

Gestión de Datos

Construcción de un Entorno Analítico (DWH) y Cálculo del CLTV de Cliente

Dr. Jacinto Velasco Rebolledo
3º Curso Ingeniería Matemática



1.- Contexto de Negocio

El modelo de negocio en el sector de la automoción abarca desde la producción hasta la postventa, integrando diferentes estrategias de fabricación, comercialización y fidelización del cliente para garantizar rentabilidad y sostenibilidad a largo plazo.

En la fase de producción, las empresas fabricantes diseñan, ensamblan y optimizan la fabricación de vehículos a través de economías de escala y sistemas de producción eficientes. La fabricación puede realizarse de manera vertical, con control sobre la mayoría de los componentes, o a través de proveedores externos que suministran partes específicas. La optimización de costos, la automatización y la digitalización han permitido mejorar la eficiencia y reducir los tiempos de entrega. Además, la personalización del producto y la adaptación a normativas medioambientales son factores clave para mantenerse competitivo en el mercado.

Una vez producidos, los vehículos entran en el circuito de venta y distribución. Tradicionalmente, la comercialización se ha realizado a través de concesionarios, pero la digitalización ha impulsado nuevas formas de venta directa por parte de los fabricantes, así como plataformas online que permiten a los clientes configurar y comprar su vehículo sin necesidad de intermediarios. Existen diferentes modelos de comercialización, desde la venta directa hasta el leasing, el renting o incluso la suscripción, permitiendo a los consumidores acceder a un vehículo sin necesidad de adquirirlo en propiedad. La financiación y los seguros asociados son parte fundamental del proceso de venta, generando ingresos adicionales y facilitando el acceso a los vehículos.

Dentro del modelo de negocio, la diferenciación entre fabricantes se basa en la estrategia de mercado elegida. Algunas empresas se enfocan en la producción en masa, con vehículos accesibles y costos optimizados, mientras que otras priorizan la exclusividad y la innovación tecnológica para atraer a clientes dispuestos a pagar por mayor calidad y personalización. También existen compañías que han desarrollado modelos centrados en la movilidad como servicio, ofreciendo vehículos bajo demanda en lugar de venta tradicional.

Tras la compra, la postventa se convierte en un pilar clave del negocio, ya que permite generar ingresos recurrentes a través del mantenimiento, las reparaciones, la venta de repuestos y la actualización de software en los vehículos más modernos. Los servicios de posventa pueden ofrecerse a través de redes oficiales de la marca o de talleres independientes que proporcionan alternativas más económicas. Además, con la digitalización de los vehículos, se han desarrollado nuevas fuentes de ingresos a partir de servicios conectados, suscripciones a funcionalidades específicas y mejoras de software.

En conjunto, el modelo de negocio del sector automotriz no se basa únicamente en la venta de vehículos, sino en la creación de un ecosistema de servicios que acompañan al cliente durante todo el ciclo de vida del automóvil. La fidelización del cliente y la capacidad de generar ingresos más allá de la venta inicial son clave para la sostenibilidad y competitividad de las empresas en este mercado en constante evolución.

2.- Caso de Uso: Diseño de un entorno analítico de todo el ciclo de producto

2.1.- Contexto del proyecto

En la industria automotriz, la integración y análisis de datos provenientes de diversas áreas del negocio es fundamental para optimizar la toma de decisiones. Actualmente, la información se encuentra dispersa en distintos sistemas operacionales, como ERP para la gestión de ventas y logística, CRM para la interacción con clientes, y sistemas de postventa para el mantenimiento y seguimiento de los productos. La falta de un entorno analítico consolidado dificulta la obtención de insights en tiempo real y la creación de reportes estratégicos para mejorar la eficiencia y la experiencia del cliente.

El propósito de este proyecto es diseñar e implementar un entorno analítico centralizado que permita consolidar datos de distintas fuentes en un Data Warehouse (DWH). Este sistema proporcionará información estructurada y optimizada para el análisis del ciclo completo del producto, desde la producción y venta hasta la postventa y el servicio al cliente. El sistema debe integrar datos desde diferentes tablas de los sistemas ERP, CRM, logística y postventa, construyendo un modelo de datos robusto que permita realizar análisis de tendencias, medición de desempeño y segmentación de clientes, así como optimizar la eficiencia operativa.

El proyecto incluirá las siguientes fases:

- **Identificación de fuentes de datos:** Definición de los sistemas involucrados y las tablas relevantes en ERP, CRM, logística y postventa.
- **Diseño del modelo de datos:** Creación de un modelo entidad-relación (ER) para definir la estructura de almacenamiento en el Data Warehouse.
- **Desarrollo del modelo dimensional:** Construcción de tablas de hechos y dimensiones para facilitar el análisis de datos.
- **Arquitectura del Data Lake a DWH:** Definición del flujo de datos desde el almacenamiento en la nube hasta la base de datos local.

- **Implementación del proceso ETL:** Desarrollo de un pipeline de extracción, transformación y carga (ETL) para mover los datos desde Azure hasta un servidor local.
- **Validación y explotación de datos:** Creación de dashboards y reportes para la visualización y análisis de información.

2.2.- Requerimientos del Proyecto

1. El sistema debe consolidar información proveniente de múltiples fuentes en un único entorno analítico.
2. Debe diseñarse una arquitectura escalable que soporte grandes volúmenes de datos.
3. Se debe garantizar la calidad de los datos mediante procesos de limpieza y transformación.
4. La solución debe permitir la integración con herramientas de BI para la generación de reportes y dashboards.
5. El proceso ETL debe asegurar la actualización periódica de los datos desde la nube hasta el entorno local.

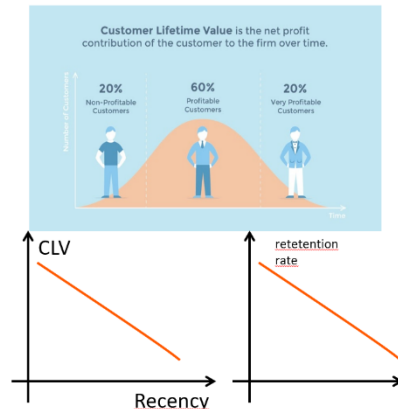
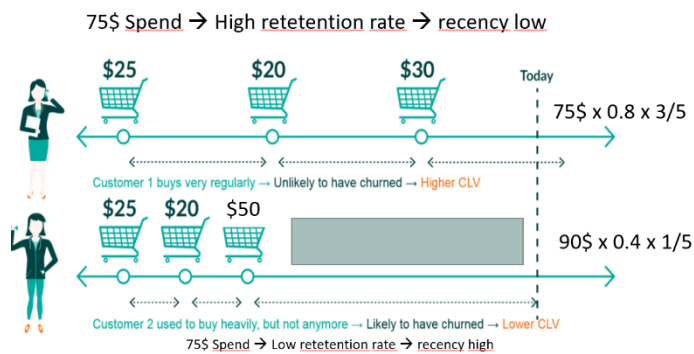
3.- Caso de Uso: Diseño de un Entorno Analítico para el Ciclo de Producto y Cálculo del Customer Lifetime Value (CLTV)

3.1.- Contexto y Necesidad

En la industria automotriz, contar con un análisis integral del ciclo de vida del producto es fundamental para optimizar las operaciones y mejorar la experiencia del cliente. Actualmente, los datos de ventas, postventa, logística y servicio al cliente se encuentran dispersos en distintos sistemas (ERP, CRM, sistemas de postventa y logística), lo que dificulta la obtención de insights estratégicos. Además, la empresa no cuenta con un modelo preciso para calcular el **Customer Lifetime Value (CLTV)**, lo que impide definir estrategias efectivas de retención y personalización de ofertas. El **CLTV** es un indicador clave que permite estimar el valor total que un cliente genera para la empresa a lo largo de su relación comercial. Su cálculo requiere consolidar información de ventas, frecuencia de compra, márgenes de beneficio y comportamiento en servicios de postventa.

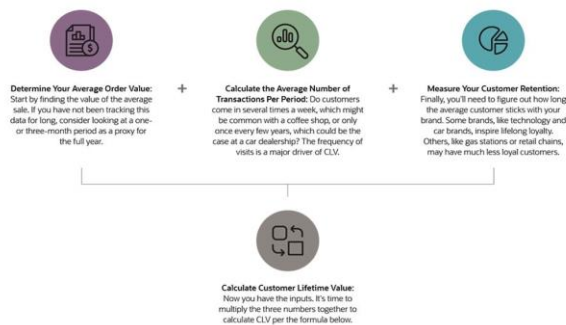
Customer lifetime value is the total revenue or value of a customer relationship over a lifetime.

Calculating CLV is important because it allows business owners to see how much profit they're making from each customer during their relationship, which provides valuable data for their marketing and sales efforts. A high CLV means that your customers leave a lot of money at your store, likely by shopping regularly or purchasing higher-priced items.



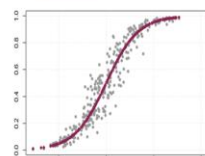
Customer lifetime value is the total revenue or value of a customer relationship over a lifetime.

How to Measure Customer Lifetime Value



$$CLV = \sum_{k=0}^n \frac{Margin_i \cdot R_t \cdot f}{(1+i)^n} - Cost_{AC}$$

where
 m_t = margin by a consumer at time t (Net sales – production cost),
 $Cost_{AC}$ = acquisition cost,
 i = discount rate or cost of capital for the firm,
 r_t = probability of customer repeat buying or being "alive" at time t ,
 N = time horizon for estimating CLV,
 f = frequency (one timers + existing)



Retention rate refers to the percentage of customers who continue paying for a product over a given timeframe

3.2.- Objetivo del Caso de Uso

El propósito de este proyecto es diseñar un **entorno analítico centralizado** que consolide información de distintas fuentes en un **Data Warehouse (DWH)**, proporcionando información estructurada para el análisis del ciclo de producto y la estimación del **Customer Lifetime Value (CLTV)**. Con este modelo, la empresa podrá identificar patrones de consumo, optimizar la segmentación de clientes y desarrollar estrategias de fidelización basadas en datos.

El proyecto incluirá las siguientes fases:

1. **Identificación de fuentes de datos:** Definición de los sistemas involucrados (ERP, CRM, logística, postventa) y selección de las tablas relevantes para el análisis del ciclo de vida del producto y el CLTV.

2. Implementación del cálculo del CLTV:

- **Cálculo de ingresos por cliente** a partir de ventas históricas.
- **Frecuencia y tasa de recompra**, con base en interacciones en postventa y CRM.
- **Cálculo del margen de beneficio** descontando costos de adquisición y servicio.
- **Estimación del tiempo de relación con el cliente** con base en datos históricos.

Dado que los clientes pueden generar ingresos a lo largo del tiempo mediante compras adicionales, mantenimiento y servicios financieros, el mejor enfoque para el cálculo del CLTV en automoción es un modelo multi-ingreso y basado en cohortes de clientes.

$$CLTV = \sum_{t=0}^N \frac{(Ingresos_t \times Margen_t \times R_t)}{(1+i)^t} - Costo_{AC}$$

Para que el CLTV refleje correctamente la realidad del negocio, se deben considerar las siguientes fuentes de ingresos y costes:

- **Venta de Vehículo Principal:** Margen sobre el precio de venta (después de incentivos y descuentos). Financiación y seguros si se venden como parte del contrato.
- **Postventa y Mantenimiento:** Servicios de mantenimiento programado. Reparaciones y repuestos (aftermarket). Venta de accesorios.
- **Servicios Financieros y Suscripciones:** Ingresos por financiamiento del vehículo (préstamos o leasing). Seguros asociados a la compra del vehículo.
- Suscripciones de software y conectividad (ejemplo: actualizaciones OTA, telemetría, navegación premium).
- **Segunda Compra o Recompra (Retención del Cliente):** Probabilidad de que el cliente vuelva a comprar otro vehículo de la marca. Programas de fidelización o leasing con recompra.

La tasa de descuento, representada como i en la fórmula del Customer Lifetime Value (CLTV), refleja el costo de oportunidad del capital y el valor temporal del dinero. En la industria automotriz, esta tasa puede variar según diversos factores, incluyendo las condiciones económicas y las políticas monetarias vigentes. Aunque no existe una tasa de descuento estándar para la industria automotriz, las tasas de interés para préstamos de automóviles nuevos han promediado alrededor del **7% en fechas recientes**.

En la **postventa automotriz**, la retención de clientes se mide en función de **cuántos clientes regresan a los talleres oficiales** después de su primera visita. Se puede calcular la **tasa de retención** en función del número de visitas al taller o la probabilidad de que un cliente vuelva en el futuro utilizando **Regresión Logística**. Hay que considerar que una persona ha abandonado si no ha vuelto al taller en 400 días desde la última visita. Para construir la tabla usar máximo tres variables.

$$p = \frac{1}{1 + e^{-P(x)}} \rightarrow P(x) = \beta_0 + \sum_{i=1}^N \beta_i \cdot x_i$$

4.- Entregables

- Diagrama del modelo de Entidad Relación
- Diagrama del Modelo Dimensional. Esquema de las tablas finales de hechos y dimensiones.
- ETL que mueva los datos del Datalake al entorno analítico. Solo el diseño de la solución, si se aplica se valorará positivamente.
- Cálculo y análisis del CLTV. El análisis se puede hacer con Power BI o cualquier herramienta de análisis de datos.
- Modelo de retención aplicable al CLTV
- PCA de la base de datos y un modelo de clústering según el CLTV
- Documento no más de 10 caras (5 hojas) con el detalle técnico del proyecto.