes, o promociones especiales a clientes de alto valor).
- **Fuentes de datos externas en tiempo real:** Opcionalmente, el modelo podría incorporar señales externas contextuales, como tendencias generales (p. ej., un aumento global en búsquedas de cierto corte de carne por alguna festividad) o información meteorológica/local que influya la demanda (por ejemplo, clima frío podría correlacionar con ventas de cortes para sopa). No obstante, el núcleo del caso de uso se basa en datos propios de comportamiento del cliente en la plataforma.



En cuanto al almacenamiento y acceso a estos datos, Carnivery podría apoyarse en tecnologías de Big Data. Por ejemplo, los eventos de navegación podrían enviarse a un sistema de mensajería en streaming como Apache Kafka, que alimente un proceso de análisis en tiempo real. Los datos históricos de pedidos y perfil de clientes podrían residir en un Data Warehouse relacional o en un repositorio NoSQL como MongoDB, el cual es una base de datos orientada a documentos muy utilizada por su flexibilidad para manejar datos semiestructurados de clientes (Pérez Galán, 2019). La información consolidada podría almacenarse en caché en una base NoSQL en memoria (ej. Redis) para responder en milisegundos a cada visita (Pérez Galán, 2019). De esta manera, el modelo dispondrá de datos actualizados al instante que combinen el histórico (lo que el cliente ha comprado antes) con el contexto actual (lo que está mirando ahora), fundamentales para generar recomendaciones precisas.

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

El área de aplicación de este modelo es principalmente la plataforma de comercio electrónico de Carnivery – es decir, su sitio web (y eventualmente aplicación móvil si existiera) donde interactúan los clientes. Es en este front-end digital donde se materializan las recomendaciones en tiempo real: por ejemplo, mediante un banner o ventana emergente que al reconocer al usuario diga “¡Bienvenido de nuevo! ¿Desea repetir su último pedido?” junto a un botón rápido, o “Recomendado para ti” mostrando productos personalizados en la página de inicio o de categoría.

Varios departamentos de negocio se ven implicados en el desarrollo e implementación de esta solución de personalización:

- Departamento de Marketing/Ventas: Es el principal impulsor, ya que busca incrementar la conversión y fomentar compras recurrentes mediante la personalización. Marketing define las reglas de negocio (p. ej., cuándo ofrecer la repetición del pedido vs. una recomendación distinta) y colabora en la selección de qué productos promocionar. Como señalan Chaffey & Ellis-Chadwick (2016), el área de marketing requiere analizar patrones de consumo y preferencias de clientes para diseñar campañas y ofertas dirigidas; en este caso, el “ofertar” en tiempo real dentro de la web es parte de esa estrategia de marketing personalizada. Asimismo, el equipo de marketing digital utilizará este modelo para mejorar la experiencia del usuario en la tienda en línea, haciendo que cada cliente sienta la plataforma adaptada a sus necesidades (lo cual puede traducirse en mayor satisfacción y lealtad).
- Departamento de Tecnología/Ingeniería (TI): Es responsable de la implementación técnica. Incluye equipos de desarrollo web (que integrarán

las recomendaciones en la interfaz de usuario del e-commerce), científicos de datos o machine learning engineers (que construyen y entrenan el modelo de recomendación) y arquitectos de datos (que diseñan la infraestructura para capturar, procesar y servir los datos en tiempo real). Por ejemplo, se requerirá coordinar la recolección de datos de navegación (posiblemente instrumentando Google Analytics o eventos propios), el entrenamiento del modelo con datos históricos (que podría realizarse en Python usando librerías de ciencia de datos como scikit-learn o frameworks de deep learning), y la puesta en producción del modelo. El despliegue en producción puede lograrse mediante un servicio en la nube que haga inferencias en milisegundos; Carnivery podría considerar herramientas como TensorFlow Serving si utiliza modelos neuronales, dada su integración con Google Cloud. El equipo de TI también asegurará que la latencia sea baja y la disponibilidad alta, para no afectar la navegación del usuario.

- Departamento de Experiencia de Cliente/Atención al Cliente: Indirectamente involucrado, en tanto una recomendación acertada contribuye a una mejor experiencia y satisfacción. Aunque este modelo actúa antes de que el cliente contacte con soporte, el área de atención al cliente puede proveer feedback sobre la recepción de estas recomendaciones (por ejemplo, si los clientes mencionan que la sugerencia de repetir pedido les resulta útil). Además, si el modelo induce a más compras recurrentes, la atención al cliente deberá gestionar menos casos relacionados a dificultades en encontrar productos, etc., y más bien enfocarse en valor añadido. En Carnivery, la fidelización es una meta transversal: atención al cliente busca fortalecer la lealtad ofreciendo soluciones personalizadas, y el modelo en cuestión es una herramienta para anticiparse a las necesidades del cliente antes de que tenga que solicitarlas.
- Operaciones y Logística: Este departamento no participa directamente en la generación de la recomendación, pero podría verse impactado por sus resultados. Un aumento en pedidos (o ciertos cambios en el mix de productos vendidos debido a recomendaciones) implicará que logística e inventarios estén preparados para responder a la demanda. Por ejemplo, si el modelo recomienda con frecuencia cierto producto complementario (p. ej., una salsa para acompañar el cerdo) y muchos clientes comienzan a agregarlo, el equipo de operaciones deberá asegurarse de contar con suficiente stock y ajustar la cadena de suministro. Por lo tanto, es importante que el área de Operaciones esté al tanto de las iniciativas de personalización en tiempo real, para alinear la oferta y la disponibilidad.

En resumen, Marketing/Ventas lidera los requisitos funcionales (mejorar conversión, upselling), TI/Data Science lidera la construcción técnica, Experiencia del Cliente aporta retroalimentación cualitativa sobre la aceptación, y Operaciones monitorea impactos en la demanda. Esta naturaleza multidisciplinaria es característica de los proyectos de Big Data y BI modernos, donde la colaboración entre áreas de negocio y técnicas resulta esencial para el éxito (Davenport & Harris, 2017).

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

La implementación de un modelo de recomendación en tiempo real en Carnivery espera lograr mejoras sustanciales en varios indicadores clave del negocio:

- Tasa de conversión más alta: Una recomendación o recordatorio personalizado disminuye la fricción en el proceso de compra y dirige al usuario hacia productos de su interés, aumentando la probabilidad de que realice una compra en esa visita. En términos cualitativos, la experiencia “se siente” más útil y eficiente para el cliente, quien encuentra más rápidamente lo que podría querer. Desde el punto de vista cuantitativo, una tasa de conversión elevada es señal de que las ofertas están alineadas con las preferencias de los clientes (Chaffey & Ellis-Chadwick, 2016). Se espera entonces que, tras introducir el modelo, la proporción de usuarios que compren en su visita crezca. Por ejemplo, si actualmente compran 2 de cada 100 visitantes (2%), quizás con recomendaciones relevantes se pueda subir a 3 de cada 100 (3%) o más. Esta mejora indicaría que la estrategia analítica está fomentando más ingresos y fidelización por visitante.
- Incremento del valor promedio de pedido (AOV): Además de convertir más visitas en ventas, el modelo puede aumentar cuánto gasta cada cliente por compra. Esto puede ocurrir de dos modos: (a) Si se sugiere repetir el último pedido, un cliente podría agregar todos esos ítems rápidamente al carrito, incluyendo productos que quizá hubiera olvidado, elevando el ticket. (b) Con recomendaciones de upselling/cross-selling (“también te podría gustar...”), se incentiva al cliente a añadir productos complementarios. Una estrategia típica es recomendar productos frecuentemente comprados juntos (por ejemplo, si el cliente está comprando un corte de carne especial, sugerirle una especie o salsa gourmet). Estas técnicas han demostrado elevar el valor de carrito en e-commerce en general. En Carnivery, un AOV mayor implica mejor rentabilidad; por ende, se espera que el modelo eleve progresivamente este indicador. De hecho, ofrecer productos complementarios ajustados al perfil del usuario tiene un impacto directo en la rentabilidad del negocio (García, 2018), al maximizar los ingresos por cliente.

- Mejoras en retención y lealtad del cliente: Una de las ventajas a largo plazo de personalizar la experiencia es que los clientes se sienten mejor atendidos y tienden a volver con más frecuencia. Si Carnivery facilita que un cliente recurrente repita su compra habitual en segundos, está proporcionando conveniencia, lo cual refuerza la fidelidad. De igual forma, sorprender al cliente con recomendaciones relevantes puede generar una grata impresión y diferenciación frente a competidores. Un aumento en la tasa de retención (clientes que realizan compras repetidas en lugar de abandonar) sería una evidencia de éxito. Retener clientes es crucial para ingresos recurrentes y reducción de costos de adquisición (Reichheld & Sasser, 1990), por lo que se espera que el modelo contribuya a mejorar este KPI. Por ejemplo, podría medirse que la frecuencia de compra promedio por cliente en un año pase de, digamos, 4 a 5 compras gracias a las recomendaciones y facilidades brindadas.
- Satisfacción del cliente y diferenciación competitiva: Aunque más difícil de medir de forma inmediata, se espera que la satisfacción (reflejada en encuestas NPS o en menor tasa de abandono de carrito) aumente. Un sitio inteligente que “recuerda” al cliente y le ahorra tiempo crea un vínculo positivo. En un mercado digital competitivo donde la personalización es clave, invertir en este modelo de machine learning posicionará a Carnivery como una empresa centrada en el cliente y en la innovación (Chaffey & Ellis-Chadwick, 2022).

En síntesis, la mejora esperada con este desarrollo es multidimensional: más ventas (por más conversiones y mayor ticket medio), mayor lealtad (por retención) y mejor percepción de la marca. Estas mejoras están alineadas con los objetivos comerciales de Carnivery de incrementar ingresos y participación de mercado a través de una gestión basada en datos. Implementar un motor de recomendaciones en tiempo real es, en esencia, llevar la estrategia tradicional de ventas cruzadas del dependiente de tienda física a la escala y velocidad del mundo digital, obteniendo beneficios tanto para la empresa (ingresos, eficiencia) como para el cliente (comodidad, relevancia) (Davenport, 2014).

### **5.5. Cuantificación de la mejora (Business Case)**

Para ilustrar cuantitativamente el impacto potencial de este modelo en el negocio de Carnivery, presentamos un breve business case. Supongamos los siguientes números base de referencia para Carnivery antes de implementar el modelo de recomendación:

- Tráfico mensual al sitio: 100.000 visitas (sesiones de usuarios).

- Tasa de conversión actual (visitantes que compran): 2.0% (es decir, 2 de cada 100 visitas terminan en una compra).
- Número de pedidos mensuales actual: 2.000 (derivado de  $100.000 \times 2\%$ ).
- Ticket o valor promedio por pedido (AOV) actual: \$50 (dólares, o su equivalente en moneda local).
- Ingresos mensuales actuales: \$100.000 (2.000 pedidos \* \$50 cada uno).
- Ahora, proyectemos las mejoras con la solución de recomendaciones personalizadas:
- Escenario de conversión mejorada: Tras implementar el modelo en tiempo real, Carnivery logra incrementar la tasa de conversión a 2.5%. Este es un aumento absoluto de 0.5 puntos porcentuales, que en términos relativos significa un 25% más conversiones que antes (pasando de 2% a 2.5%). Dicha mejora es plausible dado que las recomendaciones personalizadas hacen más eficiente la compra; de hecho, en líderes del sector se han observado incrementos notables: Amazon ha reportado conversiones de recomendación on-site de hasta un 60% en ciertos casos (Mulpuru, citado en Arsenault, 2022), aunque en promedio nuestro caso sería más modesto. Con 100.000 visitas mensuales, una conversión de 2.5% generaría 2.500 pedidos al mes (500 adicionales respecto al escenario actual).
- Escenario de ticket promedio mejorado: El modelo también fomenta que los usuarios agreguen más ítems (por cross-selling) o repitan pedidos más grandes. Conservadoramente, asumamos que el AOV sube de \$50 a \$55 (+10%) gracias a las recomendaciones (por ejemplo, un cliente añade una salsa de \$5 sugerida junto con su carne). Este aumento está respaldado por la efectividad de sugerir artículos complementarios.
- Cálculo de incremento en ingresos: Combinando ambos efectos, en el escenario post-implementación Carnivery tendría ~2.500 pedidos al mes con un ticket medio de \$55. Eso resultaría en \$137.500 de ingresos mensuales, comparado con \$100.000 antes de la personalización. En otras palabras, se obtendrían \$37.500 adicionales por mes, lo que es un +37.5% en facturación mensual. Anualmente, esta mejora representa \$450.000 adicionales en ingresos, un impacto significativo para la empresa.
- Rentabilidad y costos: Suponiendo que el margen neto sobre las ventas es, digamos, 20%, esa mejora anual de \$450k en ventas implicaría cerca de \$90k más de beneficio anual neto atribuible al modelo. Habría que contrastarlo con la inversión y costos operativos del proyecto (costos de desarrollo, infraestructura en la nube para servir el modelo en tiempo real, etc.). Si, por ejemplo, implementar y mantener el sistema costase \$30k anuales, el business case aún muestra un ROI positivo muy alto. En otras palabras, por

cada dólar invertido en la iniciativa de recomendación personalizada, retornarían varios dólares en ventas incrementales.

- Mejora en retención de clientes: Adicionalmente, podríamos cuantificar beneficios en retención. Supongamos que antes un 25% de los nuevos clientes realizaban al menos una segunda compra (tasa de retención de primer ciclo), y que, con la mejora en la experiencia y las facilidades para reordenar, esa tasa sube a 30%. Si Carnivery capta 1.000 clientes nuevos al mes, esto significa 50 clientes más retenidos que antes cada mes. Asignando un valor de vida del cliente (CLV) promedio de \$200, esos 50 clientes adicionales aportan \$10.000 extra a futuro. Sumado a la mejora en ventas inmediatas, se fortalece el valor total generado.

Cabe destacar que estos cálculos son estimaciones para construir el caso de negocio, y los resultados reales podrían variar. No obstante, son coherentes con hallazgos de la industria: las empresas que explotan la personalización y analítica de clientes suelen lograr incrementos de dos dígitos en ingresos y eficiencia. Un estudio de McKinsey señala que la personalización puede impulsar las ventas entre un 10% y 15% en retail digital, y en el caso paradigmático de Amazon, el motor de recomendaciones generó alrededor de un tercio de sus ventas. Así, el modelo de Carnivery podría justificarse plenamente si logra siquiera una fracción de ese impacto.

Finalmente, más allá del impacto financiero inmediato, existe un beneficio difícil de cuantificar, pero muy valioso: inteligencia sobre el cliente. Al desplegar este modelo, Carnivery acumulará conocimiento sobre qué recomendaciones funcionan, qué productos se piden recurrentemente, qué combos prefieren ciertos segmentos, etc. Ese feedback loop permite refinar no solo el algoritmo sino también las estrategias de negocio (gestión de inventarios, diseño de promociones, etc.), generando una ventaja competitiva sostenible (Provost & Fawcett, 2013). En suma, el business case del modelo de recomendación en tiempo real muestra mejoras sustanciales en ingresos y rentabilidad, justificando la inversión tecnológica al proveer un retorno financiero positivo y consolidar la lealtad de la clientela de Carnivery.

## **Conclusión**

Tanto el modelo de recomendación en tiempo real como el modelo de propensión de compra ilustran cómo las técnicas de Data Science y Machine Learning pueden integrarse en la estrategia de un e-commerce para impulsar resultados. El caso de la recomendación en tiempo real evidencia beneficios inmediatos en la experiencia del cliente y las ventas diarias, apoyándose en tecnologías de streaming (Kafka, Spark Structured Streaming) y bases de datos NoSQL de baja latencia (MongoDB, Redis) para ofrecer personalización instantánea. Por otro lado, el caso de propensión de

compra batch demuestra el poder de la analítica predictiva en la optimización de procesos de marketing, usando el historial masivo de datos con herramientas de big data (Spark, BigQuery) y algoritmos de machine learning en Python para mejorar la toma de decisiones semanal. En ambos casos, la justificación tecnológica –ya sea el uso de Python y sus librerías para modelado avanzado, Google Analytics como fuente de datos de comportamiento, o frameworks como Apache Spark para escalabilidad– va de la mano con la justificación de negocio: mejorar indicadores clave (KPIs) y generar ROI positivo.

Carnivery, al implementar estas soluciones, esperaría no solo mejoras cuantificables (más conversiones, mayor valor por cliente, ROI de marketing superior), sino también desarrollar capacidades internas valiosas: una cultura más data-driven, infraestructuras de datos modernizadas y equipos multifuncionales colaborando en torno a la inteligencia de negocios. Esto sienta las bases para futuros proyectos de BI y Big Data aún más sofisticados, en un ciclo de mejora continua. Como indican Provost & Fawcett (2013), la ventaja competitiva sostenible proviene de saber extraer conocimiento útil de los datos y actuar sobre él – exactamente lo que estos dos casos de uso persiguen lograr en Carnivery.

En suma, mediante un lenguaje formal apoyado en fundamentos académicos y referenciado en bibliografía, hemos identificado los objetivos, datos, ámbitos de aplicación, mejoras esperadas y cuantificación de impacto para dos casos de uso representativos. Ambos constituyen business cases sólidos de cómo la ciencia de datos puede integrarse con éxito en una empresa de comercio electrónico para potenciar su rendimiento comercial y su orientación al cliente, haciendo tangible la promesa del Big Data y la Inteligencia de Negocio en resultados empresariales concretos.



## CONCLUSIONES Y APLICACIONES

La presente tesis se enfoca en el análisis del comportamiento del cliente y la personalización de ofertas en el ámbito del comercio electrónico, ejemplificado a través del caso de estudio de Carnivery, una empresa dedicada a la venta de productos cárnicos. A lo largo del trabajo, se ha puesto de manifiesto que la integración de estrategias digitales resulta fundamental para la mejora de la experiencia del cliente, la fidelización de la clientela y el incremento de las ventas. A través de la aplicación de técnicas de análisis del comportamiento del consumidor, se logra identificar variables significativas que influyen en el proceso de decisión de compra, lo que permite la formulación de estrategias de marketing más efectivas y orientadas.

Un aspecto destacado del estudio es la implementación de modelos predictivos para la segmentación de clientes en función de su comportamiento de compra. Mediante el uso de datos históricos, se pueden prever las propensiones de compra de cada cliente y personalizar las ofertas en función de sus preferencias individuales. Esta metodología no solo contribuye a mejorar la tasa de conversión, sino que también optimiza la asignación de recursos en las campañas de marketing, dirigiendo los esfuerzos hacia aquellos clientes que presentan un mayor potencial de respuesta.

La creación de un ecosistema de Business Intelligence (BI) en Carnivery se ha planteado como una solución integral destinada a la gestión y el análisis de datos. A través de la implementación de una data warehouse y herramientas de visualización, se busca democratizar el acceso a la información en toda la organización. Esta estrategia permite a diferentes departamentos, como marketing y ventas, fundamentar sus decisiones en datos precisos, fomentando así una mayor agilidad y efectividad en las operaciones comerciales. De esta manera, se promueve una cultura organizacional orientada a la analítica y al uso de información basada en evidencias.

Asimismo, el empleo de técnicas de Big Data y aprendizaje automático ha sido propuesto como un medio para optimizar procesos y mejorar la experiencia del cliente. La personalización de la interacción en tiempo real, mediante recomendaciones basadas en el historial de navegación y compras, se presenta como una estrategia efectiva para elevar la satisfacción del cliente. Este enfoque no solo facilita el proceso de compra, sino que también potencia la lealtad, al hacer que cada consumidor se sienta valorado y comprendido, lo que es esencial en un entorno competitivo.

A través de la investigación, se ha evidenciado que la implementación de modelos de propensión y recomendación en tiempo real puede generar un impacto



significativo en los ingresos de la empresa. Al dirigir campañas de marketing a segmentos específicos y ofrecer recomendaciones personalizadas, se anticipa un aumento en la tasa de conversión y en el valor promedio de los pedidos. Estos resultados no solo contribuyen a mejorar la rentabilidad, sino que también refuerzan la relación entre la marca y sus clientes, transformando a Carnivery en un referente en el ámbito del comercio electrónico cárnico.

La aplicación de los conceptos desarrollados en un entorno concreto, como el de Carnivery, ilustra la viabilidad y efectividad de adoptar un enfoque basado en el análisis de datos. A medida que el comercio electrónico sigue en expansión, las organizaciones que integran análisis avanzados y personalización en sus estrategias estarán en una mejor posición para competir en un mercado cada vez más saturado. La tesis subraya la importancia de adaptarse a las tendencias del consumidor y de utilizar la tecnología para mejorar continuamente la oferta y la experiencia de compra.

En última instancia, se concluye que el desarrollo de capacidades analíticas y la promoción de una cultura de toma de decisiones fundamentadas en datos son esenciales para el crecimiento sostenible de empresas como Carnivery. La investigación proporciona no solo un marco teórico, sino también un enfoque práctico que beneficia a la organización en términos de eficiencia y efectividad, al tiempo que contribuye a la literatura existente en el campo de la administración de empresas y el comercio electrónico, ofreciendo un modelo que puede ser replicado en otras industrias.

**Tabla Comparativa “antes vs después de implementar business intelligence”**

<b>Aspecto Evaluado</b>	<b>Situación Actual (Antes del Proyecto)</b>	<b>Situación Proyectada (Después de Implementar BI)</b>
<b>Gestión de la información</b>	Datos dispersos entre varias plataformas sin conexión ni validación automática.	Datos consolidados en un sistema único con acceso en tiempo real.
<b>Toma de decisiones</b>	Basada en intuición o reportes manuales generados por la gerencia.	Basada en indicadores claros (KPIs) visibles en dashboards interactivos.
<b>Conocimiento del cliente</b>	No se identifica quiénes son los clientes más rentables o fieles.	Se segmentan clientes según su frecuencia de compra, ticket promedio y recompra.
<b>Eficiencia operativa</b>	Procesos de carga manual y falta de alertas automáticas.	Procesos automatizados con actualizaciones diarias y reducción de errores.
<b>Marketing y ventas</b>	Promociones generales, sin segmentación personalizada.	Ofertas automáticas y personalizadas según hábitos de compra.
<b>Competitividad digital</b>	Uso básico de herramientas digitales.	Adopción de un modelo de gestión basado en analítica e inteligencia de negocios.
<b>Cultura organizacional</b>	Dependencia de una sola persona para manejo de información (CEO).	Creación de un equipo de análisis con roles definidos y capacitación técnica.

## REFERENCIAS

- Belova, K. (2024). 15 best big data analytics tools and platforms to look out for in 2025. PixelPlex. <https://pixelplex.io/blog/best-big-data-analytics-tools/> pixelplex.io
- Carnivery. (2025). Comunicación personal, reunión de levantamiento de información con Marco [Datos no publicados].
- Davenport, T. H., & Harris, J. G. (2017). *Competing on Analytics: The New Science of Winning*. Harvard Business Review Press.
- Céspedes, Diego. Entrevista. "ENTREVISTA DIEGO CÉSPEDES PBL." Teams, 23 de septiembre de 2025. <https://app.read.ai/analytics/meetings/01K5W5P9C0Y3695WZ4SJKKA0C5> (minuto 10:05).
- Chaffey, D., & Ellis-Chadwick, F. (2016). *Digital marketing* (6<sup>a</sup> ed.). Pearson Education Limited. <https://doi.org/10.13140/RG.2.2.25371.45600>
- Chaffey, D., & Ellis-Chadwick, F. (2022). *Digital Marketing* (8.<sup>a</sup> ed.). Pearson Education.
- CleverTap. (2025, 18 de marzo). *What is RFM customer segmentation in e-commerce?* CleverTap Blog. <https://clevertap.com/blog/rfm-ecommerce/>
- Davenport, T. H. (2014). *Big data at work: Dispelling the myths, discovering the opportunities*. Harvard Business Review Press.
- Davenport, T. H., & Harris, J. G. (2017). *Competing on Analytics: The New Science of Winning*. Harvard Business Review Press.
- Delen, D., Guneri, A., & Kaya, S. (2016). Customer segmentation using data mining techniques. *Procedia Computer Science*, 111, 269-276. <https://doi.org/10.1016/j.procs.2017.07.204>
- Full Scale. (2024). *Java vs. Python decoded: 9 crucial factors for informed tech decisions*. Recuperado de <https://fullscale.io/blog/java-vs-python/> fullscale.io
- García, L. (2018). Gestión y procesamiento de datos en un data warehouse. Documento interno.
- García, R. (2018). Big data y análisis de datos en la toma de decisiones empresariales. *Revista de Ingeniería Industrial*, 31(3), 213-228. <https://doi.org/10.15446/rcii.v31n3.67769>

- García, R. (2018). **Big data y análisis de datos en la toma de decisiones empresariales**. *Revista de Ingeniería Industrial*, 31(3), 213–228. <https://doi.org/10.15446/rcii.v31n3.67769>
- Google Cloud. (2023). *Google Analytics (GA4) and BigQuery for marketing and audience insights*. Google Cloud – Use Cases. Recuperado de <https://cloud.google.com/use-case/google-analytics-bigquery>  
cloud.google.com
- Guevara, C. (2025, 15 de julio, 09:00 h). Comunicación personal. Entrevista realizada en la reunión “REUNION PBL- TESIS DE MAESTRIA UIDE”.
- Kotler, P., & Keller, K. L. (2016). *Dirección de marketing* (14a ed.). Pearson Education.
- Pérez Galán, J. L. (2019a). *Business Intelligence y Big Data – Lectura Semana 8*. EIG Business School. (Contenido académico sobre tecnologías Big Data)
- Pérez Galán, J. L. (2019b). *Business Intelligence y Big Data – Lectura Semana 8 (Herramientas Open Source)*. EIG Business School. (Contenido académico sobre Hadoop, Spark, NoSQL, etc.)
- Provost, F., & Fawcett, T. (2013). *Data science for business: What you need to know about data mining and data-analytic thinking*. O'Reilly Media.
- Pykes, K. (2022). *PyTorch vs TensorFlow vs Keras para el aprendizaje profundo: Guía comparativa*. DataCamp. Recuperado de <https://www.datacamp.com/es/tutorial/pytorch-vs-tensorflow-vs-keras>  
datacamp.comdatacamp.com
- Reichheld, F. F., & Sasser, W. E. (1990). Zero defections: Quality comes to services. *Harvard Business Review*, 68(5), 105–111. <https://doi.org/10.1177/0972262916677365>
- Reynolds, B. (2025). Choosing the right cloud data warehouse: A deep dive into Amazon Redshift vs competitors. Baytech Consulting. <https://www.baytechconsulting.com/blog/amazon-redshift-a-comprehensive-analysis-2025> baytechconsulting.com
- Reynolds, B. (2025). Google BigQuery uncovered: Architecture, features, and strategic comparisons. Baytech Consulting. <https://www.baytechconsulting.com/blog/google-bigquery-2025> baytechconsulting.com
- Reynolds, B. (2025). Inside Azure Synapse Analytics: Capabilities, competitive edge, and when to use it. Baytech Consulting.

<https://www.baytechconsulting.com/blog/azure-synapse-analytics-2025>  
baytechconsulting.com

Rodríguez, H. (2019). Optimización de flujos de información para e commerce. Documento interno.

Rodríguez, M. (2019). **Transformación digital en pequeñas y medianas empresas: Caso de estudio en comercio electrónico.** *Revista Digital de Innovación y Tecnología*, 12(2), 45–59. <https://doi.org/10.17583/rdit.2019.4418>

Ruczyński, Ł. (2025, 26 agosto). *ML frameworks compared: Scikit-learn, TensorFlow, PyTorch and more (updated)* [Blog]. Netguru. Recuperado de <https://www.netguru.com/blog/top-machine-learning-frameworks-compared> netguru.comnetguru.com

Ryzhokhin, A. (2024). Big data in logistics: Key benefits & real use cases. Ardas. <https://ardas-it.com/big-data-in-logistics-key-benefits-real-use-cases> ardas-it.com

Ryzhokhin, A. (2024, 14 de octubre). *Big data in logistics: Key benefits & real use cases.* Ardas. <https://ardas-it.com/big-data-in-logistics-key-benefits-real-use-cases>

Shapiro, C., & Varian, H. R. (1999). Information rules: A strategic guide to the network economy. *Harvard Business Review Press*.

*Tesis Carnivery – Entrega 6* (2023). Análisis interno de Carnivery: diagnóstico de datos y requerimientos de BI. (Documento inédito proporcionado por el equipo Carnivery)

*Tesis Carnivery – Entrega 7* (2023). Evaluación de plataformas Big Data en la nube para Carnivery. (Documento inédito proporcionado por el equipo Carnivery)

Tuzhilin, A. (2024). The power of big data in transforming business decision making. University of the People Leadership Blog. <https://www.uopeople.edu/blog/big-data-impacts-business-decisions/> uopeople.edu

Tuzhilin, A. (2024, 7 de octubre). *The power of big data in transforming business decision-making.* University of the People Leadership Blog. <https://www.uopeople.edu/blog/big-data-impacts-business-decisions/>

Zoho One. (2025). Información de uso parcial de módulos. Reporte interno de aplicativos.

