

# ESTRATEGIAS DE NEGOCIOS CON PROCESOS DE MARKOV: ANÁLISIS DE RIESGOS DE DESACTIVACIÓN, REACTIVACIÓN Y ESTRATEGIAS DE VENTAS

B. Itzelt Gómez Catzín<sup>1</sup>, Rebeca Koch Torres<sup>1</sup>, Gabriela Lujan<sup>1</sup>, María F. Gamboa Martinez<sup>1</sup> and  
Andrea Ruía Álvarez<sup>1</sup>

<sup>1</sup>Tecnológico de Monterrey, Campus Guadalajara

---

## Abstract

*This study explores the application of Markov chains to analyze the state transitions of products within a sales portfolio. By examining the behavior of a sample of products, we developed transition matrices to represent the probabilities of moving between active and inactive states. Our analysis reveals the presence of both ergodic and non-ergodic chains among the products, with implications for predicting long-term sales behavior. We identified approximately 538 products with ergodic chains, indicating stable sales patterns, and 127 products with non-ergodic chains, reflecting more unpredictable sales dynamics. The findings are used to guide strategic decisions on product activation and deactivation, with a focus on optimizing inventory and maximizing profitability. Additionally, we calculated the Customer Lifetime Value (CLV) to evaluate customer worth, which supports targeted retention strategies and enhances overall business performance.*

**Keywords:** Markov Chains, State Transitions, Product Lifecycle, Ergodic Chains, Sales Prediction, Customer Lifetime Value, Inventory Management, Product Deactivation

---

## 1 INTRODUCCIÓN

PiSA Farmacéutica es una de las empresas socialmente responsables que más contribuye al desarrollo de México. Como empresa mexicana, se ha dedicado a ofrecer productos y servicios integrales para los sectores de salud pública y privada en México, Estados Unidos, Latinoamérica y el Caribe. PiSA se destaca por su gran capacidad de producción y su vasta experiencia en el sector, operando 14 plantas estratégicamente ubicadas, con más de 20,000 colaboradores y un portafolio que incluye más de 1,500 marcas de medicamentos y 17 líneas de especialidad.

Nuestro equipo tuvo la oportunidad de visitar una de sus plantas de producción, donde pudimos observar de primera mano la amplia gama de productos y servicios que PiSA no solo ofrece en México, sino que también exporta a varios países. El trabajo logístico, el análisis detallado y el esfuerzo detrás del crecimiento y expansión internacional de la empresa son impresionantes.

No obstante, siempre existen áreas de mejora. A través de la tecnología y decisiones basadas en datos, es posible lograr resultados aún más impactantes. El reto de este proyecto es fortalecer la

relación con los clientes actuales de PiSA, asegurando que continúen consumiendo sus productos de manera sostenida.

En este proyecto, estamos realizando un análisis profundo de los datos proporcionados por Laboratorios PiSA para estudiar el comportamiento de sus clientes en cuanto a la compra y uso de productos de su catálogo durante un período determinado. Nos enfocamos en tres aspectos clave: adquisición, retención y pérdida de clientes. Al comprender estos patrones de comportamiento utilizando cadenas de Markov, buscamos optimizar las estrategias comerciales de la empresa, con el objetivo de maximizar sus ventas por producto, la retención de clientes y fomentar su lealtad a largo plazo.

## 2 METODOLOGÍA

Como mencionamos anteriormente, este proyecto se centrará en comprender los patrones de comportamiento de los productos utilizando cadenas de Markov, con el objetivo principal de maximizar su retención. Sin embargo, antes de profundizar en el análisis, es fundamental revisar algunos conceptos teóricos clave que son esenciales para llevar a cabo este estudio de manera efectiva.

### 2.1 Cadena de Markov a tiempo discreto

Una cadena de Markov a tiempo discreto es un modelo matemático ampliamente utilizado para modelar sistemas que evolucionan en el tiempo, con la particularidad de que es un tipo especial de proceso estocástico en el cual la probabilidad de que ocurra un evento depende únicamente del estado inmediato anterior, lo que se conoce como propiedad de memoria limitada. (Dr. Garrido J., 2024)

Un proceso a tiempo discreto  $\{X_n\}_{n=0}^{\infty}$  es una cadena de Markov si y sólo si para toda  $n \in \mathbb{N}$  y  $\{x_0, \dots, x_{n+1}\} \in \mathcal{E}$  se cumple la propiedad de Markov, esto es:

$$P(X_{n+1} = x_{n+1} \mid X_n = x_n, \dots, X_0 = x_0) = P(X_{n+1} = x_{n+1} \mid X_n = x_n)$$

### 2.2 Probabilidad de transición

En el contexto de las cadenas de Markov, la probabilidad de transición se define como la probabilidad de que un sistema pase de un estado a otro en un solo paso de tiempo.

Sea  $\{X_n\}_{n=0}^{\infty}$  un proceso estocástico a tiempo discreto con espacio de estados  $\mathcal{E}$  discreto. Para cualquier  $i, j \in \mathcal{E}$ , la probabilidad de pasar de un estado  $i$  a un estado  $j$  en un paso se conoce como probabilidad de transición y está dada por:

$$P(X_{n+1} = j \mid X_n = i)$$

### 2.3 Matriz de transición

La matriz de transición, como su nombre lo indica, es una matriz que contiene todas las probabilidades de transición de un paso entre estados, donde cada entrada indica la probabilidad de moverse

de un estado específico a otro.

$$P = \begin{pmatrix} p_{11} & p_{12} & \cdots \\ p_{21} & p_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}$$

$$p_{ij} = P(X_1 = j \mid X_0 = i)$$

Las matrices de transición tienen propiedades importantes de probabilidad, como las siguientes: probabilidades no negativas, es decir, todos los elementos de la matriz deben ser valores entre 0 y 1, y la suma de las probabilidades en cada fila es igual a 1 (Dr. Garrido J., 2024).

#### 2.4 Cadena de dos estados

Las cadenas de Markov pueden operar con espacios de estados infinitos, pero en este proyecto en particular trabajamos con una cadena de Markov de dos estados. Por ello, es fundamental explicar en detalle su estructura y cómo la utilizamos en nuestro análisis.

Sea  $\{X_n\}_{n=0}^{\infty}$  un proceso a tiempo discreto con espacio de estados  $\mathcal{E} = \{0, 1\}$ .

Su matriz de transición está dada por:

$$P = \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix}$$

donde  $p, q \in (0, 1)$ .

#### 2.5 Matriz de transición en $n$ -pasos

Una matriz de transición a  $n$ -pasos contiene las probabilidades de que el sistema pase de un estado a otro en exactamente  $n$  pasos. A diferencia de la matriz de transición de un solo paso, esta matriz permite analizar las probabilidades de llegar a un estado específico después de un número determinado de iteraciones.

Estas matrices son extremadamente útiles y presentan ventajas sobre la matriz de transición de un solo paso. Permiten predecir comportamientos a largo plazo y, bajo ciertas condiciones, ayudan a determinar si el sistema alcanzará un estado estable, en el que las probabilidades de estar en cada estado ya no cambian con el tiempo. Esto facilita el análisis de probabilidades y proporciona una visión más completa del comportamiento del sistema a lo largo del tiempo.

#### 2.6 Cadenas ergódicas y convergencia

Como mencionamos anteriormente, bajo ciertas condiciones, una matriz de transición a  $n$ -pasos puede converger a un estado estable o estacionario, donde las probabilidades de estar en cada estado ya no cambian con el tiempo. Este tipo de comportamiento ocurre en cadenas de Markov que se

conocen como cadenas ergódicas. Para que una cadena de Markov sea ergódica y alcance un estado estacionario, debe cumplir con las siguientes características:

- **Irreducibilidad:** Una cadena es irreducible si todos sus estados se comunican entre sí, es decir, es posible alcanzar cualquier estado desde cualquier otro. Esto implica que no hay subconjuntos aislados dentro del espacio de estados.
- **Aperiodicidad:** Una cadena es aperiódica si el número de pasos necesarios para regresar a un estado no es múltiplo de un número mayor que uno. Esto significa que la cadena no quedará atrapada en ciclos regulares, garantizando que el proceso no siga un patrón cíclico rígido.
- **Recurrencia Positiva:** Un estado es recurrente positivo si el tiempo esperado para regresar a él es finito. Para asegurar la convergencia, es necesario que todos los estados de la cadena sean recurrentes positivos, lo que previene que la cadena se desvíe indefinidamente sin regresar a ciertos estados.

Cuando una cadena de Markov cumple con estas tres propiedades, se considera ergódica. Las cadenas ergódicas garantizan que, a medida que se incrementa el número de pasos, la probabilidad de estar en cualquier estado particular converge a un valor estacionario, sin importar el estado inicial. Esta distribución estacionaria refleja el comportamiento a largo plazo de la cadena (Dr. Garrido J., 2024).

### 3 RESULTADOS OBTENIDOS

#### 3.1 *Análisis Exploratorio de la Base de Datos*

##### 3.1.1 *Comprensión de la base de datos*

La base de datos utilizada en este análisis fue proporcionada por Laboratorios PISA y contiene cuatro columnas clave para nuestro estudio:

- **fecha:** la fecha que se realizó la venta.
- **id\_material:** código del producto que se vendió.
- **id\_cliente:** código del cliente.
- **ventas:** la venta en valor monetario.

Es fundamental comprender la cantidad de datos disponibles y las variables presentes en esta base de datos para llevar a cabo un análisis adecuado. En total, contamos con 830,517 registros por columna. Esta gran cantidad de datos es beneficiosa, ya que permite realizar un análisis exhaustivo.

Es importante señalar que las columnas *ID Material* y *ID Cliente* solo contienen identificadores numéricos que no proporcionan información adicional sobre el cliente o el producto. Esto limita nuestra capacidad para inferir detalles específicos sobre ellos. En cambio, las columnas *Fecha* y *Ventas* nos ofrecen valores concretos que permiten realizar análisis más profundos sobre las ventas y su

distribución en el tiempo.

En cuanto al tipo de datos, las columnas *ID Material* y *ID Cliente* contienen datos de tipo entero (int), la columna *Ventas* está compuesta por valores de tipo doble precisión (dbl) y la columna *Fecha* se encuentra en formato de texto (chr). Conocer el tipo de datos con el que estamos trabajando y asegurarnos de que cada uno es coherente con su respectiva columna es beneficioso para nuestro proyecto, ya que garantiza que estaremos utilizando los formatos adecuados para nuestro análisis.

Durante nuestras visualizaciones encontramos que el periodo de tiempo de nuestros datos estaban divididos en dos, de 2011 a 2014, con un salto en el tiempo de 2021 a 2024.

### 3.1.2 *Limpieza de Datos*

Antes de comenzar a trabajar con nuestros datos, es fundamental asegurarnos de que no haya valores faltantes en la base de datos. Durante la revisión inicial de la base de datos, confirmamos que cada una de las cuatro columnas contiene exactamente 830,517 registros, lo que indica que no hay datos faltantes.

Por otro lado, al examinar detalladamente los datos, observamos una irregularidad: aunque tenemos registros desde 2011 hasta 2024, parece haberse hecho una combinación de dos bases de datos de distintos periodos. Esto es evidente por un salto temporal, ya que no hay registros entre 2014 y 2020, reanudándose los datos en 2021. Trabajar con esta discontinuidad de siete años presenta varios problemas en el análisis, ya que ese periodo sin actividad podría ser interpretado como estados inactivos, lo cual distorsionaría el comportamiento real de los productos.

Dado que se trata de un intervalo considerable, que además incluye una pandemia, consideramos que los datos anteriores a 2021 no proporcionarían información útil para nuestros objetivos actuales. Por ello, decidimos imputar esos valores para continuar con un análisis más coherente y preciso.

Además, para llevar a cabo un análisis más enriquecedor para la empresa, consideramos que, dado que disponemos de los datos de ventas diarias de cada producto, sería más útil agrupar las ventas por mes mediante una suma. Esto nos permitirá obtener una visión más clara del comportamiento de las ventas, facilitando la identificación de patrones y mejorando la precisión de los resultados.

### 3.1.3 *Estadísticas Descriptivas*

Las estadísticas descriptivas son fundamentales para comprender la distribución de nuestros datos. En el análisis exploratorio, generamos estas estadísticas para cada columna para obtener una visión general de su comportamiento. Como mencionamos anteriormente, las columnas *ID Cliente* y *ID Material* solo contienen identificadores que no aportan información cuantificable significativa. Por lo tanto, calcular estadísticas descriptivas para estas columnas no es relevante.

En la columna *Ventas*, las estadísticas descriptivas mostraron un valor mínimo de -1,149,758.3. Inicialmente, asumimos que este valor, junto con otras ventas negativas, era incoherente, ya que los valores monetarios de ventas no deberían ser negativos. Sin embargo, tras una conversación con Laboratorios PISA, comprendimos que si aparecen dos o más valores negativos consecutivos, estos deben

interpretarse como una desactivación de un cliente o producto. Por otro lado, si se encuentra solo un valor negativo, debe ignorarse, continuando con la consideración del periodo como activo. Estas estadísticas descriptivas fueron clave para entender cómo abordar estos datos y tratarlos adecuadamente.

Por último, en cuanto a las ventas, observamos que el valor máximo es de 2,022,000, mientras que el promedio es de 6,131.4, lo que sugiere que la mayoría de las transacciones se concentran en torno a este valor. En lo que respecta a la columna *Fecha*, como mencionamos anteriormente, solo trabajaremos con los datos comprendidos entre 2021 y 2024, lo que nos ofrece un intervalo temporal lo suficientemente amplio para realizar un análisis exhaustivo y relevante.

### 3.1.4 Visualización de Datos y Hallazgos

Primero, es fundamental identificar los productos con mayores ventas. Para ello, se ha analizado cuáles son los diez productos con mayor y menor volumen de ventas totales en el periodo comprendido entre el 1 de enero de 2021 y el 7 de agosto de 2024. Este análisis es crucial para comprender el comportamiento de las ventas por material, ya que esta información es de gran relevancia para el desarrollo del proyecto. Conocer qué productos generan mayores ingresos permite identificar patrones de consumo, optimizar la producción y mejorar las estrategias de marketing.

Productos	Ventas Totales
591	97,903,712
601	84,710,470
893	84,654,423
590	78,186,309
772	77,445,107
5862	45,806,524
600	44,405,787
817	43,288,175
724	28,079,135
670	27,697,462

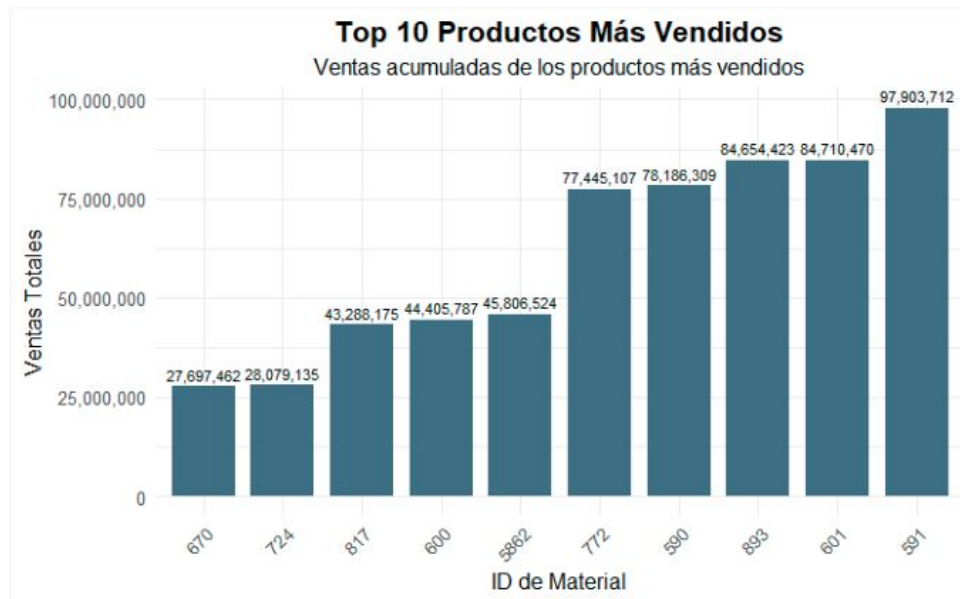
**Table 1.** Top 10 productos con mayores ventas

Productos	Ventas Totales
6929	-310.60
3488	-115.60
976	0.00
2934	0.00
6056	0.00
7207	0.00
2936	52.80
5714	62.80
3196	65.05
861	125.82

**Table 2.** Top 10 productos con menores ventas

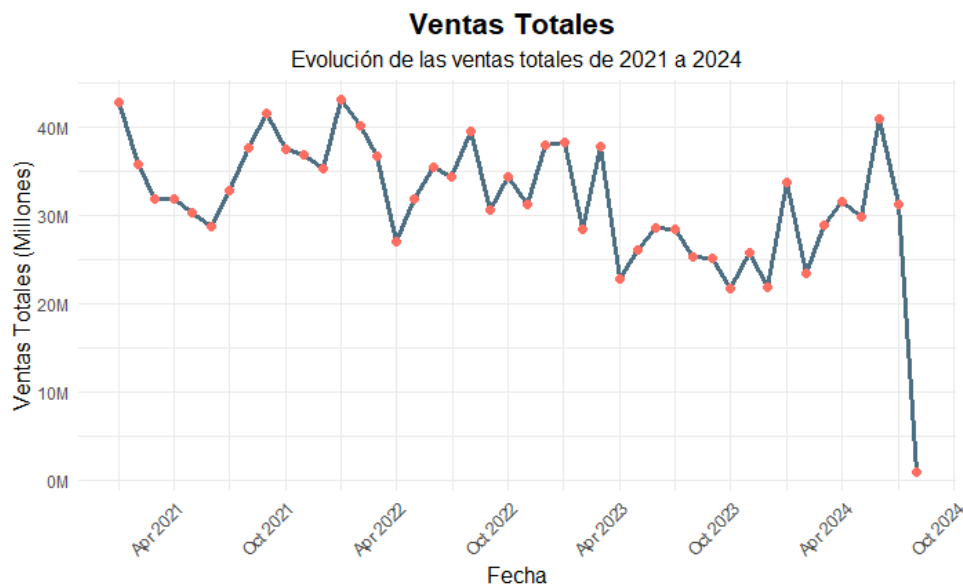
Es crucial analizar el comportamiento de las ventas totales por material, ya que esta información es sumamente relevante para el desarrollo del proyecto. Entender cuáles materiales o productos generan mayores ingresos permite identificar patrones de consumo, optimizar la producción y mejorar las estrategias de marketing.

En la Figura 1, se presentan los 10 productos con mayores ventas, lo que brinda una visión clara de los elementos más exitosos en el mercado. Este análisis no solo destaca los productos líderes, sino que también proporciona una base sólida para tomar decisiones informadas sobre dónde enfocar los recursos y esfuerzos en el futuro.



**Figure 1.** Top 10 productos con mayores ventas

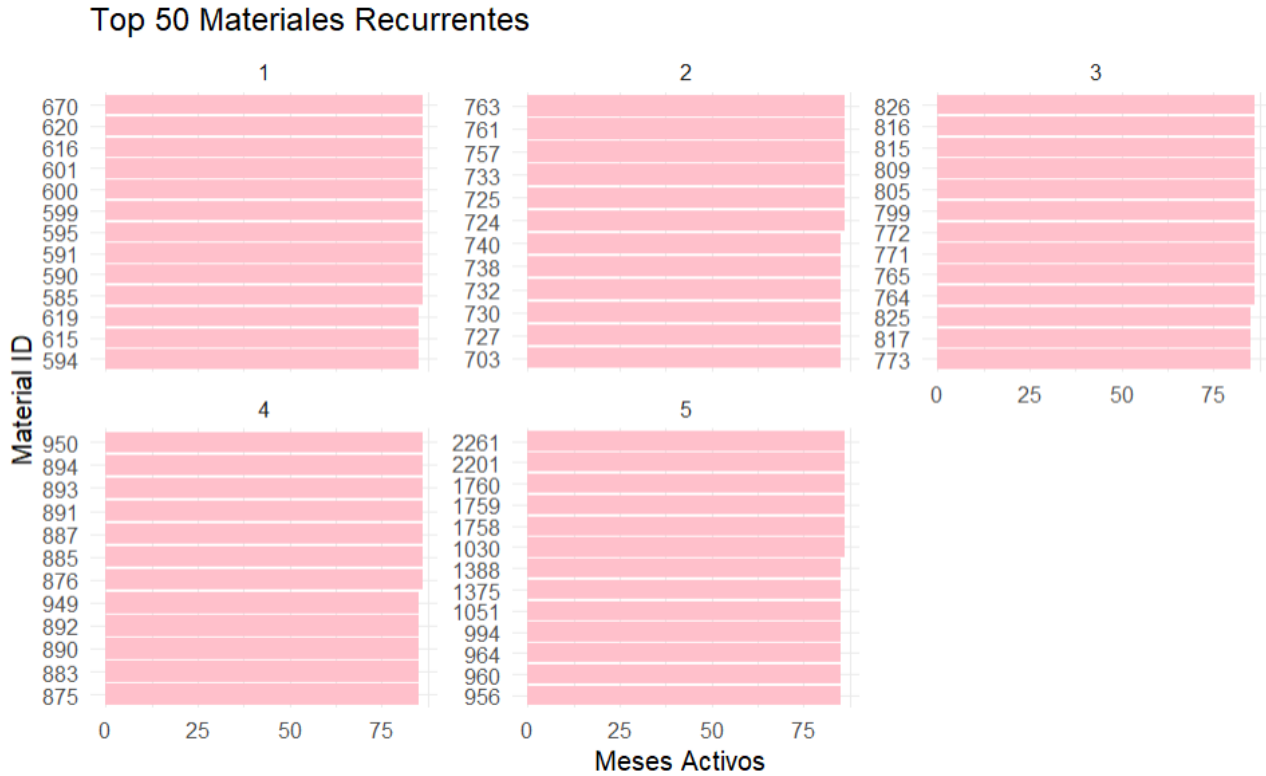
En la Figura 2 podemos ver que entre 2021 y 2023, las ventas anuales mostraron una tendencia significativa. En 2021 y 2022, las ventas se mantuvieron en niveles elevados, superando los 40 millones, lo que indica un periodo de estabilidad y éxito comercial. Sin embargo, en 2023 se observó una disminución notable en las ventas, que cayeron por debajo de los 30 millones. Este descenso podría estar relacionado con factores como cambios en la demanda del mercado, aumento de la competencia o posibles problemas internos. Es esencial investigar las causas de esta caída para ajustar las estrategias y recuperar el crecimiento en los años siguientes.



**Figure 2.** Ventas a lo largo de los años

Igualmente, este gráfico, que muestra la evolución de las ventas totales de 2021 a 2024, refleja fluctuaciones importantes en las ventas a lo largo del tiempo. Se observan picos regulares, lo que sugiere la existencia de un patrón estacional, con aumentos significativos a principios de año y caídas

en algunos meses, como septiembre. Durante los primeros dos años, las ventas alcanzaron picos de más de 40 millones, seguidos por descensos graduales, lo que sugiere la influencia de eventos específicos, como campañas de marketing o promociones estacionales. A lo largo de 2024, las ventas parecen haberse estabilizado, aunque siguen presentando fluctuaciones, manteniéndose entre los 30 y 40 millones de unidades.



**Figure 3.** Principales materiales recurrentes

Se hizo un análisis de todos los materiales y contando los meses durante los cuales cada material fue vendido, se identificaron los 50 principales materiales que tuvieron mayor actividad a lo largo de los años.

### 3.2 Planteamiento: Identificación de productos con riesgo de desactivación y probabilidad de reactivación utilizando Cadenas de Markov

#### 3.2.1 Definición de Estados

Primero, considerando que se estará analizando el comportamiento de los productos conforme se vendan o no cada mes, se tendrá una Cadena de Markov de dos estados, en donde el espacio de estados está dado por

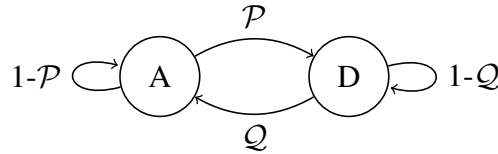
$$\mathcal{E} = \{A, D\} \quad (1)$$

siendo A el estado activo y D desactivado. Con activo nos referimos que el producto haya tenido ventas en el mes actual y desactivo en donde no tenga ventas.



### 3.2.2 Diagrama de Transición

Después, con base en los estados anteriores, definimos nuestro diagrama de transición



en donde

$\mathcal{P}$  es la probabilidad de que un producto de estar activo pase a estar desactivo al siguiente mes.  
 $1 - \mathcal{P}$  es la probabilidad de que un producto se mantenga activo de un mes a otro, es decir, siga teniendo ventas.

$\mathcal{Q}$  es la probabilidad de que un producto que está desactivo, el siguiente mes, pase a estar activo.  
 $1 - \mathcal{Q}$  es la probabilidad de que un artículo se mantenga desactivo, es decir, que no se venda

### 3.2.3 Matriz de Transición

Con base en lo anterior, se construirá la matriz de transición que usaremos para crear la de cada producto dependiendo sus probabilidades. De esta forma podremos crear una matriz de transición a  $n$ -pasos con la cual se podrá calcular la distribución de  $X_n$ , siendo esta el estado en el que se encuentre el producto, sea activo o desactivo, al tiempo  $n$ .

$$P = \begin{pmatrix} 1 - \mathcal{P} & \mathcal{P} \\ \mathcal{Q} & 1 - \mathcal{Q} \end{pmatrix}$$

### 3.2.4 Tabla de transiciones entre estados

Estos valores los obtenemos al analizar el comportamiento de cada producto entre estados, primero se realizó un análisis de cierta muestra aleatoria de cinco productos y se obtuvo la Tabla 3

Producto	$A \rightarrow A$	$D \rightarrow A$	$A \rightarrow D$	$D \rightarrow D$
603	10	8	8	17
616	43	0	0	0
629	14	1	2	26
641	30	5	6	2
682	6	6	7	24

**Table 3.** Tabla de estados por los productos elegidos aleatoriamente

en donde

$A \rightarrow A$  es el número de veces que el producto se mantuvo activo, es decir, vendiendo mes con mes  
 $D \rightarrow A$  es el número de veces que el producto de estar desactivo, pasó a estar activo, es decir, de no vender, el siguiente mes sí vendió  
 $A \rightarrow D$  es el número de veces que el producto de estar activo pasó a estar desactivo, es decir, de vender, el siguiente mes no vendió  
 $D \rightarrow D$  es el número de veces que un artículo se mantuvo desactivo, es decir, que no se vendió

Podemos observar el número de veces que transiciona de un estado a otro, por mes, para cada producto de nuestra muestra, esta información será de gran valor para los futuros análisis y encontrar el estado en el que las cadenas de Markov converjan.

### 3.3 Estimación de Probabilidades de Transición

En este caso, para calcular las probabilidades de transición de cada producto, lo que hicimos fue que con base en lo realizado en la Sección 3.2.4 creamos la matriz de cada producto como se observa en la Sección 3.2.3. Como debíamos asegurarnos que nuestra matriz fuera estocástica, la suma de las filas debía ser uno, por lo cual debíamos normalizar nuestros valores, entonces realizamos lo siguiente:

$$Suma_A = A \rightarrow A + A \rightarrow D$$

$$Suma_D = D \rightarrow A + D \rightarrow D$$

Al dividir los valores de cada estado por la suma correspondiente, obtuvimos los siguientes resultados para cada uno de nuestros productos

Producto	$A \rightarrow A$	$D \rightarrow A$	$A \rightarrow D$	$D \rightarrow D$
584	0.93333333	0.15384615	0.06666667	0.8461538
585	1.00000000	0.00000000	0.00000000	0.00000000
586	0.77142857	0.87500000	0.22857143	0.12500000
587	0.40000000	0.47826087	0.60000000	0.5217391

**Table 4.** Transiciones de estados normalizados para una muestra de cuatro productos

Con estos valores, lograremos crear nuestra matriz de transición para cada producto tomando la forma que se definió en 3.2.3, por ejemplo, para el producto con *ID 591*, la matriz de transición está dada por

$$P = \begin{pmatrix} 0.8064516 & 0.1935484 \\ 0.4166667 & 0.5833333 \end{pmatrix}$$

### 3.4 Cadenas Ergódicas

Como mencionamos en la Sección 2.6, para que una cadena sea ergódica debe de cumplir tres condiciones: irreducibilidad, recurrencia positiva, y aperiodicidad. En este caso, podemos determinar aquellas cadenas que no cumplan lo anterior al revisar sus matrices de transición, en específico, los

valores que se encuentren fuera de la diagonal principal, ya que si alguno de estos valores es cero, significaría que la conexión entre ambos estados no existe, por lo cual no cumpliría con las condiciones requeridas para que sea una cadena ergódica. Encontramos que al rededor de 538 productos cumplen las condiciones, es decir, son ergódicas, por otro lado, 127 productos no son ergódicos. Algunos de los productos que categorizamos son:

#### **Productos ergódicos**

- ELECTROLIT NARAN-MANDARI C/625ML Soor [ID 662]
- COMBINACION PI C/500ML 1FCO FLX [ID 584]
- JERINGA 20ML PLAS P/PREFURSOR C/AGUJA [ID 1000]

#### **Productos no ergódicos**

- AGUA P/IRRIGA ESTERIL C/1000ML 1FCO PLAS
- SOLUCION CS C/500ML 1FCO FLX SI
- MICROCLAVE CLEAR C/EXT MICRO 30CM 3VIA

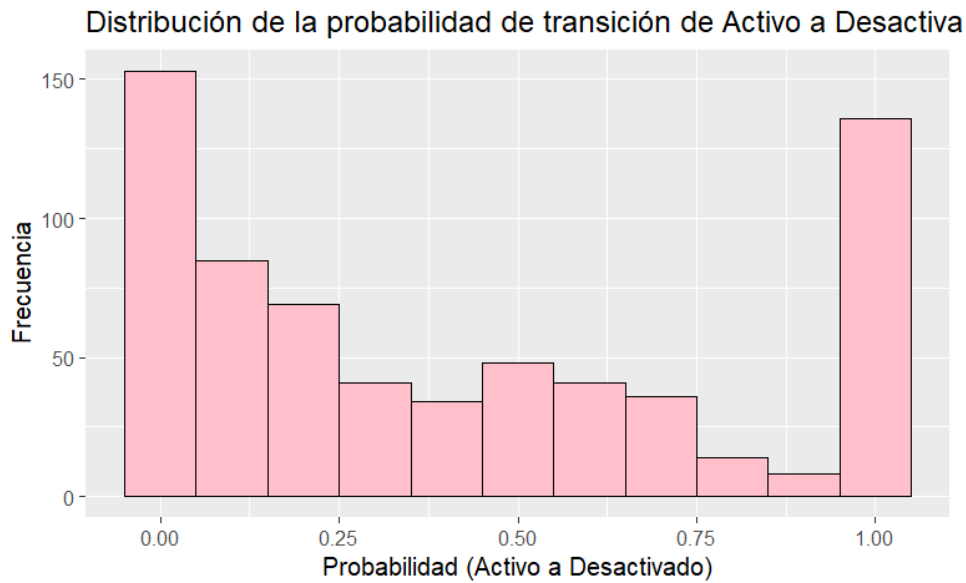
### **3.5 Supuestos y Desafíos en la Creación del Modelo**

Para realizar nuestro análisis, asumimos que el comportamiento del sistema sigue un proceso estacionario de primer orden, lo que implica que el sistema tiene una memoria limitada: el estado futuro depende únicamente del estado actual y no de los anteriores. Bajo esta premisa, consideramos que el tiempo tiende al infinito, lo que nos permite estudiar los datos bajo la dinámica en la que el sistema se estabiliza y eventualmente alcanza un equilibrio. Esta suposición facilita el análisis, al enfocarnos en un escenario donde las variaciones en el comportamiento desaparecen y se llega a un estado estable.

El hecho de que el sistema tenga memoria limitada simplifica nuestro modelo, ya que reduce la complejidad del análisis al enfocarnos solo en las transiciones de un estado al siguiente (Dr. Garrido J., 2024).

### **3.6 Impactos de desactivación y reactivación de los productos**

La gestión efectiva del portafolio de productos es crucial para cualquier empresa que aspire a mantener su competitividad y rentabilidad en un mercado dinámico. Al realizar un análisis exhaustivo de los materiales y las ventas, se pueden identificar productos que, debido a su bajo rendimiento en el mercado, podrían no justificar la inversión continua en su producción y comercialización. Este análisis nos permite tomar decisiones informadas sobre qué productos seguir ofreciendo y cuáles considerar para su desactivación. Sin embargo, es importante entender que la decisión de discontinuar un producto tiene implicaciones significativas, tanto positivas como negativas.



**Figure 4.** Distribución de la probabilidad de transición de Activo a Desactivado

En la Figura 4 se muestra la distribución de la probabilidad de transición de *Activo a Desactivado* de los productos. En el eje X se presenta la probabilidad de que un producto pase de estar activo a desactivo, con valores que van de 0 a 1. Un valor de 0 indica que no hay probabilidad de desactivación, mientras que un valor de 1 significa que la desactivación es segura. El eje Y indica la frecuencia, es decir, cuántos productos tienen la probabilidad de desactivación correspondiente al valor en el eje X.

Por ejemplo, la barra más alta del histograma se encuentra cerca del 0 en el eje X, alcanzando una frecuencia de poco más de 150 en el eje Y. Esto sugiere que más de 150 productos tienen una probabilidad muy baja de desactivarse, lo que indica que son productos estables y con alta permanencia en estado activo.

En el extremo derecho, con una probabilidad de 1, se observan los productos que casi con certeza se desactivarán. Es importante destacar que el número de productos en este estado es considerable, lo que implica que muchos productos tienen una probabilidad segura de no vender el mes siguiente.

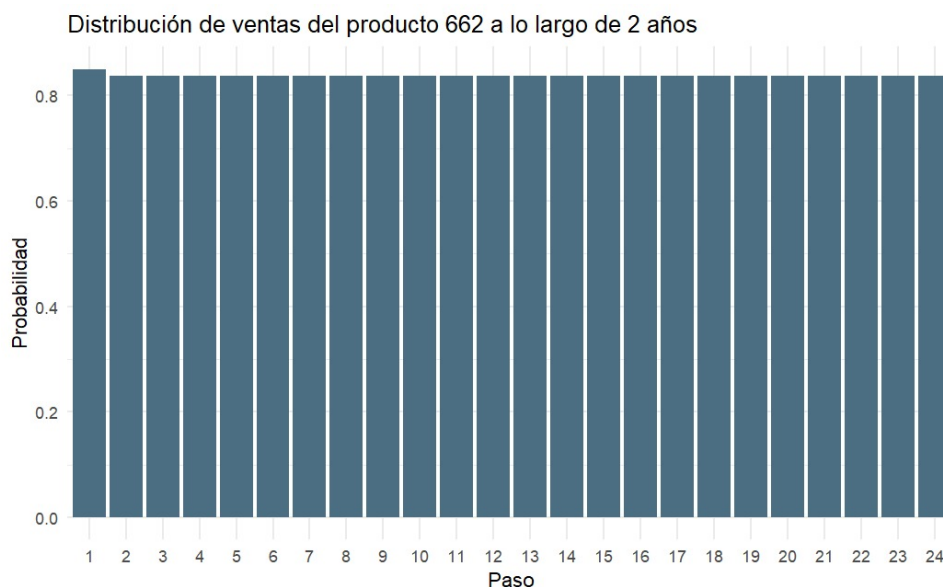
### 3.7 Proyección de Ventas

Para obtener la proyección de ventas a largo plazo de los productos en la base de datos proporcionada, fue necesario primero identificar aquellos productos cuyo comportamiento de ventas se estabilizará a largo plazo y aquellos que no. En base a los resultados de la sección 3.4, identificamos alrededor de 538 productos cuyas cadenas son ergódicas y, por lo tanto, sus ventas tendrán un comportamiento estable. Esto significa que, con el tiempo, las probabilidades de ventas de estos productos tienden a converger en un estado de equilibrio, independientemente del tiempo o del punto de partida. Es decir, se puede esperar que a largo plazo, las ventas de estos productos se distribuyan de manera constante entre diferentes estados, los cuales en este caso, serían los estados activo o desactivo en el contexto de la demanda del producto.

Por otro lado, identificamos otros 127 productos cuyas respectivas cadenas de Markov no son ergódicas, por lo que las probabilidades de ventas no se estabilizan a lo largo del tiempo. Esto indica que el comportamiento de ventas de estos productos depende significativamente de las condiciones iniciales o de eventos específicos. Este tipo de comportamiento puede estar influenciado por factores externos, estacionalidad o cambios en las preferencias de los consumidores, resultando en ventas más impredecibles.

Una vez clasificadas las cadenas de los productos en ergódicas y no ergódicas, calculamos las matrices de transición a  $n$ -pasos de cada producto. Si bien recordamos en la sección 2.5, estas matrices nos permiten analizar las probabilidades de llegar a un estado específico, en este caso, el estado activo, después de un número determinado de iteraciones, y son útiles para obtener la distribución de probabilidades, lo que resulta en las proyecciones de venta a largo plazo para cada uno de los productos.

Esta proyección se puede ilustrar en la siguiente gráfica, la cual es la distribución de ventas para uno de los productos cuyo comportamiento de ventas se estabiliza. Como podemos observar, el eje  $X$  representa el tiempo en meses, y el eje  $Y$  representa la probabilidad de que el producto se encuentre en estado activo, es decir, en venta. Como se observa, las probabilidades se mantienen constantes a través de los meses, lo que refleja el comportamiento de ventas de un producto cuya cadena es ergódica.



**Figure 5.** Distribución de ventas del producto 662 en 2 años

### 3.8 Estrategias de ventas

Por un lado, la desactivación de productos con bajas ventas puede liberar recursos valiosos, como espacio en inventario y costos de producción. Mantener productos que no se venden adecuadamente implica un costo de oportunidad, ya que esos recursos podrían ser reasignados a productos más rentables o a nuevas oportunidades de mercado. Al eliminar estos productos, la empresa puede mejorar su eficiencia operativa y concentrarse en aquellos que realmente impulsan el crecimiento y la rentabilidad.

No obstante, la decisión de desactivar un producto también conlleva riesgos. Uno de los impactos más inmediatos es la posible pérdida de ingresos, especialmente si el producto, aunque de baja rotación, tiene un grupo fiel de consumidores. La retirada de un producto puede resultar en la disminución de la lealtad del cliente, especialmente si el producto discontinuado era el único en la cartera que satisfacía sus necesidades específicas. Esto podría llevar a la migración de esos clientes hacia competidores que aún ofrecen ese tipo de producto. Además, el inventario residual no vendido puede convertirse en un problema financiero y logístico, incrementando los costos de almacenamiento y gestión de productos obsoletos.

Antes de desactivar un producto, la empresa podría realizar un análisis costo-beneficio que tenga en cuenta no solo el rendimiento en ventas, sino también el ciclo de vida del producto, su contribución a la imagen de marca y el costo de oportunidad de mantenerlo en el portafolio. Este análisis permitiría evaluar si el producto tiene el potencial de generar rentabilidad en el futuro con una inversión en marketing o si, por el contrario, su eliminación sería más beneficiosa a largo plazo. De esta manera, la decisión estaría respaldada por un enfoque más cuantitativo y estratégico, que minimice los riesgos financieros y logísticos que mencionas.

### **3.9 El valor de vida de un cliente**

Otro aspecto crucial en la gestión de una empresa es el cálculo del Valor de Vida del Cliente (*Customer Lifetime Value, CLV*). Este indicador nos permite medir cuánto valor genera un cliente para la empresa durante todo el tiempo que permanece activo. El *CLV* se puede calcular tomando en cuenta varios factores, como el historial de compras de cada cliente, la frecuencia de compra a lo largo de los meses, el valor promedio de las ventas y los ingresos totales generados por el cliente.

Un enfoque efectivo para determinar el *CLV* es analizar los ingresos totales que un cliente ha generado para la empresa, el número de meses distintos en los que ha realizado compras y el valor promedio de cada compra. Esta información no solo nos proporciona una visión clara del valor actual del cliente, sino que también permite predecir su valor futuro (The Editorial Team., 2024). Al identificar a los clientes con un alto *CLV*, la empresa puede enfocarse en ofrecerles incentivos personalizados para aumentar su lealtad y frecuencia de compra. Esto, a su vez, garantiza un flujo constante de ingresos y fortalece la relación a largo plazo con los clientes.

En resumen, la gestión adecuada de la cartera de productos y la comprensión del *CLV* son elementos esenciales para maximizar la rentabilidad y asegurar el éxito a largo plazo de la empresa. Las decisiones estratégicas basadas en estos análisis, ya sea para desactivar o reactivar productos, o para enfocar los esfuerzos en clientes de alto valor, pueden tener un impacto significativo en la sostenibilidad y el crecimiento de la empresa en el mercado competitivo actual.

Para calcular el *CLV* usamos los siguientes criterios establecidos por Qualtrics (2021):

#### **3.9.1 Valor promedio de compra**

El valor promedio de compra representa la cantidad promedio de dinero que un cliente gasta en cada transacción. Nos sirve como una medida clave para entender cuánto están dispuestos a gastar los clientes en promedio en cada venta o interacción con la empresa. Este valor ayuda a identificar la

rentabilidad de cada compra. Si el  $V_p$  es alto, significa que cada compra genera un ingreso significativo, lo que puede ser una señal positiva de productos de alto valor o una estrategia de ventas efectiva.

$$V_p = \frac{V_t}{N_p}$$

Donde:

- $V_t$  es el valor total de las compras en un periodo.
- $N_p$  es el número de compras en ese periodo.

### 3.9.2 Frecuencia promedio de compra

Una alta frecuencia de compra indica que los clientes regresan frecuentemente, lo cual puede ser un indicativo de satisfacción del cliente, lealtad o necesidad constante de los productos ofrecidos. Por otro lado, una baja frecuencia puede señalar problemas en la retención de clientes o que los productos tienen ciclos de compra largos. Esta métrica es crucial para identificar oportunidades de aumentar la frecuencia de compra mediante estrategias de marketing, programas de fidelización, o introduciendo productos complementarios que incentiven compras adicionales.

$$F_p = \frac{N_p}{N_c}$$

Donde:

- $N_p$  es el número de compras en ese periodo.
- $N_c$  es el número de clientes individuales que realizaron una compra en ese periodo.

### 3.9.3 Valor del cliente

El valor del ciclo de vida (CLV) muestra cuánto vale un cliente para la empresa a lo largo de su relación con ellos. Le ayuda a comprender qué tan rentable (o no) es un cliente o segmento de clientes en particular a lo largo de toda su relación con la empresa.

$$V_c = F_p \times V_p$$

Donde:

- $F_p$  es la frecuencia promedio de compra.
- $V_p$  es el valor promedio de compra.

$V_c$  es útil para identificar a los clientes más valiosos de la empresa. Si un cliente tiene un  $V_c$  alto, es probable que contribuya significativamente a los ingresos totales de la empresa.

### 3.9.4 Resultados CLV

id_cliente <int>	total_revenue <dbl>	total_purchases <int>	first_purchase <date>	last_purchase <date>	avg_purchase_value <dbl>	clv <dbl>
8342	82336717	553	2021-01-01	2024-08-01	148890.99	294856746
8635	69064926	411	2021-01-01	2024-08-01	168041.18	247329016
8318	64715454	629	2021-01-01	2024-08-01	102886.25	231753085
9066	45396012	760	2021-01-01	2024-08-01	59731.59	162568059
2456	41258718	234	2021-01-01	2024-07-01	176319.31	144250193
574	54975598	191	2021-01-01	2023-03-01	287830.36	118756323
7624	31008196	207	2021-01-01	2024-07-01	149798.05	108411955
7833	29541181	409	2021-01-01	2024-08-01	72227.83	105790184
2893	26965385	283	2021-01-01	2024-08-01	95284.05	96565978
7796	28502151	117	2021-01-01	2024-05-01	243608.13	94890117

Figure 6. Clientes con un CLV alto

Estos son los resultados utilizando las fórmulas dichas previamente. En la Figura 6 podemos apreciar los clientes con el CLV más alto.

#### ¿Qué significa cuando el CLV de un cliente es alto?

Cuando el **Customer Lifetime Value (CLV)** de un cliente es alto, significa que ese cliente genera un valor significativo para la empresa a lo largo del tiempo. Este valor puede derivarse de varias características, como:

- **Frecuencia de Compra Alta:** El cliente realiza compras frecuentes, lo que incrementa el flujo constante de ingresos.
- **Valor de Compra Alto:** El cliente tiende a gastar más en cada transacción en comparación con otros clientes.
- **Larga Duración de Relación:** El cliente ha estado comprando productos o servicios durante un período prolongado, lo que refleja lealtad y compromiso con la marca.
- **Bajo Costo de Mantenimiento:** Es posible que estos clientes requieran menos esfuerzo de marketing o atención al cliente, lo que los hace más rentables en comparación con otros clientes.

#### ¿Por qué es importante un CLV alto?

- **Rentabilidad:** Un cliente con un CLV alto contribuye significativamente a los ingresos y beneficios de la empresa.
- **Lealtad:** Estos clientes suelen ser más leales y es probable que sigan comprando en el futuro, lo que permite a la empresa planificar mejor su estrategia de retención.
- **Menor Costo de Adquisición:** La empresa puede invertir menos en atraer a estos clientes nuevamente, ya que suelen regresar por cuenta propia o con promociones limitadas.
- **Referencia de Otros Clientes:** Los clientes con un CLV alto también suelen recomendar la empresa a otros, lo que puede atraer a nuevos clientes de manera orgánica.



Un CLV alto indica que un cliente es valioso para la empresa y debería considerarse una prioridad en las estrategias de retención y fidelización.

id_cliente <int>	total_revenue <dbl>	total_purchases <int>	first_purchase <date>	last_purchase <date>	avg_purchase_value <dbl>	clv <dbl>
1103	121190.16	4	2021-06-01	2021-06-01	30297.540	0.000
3594	6447.10	1	2024-02-01	2024-02-01	6447.100	0.000
7978	676451.78	1	2022-08-01	2022-08-01	676451.780	0.000
8209	0.00	2	2021-06-01	2021-07-01	0.000	0.000
8333	2719.68	1	2021-02-01	2021-02-01	2719.680	0.000
8240	14439.64	2	2021-01-01	2021-02-01	7219.820	1225.541
3452	12276.36	8	2023-04-01	2024-06-01	1534.545	14351.829
305	18598.06	5	2022-09-01	2023-08-01	3719.612	17006.850
3451	23832.84	16	2023-08-01	2024-07-01	1489.553	21859.005
3440	36533.84	26	2023-06-01	2024-07-01	1405.148	39609.584

**Figure 7.** Clientes con el CLV negativo o bajo

Estos son los resultados utilizando la formula dicha previamente. En la Figura 7 podemos apreciar los clientes con el CLV más bajo.

### ¿Qué significa cuando el CLV de un cliente es bajo?

Cuando el **Customer Lifetime Value (CLV)** de un cliente es bajo, indica que ese cliente genera un valor relativamente bajo para la empresa a lo largo del tiempo. Esto puede deberse a varios factores:

- **Baja Frecuencia de Compra:** El cliente realiza compras poco frecuentes, lo que limita el flujo constante de ingresos.
- **Valor de Compra Bajo:** El cliente tiende a gastar menos en cada transacción en comparación con otros clientes.
- **Relación Corta con la Empresa:** El cliente ha comprado productos o servicios por un período limitado, lo que refleja una baja lealtad o poca conexión con la marca.

### ¿Por qué es problemático un CLV bajo?

- **Baja Rentabilidad:** Los clientes con CLV bajo no contribuyen significativamente a los ingresos y beneficios de la empresa.
- **Mayor Costo de Adquisición:** Si el costo de adquirir estos clientes es alto y no se logra compensar con un CLV suficiente, la empresa podría estar perdiendo dinero en estas relaciones.
- **Dificultad para Retenerlos:** Es probable que estos clientes sean más difíciles de retener, ya que no tienen una relación fuerte con la empresa o no encuentran un valor claro en los productos o servicios.
- **Escaso Impacto en el Crecimiento:** Dado que aportan poco valor, estos clientes no impulsan el crecimiento a largo plazo de la empresa ni ayudan a atraer nuevos clientes mediante recomendaciones.

En resumen, un CLV bajo indica que el cliente no está aportando un valor significativo a la empresa, lo que puede requerir una reevaluación de las estrategias de marketing, retención o incluso desinversión en estos clientes.

## 4 DISCUSIÓN DE RESULTADOS

A partir del análisis realizado, identificamos un patrón preocupante: un número significativo de clientes pasa de ser activo a inactivo después de haber estado activos durante varios meses. Esto representa un desafío para la empresa, ya que estos clientes no muestran un comportamiento de compra constante, lo que puede afectar negativamente los ingresos y la planificación de inventarios.

El paso de clientes de activos a inactivos sugiere que, aunque inicialmente satisfechos, los clientes no mantienen una relación continua con la empresa. Esta fluctuación podría deberse a la competencia, a la falta de incentivos para continuar comprando, o a otros factores externos que afectan su comportamiento de compra.

Para mitigar este problema y fomentar la lealtad del cliente, proponemos la implementación de un programa de domiciliación automática. Este programa permitiría a los clientes suscribirse a un plan donde se les cobraría automáticamente cada mes por productos o servicios seleccionados, asegurando así una compra regular sin necesidad de que el cliente realice un pedido manualmente cada vez.

La implementación de un programa de domiciliación automática ofrecería varios beneficios significativos para la empresa. En primer lugar, garantizaría ingresos constantes y predecibles, lo cual es crucial para la planificación financiera y operativa. Además, al automatizar las compras, se fomentaría la lealtad del cliente, ya que se establecerían como compradores recurrentes, mejorando así la retención y fortaleciendo la relación a largo plazo con la empresa. Al mismo tiempo, se reduciría el riesgo de que los clientes opten por la competencia, ya que recibirían sus productos o servicios de manera automática y sin esfuerzo adicional. Finalmente, ofrecería una mejor experiencia al cliente al eliminar la necesidad de realizar pedidos frecuentes, lo que incrementaría la satisfacción y la fidelidad hacia la marca.

## 5 CONCLUSIONES

Este análisis de activación y desactivación de productos utilizando un modelo de cadenas de Markov nos ha permitido identificar patrones clave en el comportamiento de los productos dentro del portafolio de PiSA. Los resultados muestran que aproximadamente el 80% de los productos presentan cadenas ergódicas, lo que indica estabilidad a largo plazo, mientras que el resto de los productos son más impredecibles y requieren estrategias de ventas específicas. Las recomendaciones estratégicas derivadas de este estudio incluyen la optimización de inventarios y la desactivación de productos con baja probabilidad de reactivación, así como el enfoque en la retención de clientes de alto valor a través de programas de fidelización y domiciliación automática.

Este análisis resalta la importancia de tomar decisiones basadas en datos para mejorar la rentabilidad y eficiencia operativa. La implementación de un enfoque más sistemático en la gestión del portafolio de productos puede ayudar a PiSA a mantener su liderazgo en el mercado y mejorar su capacidad para responder a las fluctuaciones de demanda. Además, el cálculo del CLV ofrece una herramienta valiosa para personalizar estrategias de marketing, enfocándose en los clientes más rentables y maximizando su valor a lo largo del tiempo..

## **A MATERIAL SUPLEMENTARIO**

El código completo utilizado para el desarrollo del proyecto se encuentra en el siguiente repositorio de Drive: **Repositorio**.

## **REFERENCES**

Dr. Garrido J. (2024). Cadenas de markov.

Qualtrics (2021). How to calculate customer lifetime value.

The Editorial Team. (2024). What is customer lifetime value (clv) how to calculate it.