



Universidade de Vigo

Trabajo Fin de Máster

Simulador Ciclo de Tienda

Maria Erika Francesch Domenech

Máster en Técnicas Estadísticas

Curso 2023-2024

Propuesta de Trabajo Fin de Máster

| |
|--|
| Título en galego: Simulador Ciclo de Tenda |
| Título en español: Simulador Ciclo de Tienda |
| English title: Store Cycle Simulator |
| Modalidad: Modalidad B |
| Autora: Maria Erika Francesch Domenech, Universidade da Coruña |
| Directores: Rosa María Crujeiras Casais, Julio González Díaz, Universidade de Santiago de Compostela |
| Tutora: Carmela Cuevas García, Zara-Inditex |
| Breve resumen del trabajo: Construcción de un modelo que simule la situación de la tienda, incluyendo todo tipo de operativas, para mejorar la eficiencia de las tiendas ZARA, priorizando la optimización de los envíos sin perjudicar las ventas, minimizando las unidades devueltas al almacén. |

Doña Rosa María Crujeiras Casais, Catedrática de la Universidade de Santiago de Compostela, don Julio González Díaz, Profesor titular de la Universidade de Santiago de Compostela, doña Carmela Cuevas García, Responsable Control de Gestión Comercial de Zara-Inditex informan que el Trabajo Fin de Máster titulado

Simulador Ciclo de Tienda

fue realizado bajo su dirección por doña Maria Erika Francesch Domenech para el Máster en Técnicas Estadísticas. Estimando que el trabajo está terminado, dan su conformidad para su presentación y defensa ante un tribunal.

En A Coruña, a 3 de Junio de 2024.

La directora:
Doña Rosa María Crujeiras Casais

El director:
Don Julio González Díaz

La tutora:
Doña Carmela Cuevas García

La autora:
Doña Maria Erika Francesch Domenech

Declaración responsable. Para dar cumplimiento a la Ley 3/2022, de 24 de febrero, de convivencia universitaria, referente al plagio en el Trabajo Fin de Máster (Artículo 11, [Disposición 2978 del BOE núm. 48 de 2022](#)), **el/la autor/a declara** que el Trabajo Fin de Máster presentado es un documento original en el que se han tenido en cuenta las siguientes consideraciones relativas al uso de material de apoyo desarrollado por otros/as autores/as:

- Todas las fuentes usadas para la elaboración de este trabajo han sido citadas convenientemente (libros, artículos, apuntes de profesorado, páginas web, programas, . . .)
- Cualquier contenido copiado o traducido textualmente se ha puesto entre comillas, citando su procedencia.
- Se ha hecho constar explícitamente cuando un capítulo, sección, demostración, . . . sea una adaptación casi literal de alguna fuente existente.

Y, acepta que, si se demostrara lo contrario, se le apliquen las medidas disciplinarias que correspondan.

Agradecimientos

Quiero expresar mi profundo agradecimiento por la oportunidad de realizar mi trabajo de fin de máster al grupo de Zara Control de Gestión Comercial en Inditex. Fue una experiencia enriquecedora, y agradezco especialmente a mi tutora, Mela, por permitirme unirme a su equipo de trabajo y contribuir activamente al desarrollo del proyecto.

A mi tutor, Julio, le estoy especialmente agradecida por compartir sus amplios conocimientos, por guiarme en el diseño del proyecto, por su implicación y, sobre todo, por su infinita paciencia en cada reunión que hemos tenido.

A mi tutora Rosa, por sus valiosos consejos, guía y ayuda en la estructura del trabajo, así como por su asesoramiento académico y sus aportes cruciales en el análisis.

A Belén, que siempre estuvo disponible para aclarar todas mis dudas y ofrecerme su asesoramiento. Asimismo, a todo el equipo de Zara Control de Gestión del Retail y Zara Business Analytics, quienes desde el primer momento me hicieron sentir parte del grupo y colaboraron activamente en la resolución de todas mis cuestiones.

Agradezco también a cada uno de mis profesores, quienes me proporcionaron las herramientas y conocimientos necesarios para la realización de este trabajo.

Finalmente, mi más sincero agradecimiento a mi familia, por su amor y apoyo incondicional, y a mis amigas, siempre presentes y apoyándome en cada paso.

Índice general

| | |
|---|-----------|
| Resumen | XI |
| 1. Introducción | 1 |
| 1.1. Grupo Indítex | 2 |
| 1.1.1. Zara | 2 |
| 1.1.2. Departamentos de compra de Zara Señora | 2 |
| 1.2. Problemática | 3 |
| 1.3. Objetivo | 4 |
| 1.4. Descripción de los datos | 4 |
| 1.4.1. Obtención y procesamiento de los datos | 5 |
| 1.4.2. Definición de las variables | 5 |
| 1.5. Análisis descriptivo | 7 |
| 1.5.1. Análisis de las devoluciones | 7 |
| 1.5.2. Análisis del destocaje de cliente | 9 |
| 1.5.3. Análisis del destocaje sin cliente | 10 |
| 2. Análisis de Comportamiento de los Compradores | 13 |
| 2.1. Introducción métodos clustering | 13 |
| 2.2. Tipos de clúster | 14 |
| 2.2.1. k -means | 14 |
| 2.2.2. PAM | 14 |
| 2.2.3. Jerárquico | 15 |
| 2.3. Evaluación y validación de los métodos de clúster | 16 |
| 2.3.1. Validación interna | 17 |
| 2.3.2. Estabilidad | 17 |
| 2.4. Objetivo | 19 |
| 2.5. Procesamiento y validación | 19 |
| 2.6. Resultados | 23 |
| 2.7. Conclusiones y consideraciones para la aplicación práctica | 27 |
| 3. Modelo Ciclo de tienda | 29 |
| 3.1. Problema y objetivo | 29 |
| 3.1.1. Descripción del problema | 29 |
| 3.1.2. Objetivo del modelo | 30 |
| 3.2. Descripción del modelo | 30 |
| 3.3. Desarrollo del modelo | 31 |
| 3.3.1. Conjunto | 31 |
| 3.3.2. Parámetros | 31 |
| 3.3.3. Variables | 32 |
| 3.3.4. Función objetivo y restricciones | 32 |

| | |
|---|-----------|
| 3.3.5. Características del problema de optimización | 33 |
| 3.4. Estudio computacional | 34 |
| 3.5. Validación y resultados | 35 |
| 3.6. Interpretación del modelo de ciclo de tienda | 38 |
| 3.7. Conclusiones y Consideraciones para su Aplicación Práctica | 39 |
| 4. Conclusiones del Análisis y Consideraciones para su Aplicación Práctica | 41 |
| A. Glosario de Variables de negocio | 43 |
| B. Tablas | 47 |
| Bibliografía | 55 |

Resumen

Resumen en español

Este proyecto desarrolla un modelo de la situación del ciclo de tienda para optimizar la gestión de las tiendas ZARA y analizar el comportamiento de los departamentos de compra. Las tiendas, que reciben mercancía dos veces por semana, deben gestionar eficientemente el espacio limitado de exposición para integrar nuevos productos y manejar devoluciones, reubicando o retornando otros artículos al centro de distribución.

El objetivo de maximizar el espacio de exposición lleva a simular y evaluar estrategias operativas que priorizan los departamentos de compra según su desempeño. Este enfoque es crucial para optimizar el flujo del producto y minimizar devoluciones.

Adicionalmente, el proyecto incluye un análisis del comportamiento de los departamentos de compra en cada tienda. Este estudio se realiza a través del análisis de agrupación para comprender mejor las decisiones de compra y las tendencias en cada mercado. El conocimiento adquirido permite identificar patrones clave que pueden influir en la selección de productos, ayudando a los departamentos a ajustar sus estrategias de compra según las necesidades del mercado y las tendencias de consumo.

El modelo simula diversas configuraciones de los envíos a la tienda, evaluando su impacto en la logística y gestión de inventario, y proporciona insights sobre cómo los ajustes pueden afectar las operaciones generales.

English abstract

This project develops a simulation model to optimize the management of ZARA stores and analyze the behavior of the purchasing departments. The stores, which receive merchandise twice a week, must efficiently manage limited space to integrate new products and handle returns, relocating or returning other items to the distribution center.

The goal of maximizing display space leads to simulate and evaluate operational strategies that prioritize purchasing departments based on their performance. This approach is crucial for optimizing product flow and minimizing returns.

Additionally, the project includes an analysis of the behavior of the purchasing departments in each store. This study is conducted through cluster analysis to better understand purchasing decisions and internal trends. The knowledge gained enables the identification of key patterns that can influence product selection, assisting departments in adjusting their purchasing strategies according to market needs and consumption trends.

The model simulates various store configurations, evaluating their impact on logistics and inventory management, and provides insights on how adjustments can affect general operations.

Capítulo 1

Introducción

En este primer capítulo se ofrece una visión general del entorno en el que se sitúa el proyecto: el sector de la moda y el retail, dentro del grupo Inditex y su principal marca Zara donde se proporcionarán todos los conceptos necesarios del modelo de negocio, el cual se complementa con el [Anexo A](#) donde se encuentran el glosario de las definiciones necesarias para la comprensión del trabajo de fin de máster.

La moda es una industria vital que influye en economías globales y tendencias culturales, en las que empresas como Inditex lideran este campo con innovaciones en diseño y gestión. En un mundo dominado por rápidos cambios en las preferencias de consumo y avances tecnológicos continuos, Inditex, a través de su principal marca, Zara, ha revolucionado el sector de la moda con su modelo de negocio centrado en el cliente y su excepcional capacidad para responder a las demandas del mercado con rapidez y eficacia, caracterizado por la rápida rotación de inventario y actualización de productos en tienda.

Sin embargo, este modelo presenta desafíos logísticos significativos, especialmente en la gestión eficiente del espacio y en el inventario de las tiendas. Las tiendas de Zara enfrentan un flujo constante de mercancías nuevas, reposiciones y devoluciones de clientes, lo que complica la gestión del limitado espacio de almacenamiento y exposición.

Con la introducción de un nuevo concepto de tienda que maximiza el espacio de exposición reduciendo el almacén, Zara busca mejorar la experiencia de compra pero, al mismo tiempo, esto incrementa la complejidad en la gestión de stock. Este escenario crea un 'efecto expulsión' donde el espacio para nuevos artículos se hace cada vez más escaso y requiere de estrategias sofisticadas para 'limpiar el stock' de manera eficiente, devolviendo menos unidades al centro de distribución y manteniendo o incrementando las ventas.

El objetivo de este proyecto es desarrollar y validar un modelo que permita analizar y simular las operaciones de la tienda bajo diferentes escenarios, considerando variables como número de envíos, peso de las devoluciones de clientes, unidades destocadas, y nivel de ventas. Este trabajo se enfocará en dos componentes fundamentales de la estadística aplicada: la optimización y la estadística de agrupaciones. Por ello, el desarrollo del modelo buscará optimizar los procesos de envío y exceso de stock sin comprometer el desempeño comercial de la venta, contribuyendo a una gestión más eficiente de las tiendas Zara. Los resultados esperados incluyen una clasificación y un modelo de ciclo de tienda que detalla los procesos de la tienda y estrategias de optimización para la mejora de la eficiencia operativa y la satisfacción del cliente.

1.1. Grupo Inditex

El grupo Inditex¹ (Industria de Diseño Textil SA.) es el mayor distribuidor de moda del mundo. Con sede central en Arteixo (A Coruña) está formado por las distintas **cadena**s: Zara, Zara Home, Massimo Dutti, Pull & Bear, Bershka, Oysho y Stradivarius. Contando con una presencia en 213 **mercados** mediante canales online y sus más de 5800 tiendas.

1.1.1. Zara

Zara es la primera marca y buque insignia del grupo, donde las ventas de Zara y Zara Home han representado en 2023 un 72.5 % total de la facturación del grupo Inditex. La primera tienda de Zara se abre en 1975 en la Calle Juan Flórez en A Coruña y en 1985 se crea el grupo Inditex con el fin de agrupar la producción y ventas del resto de cadenas. Revolucionando por completo el sector de la distribución textil por los siguientes motivos:

- La tienda y el cliente son el eje central del modelo de negocio, por lo que los trabajadores en tienda representan un 86 % del personal que forma Inditex, adaptándose a las expectativas y exigencias del cliente de una manera eficaz, recopilando la información obtenida en la **tienda física** como en el canal online.
- Relación calidad-precio, ofreciendo las últimas tendencias al cliente a precios razonables de forma continua y rápida.
- Renovación de stock 2 veces por semana, con novedades y reponiendo las unidades vendidas, mediante un sistema de logística único contando con 12 centros de distribución que operan en un tiempo máximo de entrega de 48 horas.
- Rapidez en el proceso de producción mediante la distribución textil centralizada. Acortando el tiempo a 3 semanas desde el diseño, a la disposición del producto para la venta, reaccionando rápidamente a las nuevas tendencias y necesidades de cada mercado. Para ello, cuenta con más de 700 diseñadores para adaptarse a las demandas de los clientes.

Zara se divide en 3 secciones: Señora, Caballero y Niño. El proyecto de fin de máster se centra en el estudio del comportamiento del ciclo de tienda de la sección de Zara Señora, la cual se divide en tres áreas: Zara Woman (WOMAN), BASIC y Trafaluc (TRF). Zara Woman es la sección imagen de la marca caracterizada por la calidad de los tejidos. La sección BASIC se caracteriza por ser una línea más casual con prendas básicas y TRF es la sección tendencia destinada a un público más joven o que buscan modelos más atrevidos.

1.1.2. Departamentos de compra de Zara Señora

Los departamentos de compra² juegan un papel fundamental en la cadena de negocio de venta, ya que son los responsables desde el lanzamiento y seguimiento de la producción hasta la puesta en venta del producto en la tienda.

Los principales departamentos de compra de Zara Señora, los cuales se denominarán **compradores** a partir de este punto, son: WOMAN, BASIC, TRF, PUNTO y CIRCULAR. Por lo tanto, no se debe confundir con las personas individuales que acuden a la tienda, ya que estos son denominados **clientes**.

A continuación se realiza una descripción de cada uno de estos compradores:

¹<https://www.inditex.com/itxcomweb/es/home>

²<https://www.inditexcareers.com/portalweb/es/home>

- **WOMAN** es el comprador que destaca por la calidad de sus tejidos, captando las tendencias más sofisticadas donde el precio unidad/prenda es más elevado.
- **TRF** este comprador esta caracterizado por ser una línea más joven y atrevida en cuanto a tendencias, el precio por unidad es el menos elevado.
- **BASIC** es el comprador caracterizado por propuestas de prendas más básicas, donde se pueden encontrar diseños atemporales, su precio medio por unidad tiende a ser menos elevado que WOMAN.
- **PUNTO** lo componen prendas de punto que se encuentran representadas en todas las secciones de la tienda, debido a que fabrica para las 3 principales áreas: WOMAN, BASIC y TRF. Complementando de esta manera a las distintas colecciones. Su coste por unidad viene marcado por la calidad del tejido de punto. La representación de este comprador vendrá marcada por la estacionalidad dependiendo del tipo de [campaña](#).
- **CIRCULAR** este comprador debe su nombre al tipo de máquina que se utiliza para la obtención del tejido circular ³. Este comprador también fabrica para las 3 principales áreas, por lo que se encontrarán sus prendas distribuidas en todas las secciones de la tienda. Se caracteriza principalmente de producir monoproducto, camisetas y prendas de material de felpa.

Los compradores WOMAN y TRF son los encargados de marcar la tendencia de la tienda. Esto dependerá de la tipología de cliente y de la localización de la tienda, ya que ciertos tipos de productos pueden ser mejor comprendidos y recibidos por el cliente dependiendo de estos factores. Los compradores BASIC y CIRCULAR son los compradores con un volumen de entrada mayor por unidades en la tienda, lo cual cabe esperar que tenga también un mayor volumen de salida mediante venta y [destocajes](#).

Los compradores WOMAN, BASIC y TRF sacan sus colecciones representando una tipología de cliente, ya que cada uno de estos compradores divide su colección atendiendo a los diferentes estilos de mujeres⁴, segmentando así a la clientela de Zara dentro de sus tres áreas de diseño. Por ello, la tienda se distribuirá en 3 principales secciones diferenciadas por estos tres compradores, mientras que los compradores PUNTO y CIRCULAR se verán representados a lo largo de toda la tienda integrándose dentro de los conceptos de las colecciones de las tres principales áreas.

1.2. Problemática

Desde el área de Control de Gestión Retail que forma parte del área de Control de Gestión Comercial de Zara, se propone un proyecto que simule la situación del [ciclo de tienda](#). Que represente todas las entradas y salidas de los artículos a la tienda. Primero los artículos entran a la tienda como unidades enviadas desde el [centro de distribución](#) y salen como ventas o como destocajes, pero los productos vendidos pueden volver a entrar en el ciclo de tienda como devoluciones y los productos destocados pueden volver a entrar para su [reaprovechamiento](#) antes de las rebajas o por unificación de producto. El ciclo de tienda se describe en profundidad en el Capítulo 3 mediante la Figura 3.1.

Este departamento se encarga de analizar el retail mediante los procesos operativos y estratégicos de Zara. Para ello, colabora estrechamente con diversos departamentos dentro de la compañía, incluidos los equipos de Dirección Comercial, Dirección de Producto, compradores, product managers, estudio, Merchandising, Distribución, Internacional, Proyectos y Sistemas.

³Consiste en la fabricación de tela mediante uno o varios hilos formando una malla transversal, produciendo tejidos tubulares.

⁴ZARA Camp Training Program-Formación Interna de la empresa

La necesidad del proyecto surge con la idea de mejorar su estrategia de negocio para ayudar a replantear las estrategias de envío y venta. Esto es debido a que las tiendas ZARA reciben mercancías dos veces por semana, incluyendo nuevos productos y reposiciones, además de gestionar las devoluciones de los clientes.

Estos procesos requieren una administración efectiva del espacio de exposición de venta y el limitado espacio de almacenaje, necesitando una integración adecuada de los nuevos artículos y una reubicación o retorno al centro de distribución de otros. Afrontando una problemática a nivel destocajes, lo que conlleva a planificar una estrategia donde la optimización del envío prevalezca en que las ventas se maximicen minimizando la expulsión de producto de vuelta a almacén.

1.3. Objetivo

Este proyecto se centra en proporcionar las herramientas necesarias para la mejora de la estrategia comercial de Zara Señora a través de un análisis exhaustivo del ciclo de la tienda. Mediante un estudio de los compradores principales de Zara Señora y el desarrollo de un modelo de optimización del envío de unidades, se busca minimizar los destocajes y maximizar las ventas. Para alcanzar este propósito, el trabajo de fin de máster se divide en 2 partes:

- **Análisis del comportamiento de los compradores:** se realiza un análisis detallado de los distintos compradores de Zara Señora en función del mercado, atendiendo así a las ventas, devoluciones y destocajes. Este análisis se realizará mediante la aplicación de técnicas de agrupación, tipo clúster, para identificar patrones y tendencias en los datos. Los modelos necesarios para llevar a cabo este análisis serán proporcionados y detallados en el [Capítulo 2](#), el cual permitirá analizar y reaccionar de manera visual y rápida, para poder realizar un seguimiento semanal o planificar una estrategia de compra en futuras campañas, atendiendo a la información de las agrupaciones obtenidas. Este algoritmo se desarrolla mediante Rstudio [2] versión 4.2.1 generando un archivo en Excel cada vez que se ejecuta, para determinar de manera rápida a que grupo pertenece cada comprador dentro de cada tienda.
- **Desarrollo de un modelo de ciclo de tienda:** se desarrolla un modelo de ciclo de tienda para optimizar el envío de unidades con el fin de aumentar las ventas con respecto al año anterior y minimizar los destocajes, evitando problemas de operativa. Estos modelos se implementarán y desarrollarán mediante el programa AMPL IDE (A mathematical language program) [1], se desarrolla también en Solver, programa de complemento de Microsoft Excel, para que resulte más fácil de interactuar desde los distintos departamentos, lo que beneficia al usuario final. Esta parte del trabajo se describe en el [Capítulo 3](#).

Al cumplir con estos objetivos, se espera proporcionar a Control de Gestión Comercial de Zara las herramientas necesarias para una toma de decisiones más informada y estratégica respecto a la gestión del ciclo de la tienda. La implementación práctica de las estrategias resultantes queda a cargo del equipo operativo, quienes podrán utilizar y seguir desarrollando los modelos, que con los insights obtenidos podrán optimizar la gestión diaria y planear las estrategias de futuras campañas.

1.4. Descripción de los datos

A continuación se detalla la obtención y el procesamiento de los datos necesarios para el trabajo de fin de máster, junto con la descripción de las variables utilizadas en cada uno de los capítulos.

1.4.1. Obtención y procesamiento de los datos

La obtención de los datos de estudio ha sido un proceso que ha implicado a varios departamentos de la empresa. Desde Control de Gestión Comercial se ha tenido acceso a los datos mediante el equipo de Business Analytics que proporcionó el acceso a varias tablas maestro en formato SQL y tablas de reaprovechamiento vía PowerBI, también se ha accedido a los datos mediante la aplicación del departamento correspondiente de Control de Gestión Comercial, para acceder a datos históricos de continuidad de artículos y de campaña acumulados junto con las bases de datos proporcionadas por Control de Gestión del Retail en Excel, con los datos de campaña acumulados.

Para la realización del trabajo de fin de máster se realiza una selección de los datos para trabajar finalmente con 3 bases de datos, el acumulado de ventas y devoluciones de la campaña Invierno 2022 y 2023, el flujograma de negocio de las mismas campañas para obtener los datos de reaprovechamiento de cliente y el coste unitario de destocaje en tienda.

Mediante estas bases de datos proporcionadas se procesan los datos de toda la campaña de Invierno 2023, exceptuando el período de rebajas. Junto con estos datos se agregan los de la campaña de Invierno 2022, la cual se analiza para hacer una comparativa en el comportamiento de los compradores a nivel devoluciones y destocajes. Esta base de datos representa las unidades acumuladas desde el inicio de campaña y que marcan el éxito de las unidades enviadas y vendidas por mercado, [clasificación de tienda](#), tienda y comprador.

Todos estos datos son de todos las tiendas pertenecientes a todos los mercados operativos, para todos los compradores de las 3 secciones de Zara: Señora, Caballero y Niño. Por ello, para proceder al análisis se realiza un procesamiento que consta en una depuración para un posterior filtrado, donde se eliminan las tiendas que no han sido operativas en la campaña de Invierno 2023 debido a no ser una [tienda activa](#) en esa campaña. Se filtra por sección escogiendo Zara Señora, también mediante el tipo de producto limitando el análisis a artículos de ropa. Se filtra por mercado, en el caso del análisis descriptivo se analizan los mercados que producen más ventas y son más importantes para el negocio de Zara Señora, que son los que constituyen el Core Business: Alemania, España, Estados Unidos, Francia, Italia, Japón, Mainland China y Reino Unido. Por último se seleccionan por comprador, escogiendo los que producen más ventas para la sección de Zara Señora: WOMAN, BASIC, TRF, CIRCULAR y PUNTO.

Finalmente se procede a la creación de 3 nuevas variables: [ratio devolución](#), [ratio destocaje cliente](#) y [ratio destocaje sin cliente](#), las cuales se utilizarán para la agrupación clúster para el análisis del comportamiento de los compradores y el modelo de optimización, que se definen en la siguiente sección.

1.4.2. Definición de las variables

Las variables utilizadas a lo largo del proyecto son las siguientes:

- Unidades enviadas.
- Unidades vendidas: unidades vendidas sin descontar las devoluciones.
- **Éxito**: ratio de unidades de ventas en función de las unidades enviadas.
- Ratio devolución: unidades devueltas en función de las unidades vendidas.

$$\text{Ratio devolución} = \frac{\text{unidades devueltas}}{\text{unidades vendidas}}$$

- Destocajs cliente: destocaje producido por las devoluciones del cliente (véase a continuación la descripción detallada).

- Ratio destocaje cliente: unidades destocadas de cliente en función de las unidades enviadas.

$$\text{Ratio destocaje cliente} = \frac{\text{unidades destocaje cliente}}{\text{unidades enviadas}}$$

- Destocajes sin cliente: destocaje producido por las operativas de tienda. (véase a continuación la descripción detallada).
- Ratio destocaje sin cliente: unidades de destocaje sin cliente en función de las unidades enviadas.

$$\text{Ratio destocaje sin cliente} = \frac{\text{unidades destocaje sin cliente}}{\text{unidades enviadas}}$$

- Reaprovechamiento de cliente (R_C): porcentaje de las unidades vendidas en la tienda que proviene de devoluciones del cliente.
- Reaprovechamiento de destocaje (R_D): porcentaje de unidades vendidas que provienen de unidades destocadas que se vuelven a distribuir.
- Gestión del reaprovechamiento de cliente (G_C): porcentaje de gestión de reaprovechamiento de cliente.
- Gestión del reaprovechamiento de destocaje (G_D): porcentaje de reaprovechamiento de las unidades que provienen del destocaje sin cliente.

Para comprender y abordar eficazmente la problemática de los destocajes, se propone una división en dos categorías en función de como se producen:

- Destocaje cliente: este tipo de destocaje surge de las devoluciones del cliente de productos que ya no están en la zona de exposición de venta la tienda física, por tanto no se podrían reaprovechar en la tienda. Automáticamente al escanearlas en la zona de cajas se origina una devolución a almacén. Esta situación implica una carga adicional para el personal de cajas, lo cual conlleva a un problema de operativa. Este tipo de destocaje no se puede controlar en el aspecto de que el cliente dispone de 1 mes desde la compra para la devolución del artículo, y la rotación del producto de la tienda es dinámico adaptándose al nuevo producto. Además existe la posibilidad de devolución en tienda física de artículos comprados en otras tiendas del país, o adquiridos mediante el canal online. Esto implica que una tienda física reciba mediante devoluciones de cliente, artículos que nunca ha tenido en exposición de venta. Mediante un análisis de agrupación se busca identificar las tiendas y los compradores que tienen más predisposición a experimentar este tipo de devoluciones, lo que permitirá ajustar las estrategias de envío y operativas específicas para ese tipo de tiendas.
- Destocaje sin cliente: este tipo de destocaje surge de diversas operativas de la tienda, como agrupación de producto entre tiendas y problemas de [capacidad](#) en el almacén, entre otros. Este tipo de destocaje es controlado por las decisiones de negocio desde la central de la compañía. Por lo que puede ser analizado y gestionado considerando el tipo de comprador, el volumen de unidades y el ciclo de venta en la tienda. La simulación del ciclo de venta permite predecir las unidades que se deberían enviar a cada tienda, teniendo en cuenta las ventas esperadas, el destocaje y reaprovechamiento de estas unidades, basándose en datos históricos.

Cada tipo de destocaje producirá un reaprovechamiento de unidades distinto, es decir, las unidades de reaprovechamiento de cliente, serán las unidades devueltas por el cliente en tienda física que vuelven a venderse en tienda. Mientras que el reaprovechamiento producido por el destocaje sin cliente, son las unidades vendidas que vuelven a ponerse en venta después de haber sido destocadas.

1.5. Análisis descriptivo

A continuación se realiza un análisis descriptivo de las siguientes variables: unidades enviadas, unidades vendidas, unidades de devolución, unidades de destocaje de cliente, y unidades del destocaje sin cliente.

Primero se estudia la correlación y distribución de las variables de interés, mediante una matriz de gráficos de dispersión, donde se obtiene un análisis visual de la distribución de las variables y las correlaciones par a par. Tal como se aprecia en la Figura 1.1. existe una fuerte correlación positiva entre las unidades enviadas y vendidas. Mientras que las demás variables no indican correlación.

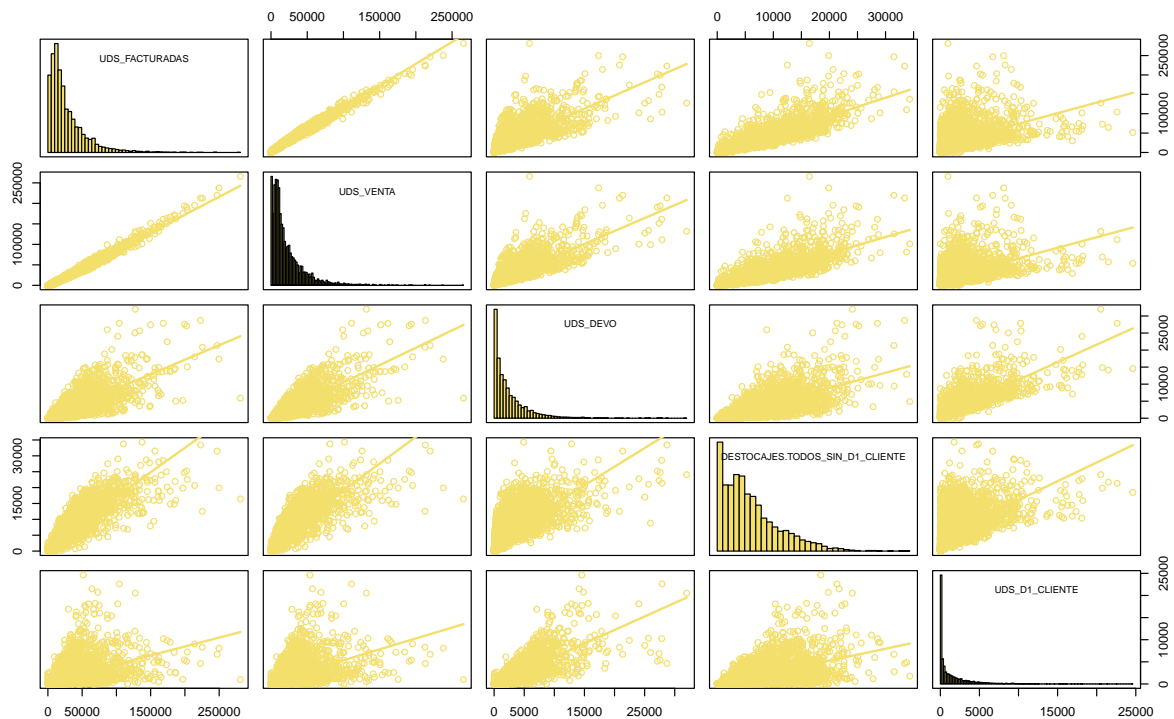


Figura 1.1: Relaciones por pares entre las variables de estudio. Variables por filas: unidades enviadas=facturadas, unidades venta bruta, unidades devolución, unidades de destocaje sin cliente y unidades de destocaje cliente. En la diagonal principal se ve representado el histograma de la distribución de cada variable, los gráficos fuera de la diagonal representan la relación par a par de las variables.

Contando con un número de tiendas a analizar total de 783, distribuidas de la siguiente manera por país, tal como indica la Tabla 1.1. Ordenados de mayor a menor por el número de tiendas, donde se observa que España es el mercado con más tiendas.

1.5.1. Análisis de las devoluciones

Se analiza cual es el comprador que en función de su venta recibe más devoluciones en la tienda en todos los mercados analizados del Core business. Mediante la Figura 1.2 izquierda, se observa que el comprador BASIC es el que recibió más devoluciones en tienda durante la campaña de Invierno 2023,

Tabla 1.1: Número de tiendas por país a analizar en el mercado Core Business.

| MERCADOS | TIENDAS |
|----------------|---------|
| ESPAÑA | 171 |
| FRANCIA | 113 |
| MAINLAND CHINA | 108 |
| ESTADOS UNIDOS | 104 |
| ITALIA | 91 |
| JAPÓN | 68 |
| ALEMANIA | 67 |
| REINO UNIDO | 61 |
| TOTAL | 783 |

seguido de CIRCULAR. Lo que se relaciona al ser los compradores que más unidades se envían y se venden en la tienda. Sin embargo, al realizar la comparación del ratio de devoluciones en función de las ventas, como se puede observar a la derecha de la Figura 1.2, los compradores que reciben más devoluciones en función de las ventas son WOMAN y TRF. Lo que indica que a nivel volumen se devuelven más unidades de BASIC y CIRCULAR en los mercados del Core Business, sin embargo si se compara en función de las ventas los compradores que experimentan mayores devoluciones son TRF y WOMAN. Esto es indicativo a que son compradores que experimentan un mayor número de devoluciones en función de sus ventas comparados con los demás.

Este comportamiento es el mismo que se observa en la campaña Invierno 2022, a diferencia del comprador WOMAN que experimenta menos devoluciones en unidades en Invierno 2023, debido a que sus ventas fueron menores con respecto al año anterior. Mientras que los demás compradores se comportan de la misma manera.

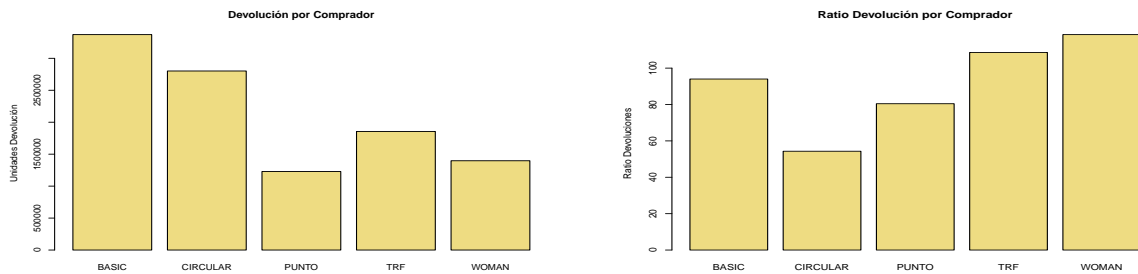


Figura 1.2: Izquierda: Representación de las unidades de Devolución por Comprador. Derecha: Representación del Ratio de Devolución en función de las ventas por Comprador.

Se realiza un estudio de la variable de Unidades de Devolución dependiendo del comprador y

mercado, se observa en la Figura 1.3 que el mercado con mayor volumen de devoluciones es España, seguido de Estados Unidos, Francia y Reino Unido. Siendo Alemania, Italia, Mainland China y Japón los países con menos devoluciones.

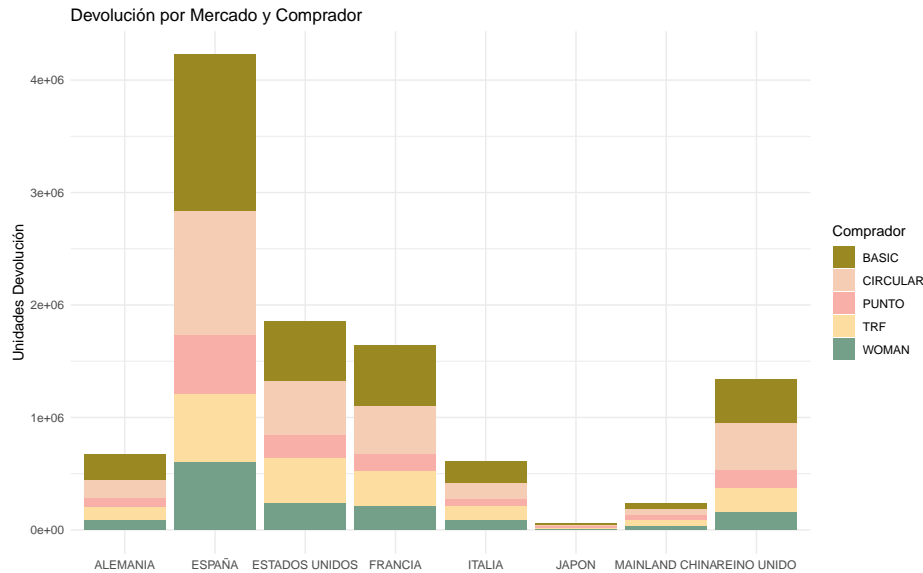


Figura 1.3: Representación de las unidades de Devolución por mercado y comprador.

1.5.2. Análisis del destocaje de cliente

Atendiendo a los destocajes de cliente, a la izquierda de la Figura 1.4 se observa la misma tendencia que en el caso de las devoluciones, donde BASIC es el comprador que más destocaje de cliente producen en el Core Business en unidades. Sin embargo, al ponerlo en función de las unidades enviadas, a la derecha de la Figura 1.4, TRF es el que tiene un ratio mayor, esto es debido a que hay una serie de tiendas⁵ localizadas en zonas donde la tipología de cliente no corresponde a TRF, lo que conlleva que reciban prendas de TRF cuando no están en exposición. Para la identificación de grupos de tiendas que sufren este tipo de destocaje de Cliente, se realizará una agrupación en clústeres que vendrá marcada por el destocaje y devoluciones del cliente.

En comparación con la campaña de Invierno 2022, se observa que los destocajes de cliente han aumentado hasta 4 veces para los compradores de BASIC y CIRCULAR y hasta 3 veces para el resto de compradores, lo que implica un problema de operativa en tienda. La razón de este aumento puede ser debido a un cambio de política de cobro en las devoluciones de artículos comprados vía online, donde su devolución en tienda no asume ningún costo al cliente. Por lo que este aumento podría explicarse a que la devolución de artículos de origen de compra online que terminan devueltos en tienda física, pero no están en exposición, provocando este aumento de destocaje de cliente.

En la Figura 1.5, se observa la división de destocajes de cliente en cada mercado, donde España sigue siendo el país con mayor tasa de destocajes de cliente, seguida de Estados Unidos y Reino Unido.

⁵Mercado España: Juan Flórez (A Coruña) y Constitución (Valladolid), Mercado Francia: Passy (París), Sevres (París), Mercado Reino Unido: Victoria (Londres)

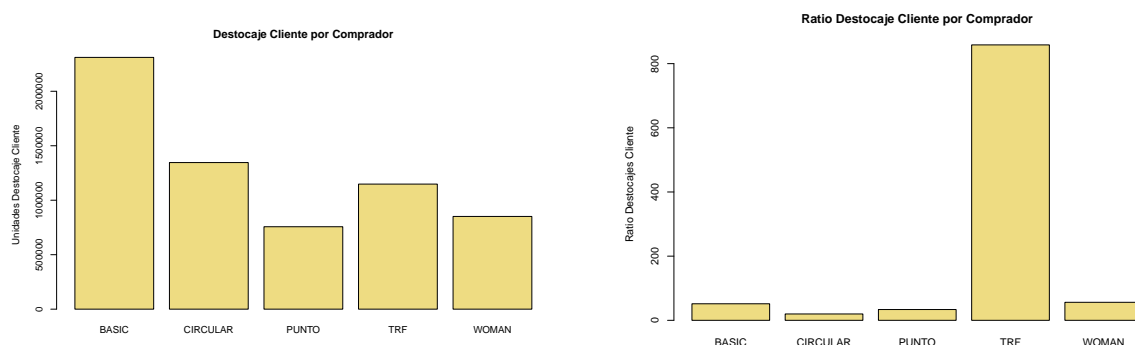


Figura 1.4: A la izquierda: Representación de Destocajes de Cliente en función de los compradores. A la derecha: Representación del Ratio de Destocajes de Cliente en función de los compradores.

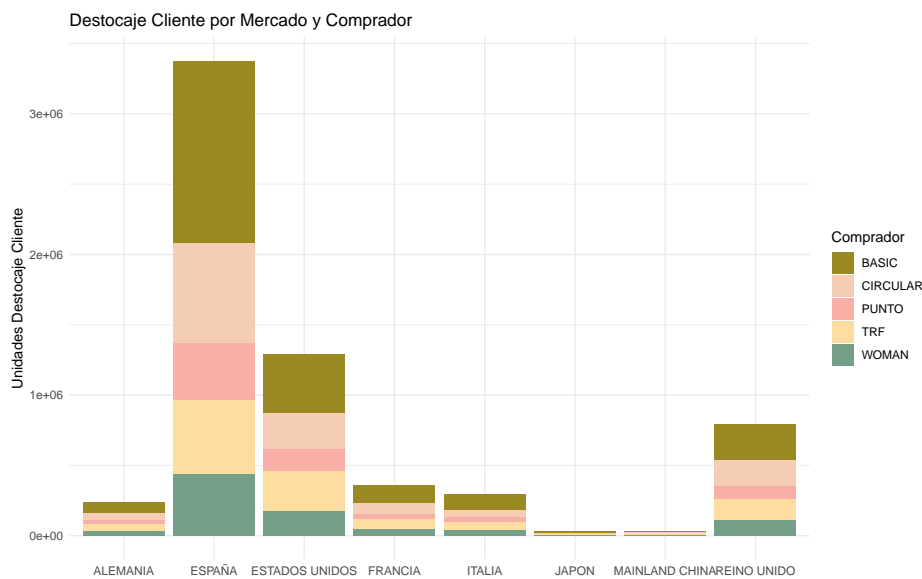


Figura 1.5: Representación del Ratio de Destocajes de Cliente por mercado y comprador.

1.5.3. Análisis del destocaje sin cliente

Atendiendo al destocaje sin cliente por comprador, a la izquierda de la Figura 1.6. Los compradores principales que originan el mayor destocaje en unidades son CIRCULAR seguido de BASIC. Lo que indica que estos dos compradores destocan a almacén un mayor volumen de unidades. Al realizar el ratio respecto a las unidades enviadas, Figura 1.6 a la derecha, TRF es el comprador mayoritario, es decir es el que destoca más de vuelta a almacén comparado con los demás respecto a las unidades enviadas, esto es debido a dos tiendas españolas ⁶ por no contar con zona de exposición del área TRF y consolidan el producto a otras tiendas mediante destocaje decidido por negocio. En comparación con la anterior campaña de Invierno 2022, se destoca más unidades para los siguientes compradores: CIRCULAR, PUNTO y TRF.

⁶Juan Flórez (A Coruña) y Constitución (Valladolid)

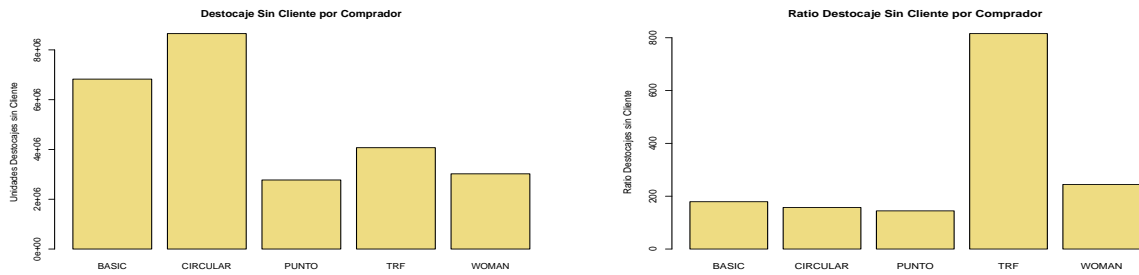


Figura 1.6: Representación de los Destocajes Sin Cliente y el Ratio de Destocajes Sin Cliente en función de los compradores.

Si se analiza el Destocaje Sin Cliente en los diferentes mercados que conforman el Core Business, Figura 1.7. España sigue siendo el país que más destoca seguido de Estados Unidos, Francia y Italia y Reino Unido, puesto que son los países con mayor numero unidades enviadas a las tiendas.

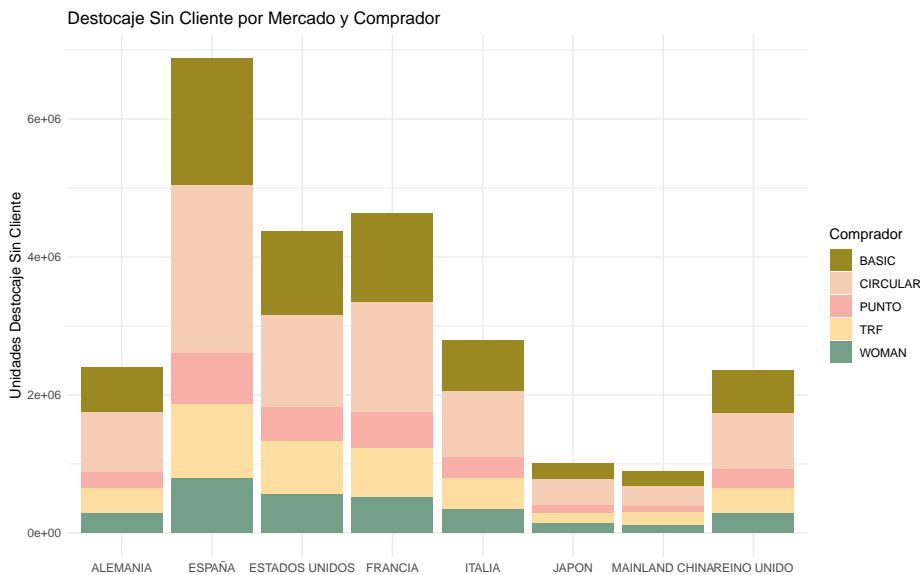


Figura 1.7: Representación del Ratio de Destocajes Sin Cliente por mercado y comprador.

Mediante este análisis descriptivo se deduce que el destocaje sin cliente es proporcional a las unidades enviadas por mercado y comprador. Sin embargo, el Ratio de devoluciones en función de las unidades vendidas varía en función de la cultura de los mercados, donde Italia y Japón son los países con menos ratio de devolución.

Cabe destacar que hay tiendas que debido a la tipología de cliente reciben un mayor volumen de devoluciones de prendas que no están en exposición, lo que origina un mayor destocaje de cliente.

Por todos estos motivos se realiza a continuación, en el [Capítulo 2](#), una agrupación de tiendas del los mercados del Core Business para disponer de una herramienta que permita visualizar este tipo de comportamiento. Mientras que en el [Capítulo 3](#) se desarrolla un modelo de optimización para minimizar los destocajes mediante las unidades enviadas, sin perjudicar el aumento de la venta.

El análisis descriptivo realizado proporciona una comprensión detallada de cómo las diferentes variables: unidades enviadas, vendidas, devueltas, y destocajes tanto de cliente como sin cliente, interactúan entre sí y varían entre los distintos mercados y diferentes compradores. Se identifican una serie de tendencias y patrones significativos que destacan las diferencias en el comportamiento del cliente y las operaciones de las tiendas en los distintos mercados del Core Business. Estos hallazgos subrayan la necesidad de una estrategia más personalizada y adaptada para abordar los distintos comportamientos de cada mercado y comprador, especialmente en lo que respecta a la gestión de inventario y la optimización de las operaciones en tienda.

El siguiente paso será realizar una agrupación detallada de las tiendas dentro de los mercados del Core Business. Esta clasificación ayudará a entender mejor los patrones de los compradores, y será tratada en profundidad en el [Capítulo 2](#). Posteriormente, en el [Capítulo 3](#), se desarrolla un modelo de optimización que se enfoca en minimizar los destocajes a través de una gestión más eficaz de las unidades enviadas, buscando así no solo reducir costes, sino también mejorar la satisfacción del cliente y aumentar las ventas sin incrementar las devoluciones a almacén.

Capítulo 2

Análisis de Comportamiento de los Compradores

2.1. Introducción métodos clustering

La agrupación mediante clústeres es una técnica fundamental en el análisis multivariante no supervisado, que clasifica datos en grupos homogéneos basándose en características compartidas, (veáse [Cuadras \(1996\)](#)). Esta metodología se divide en dos enfoques principales: los métodos jerárquicos y de particionamiento. Los métodos jerárquicos construyen una jerarquía de clústeres y son óptimos cuando la estructura de los datos y el número de clústeres son desconocidos, proporcionando una visualización clara mediante dendrogramas. Por otro lado, los métodos de particionamiento, como el k-medias, son preferidos para conjuntos de datos voluminosos donde se tiene una estimación previa del número de grupos.

Estas técnicas de agrupamiento encuentran aplicación en una variedad de campos. En el caso de la industria textil esta agrupación puede utilizarse para segmentar a los clientes según sus preferencias de moda, optimizar la logística de distribución de los diferentes productos o incluso una recomendación de producto demostrando su versatilidad y capacidad para revelar patrones ocultos en los datos. En el caso de estudio se usará esta clasificación en función de los compradores (Departamentos de Compra) de Zara para analizar su comportamiento en cada tienda en función de sus ventas, devoluciones y destocajes a almacén producidos por las devoluciones de cliente como los destocajes promovidos por las decisiones de negocio tomadas desde central. Además se proporciona el código específico necesario para su implementación práctica, facilitando así la replicación del análisis y adaptación del estudio a diferentes departamentos de compras y secciones.

Permitiendo una agrupación mediante las variables de interés: unidades enviadas, unidades vendidas, ratio devolución, ratio destocaje cliente y ratio destocajes sin cliente. Permitiendo un análisis de combinaciones de tiendas y principales compradores (departamentos de compra: BASIC, CIRCULAR, PUNTO, TRF y WOMAN) de sección señora de Zara. Para los distintos mercados principales, llamados Core Business.

Cada observación se trata como un vector x en un espacio n -dimensional, donde n representa el número de variables en cada vector. En el caso de estudio disponemos de las 5 variables descritas anteriormente.

2.2. Tipos de clúster

2.2.1. k -means

La agrupación k -means, o k -medias, es un método de particionamiento, introducido por [McQueen \(1967\)](#), el cual se basa en dividir los datos en k grupos C_k , en base a los valores de las variables observadas en cada individuo. Agrupa minimizando la suma de las diferencias la cuadrado de las distancias euclídeas entre cada elemento x , de longitud n que son el número de variables y el centroide del clúster u_k , lo que dependerán de la proximidad del espacio n -dimensional, siendo n el número de variables.

$$\text{mín} \sum_{k=1}^K \sum_{x \in C_k} |x - \mu_k|^2 \quad (2.1)$$

Siguiendo el siguiente criterio de optimalidad que implica que el algoritmo busca agrupar los puntos de datos de tal manera que los puntos dentro de cada clúster estén lo más cercanos posible a un punto central común, el centroide, (véase [Kassambara \(2017\)](#)). Esto significa que se busca una configuración de clústeres donde los puntos de datos dentro de cada clúster sean similares entre sí, minimizando así la varianza interna del clúster. La optimización se realiza iterativamente ajustando los centroides y reasignando los puntos a los clústeres hasta que se alcanza una convergencia, definida por un cambio insignificante en los centroides o por la estabilización de las asignaciones de clústeres.

Algoritmo k -means

1. Inicialización: determinar el número K de clúster iniciales, donde se selecciona aleatoriamente k objetos del conjunto de datos como centros iniciales de esos clústeres.
2. Asignación: de cada observación al centroide más cercano, basándose en la distancia $d(x, \mu_k i)$.

$$d(x, \mu_k i) = \sqrt{\sum_{i=1}^n (x_i - \mu_k i)^2} \quad (2.2)$$

3. Actualización: del centroide de cada clúster calculando los nuevos valores medios de todos los puntos de datos en el clúster. El centroide de un clúster k es un vector de longitud n que contiene las medias de todas las variables para las observaciones en dicho clúster.
4. Convergencia: minimiza iterativamente la suma total de cuadrados dentro de los clústeres, repitiendo los pasos de asignación y actualización hasta que las asignaciones de clústeres dejen de cambiar o se alcance el número máximo de iteraciones.

2.2.2. PAM

La agrupación PAM (Partitioning Around Medoids) es otro método de particionamiento introducido por Kaufman y Rousseeuw en 1987, el cual basa su búsqueda en k clústeres utilizando medoides en vez de centroides para representar las agrupaciones. El número k clúster es fijado de antemano como en el agrupamiento k -medias, sin embargo este método ofrece una alternativa robusta.

En el algoritmo de PAM, el criterio de optimalidad se centra en minimizar la suma de las distancias entre cada punto y el medoide de su clúster asignado. A diferencia del k -means, que minimiza las distancias cuadradas al centroide, PAM minimiza la suma de las distancias reales (no al cuadrado) hasta el medoide, proporcionando robustez frente a valores atípicos y una medida más directa de la dispersión dentro del clúster.

Algoritmo PAM

1. Inicialización: determinar el número K de clúster iniciales, donde se selecciona aleatoriamente k objetos del conjunto de datos como medoides ¹, punto medio del grupo.
2. Asignación: de cada observación al medoide m más cercano, basándose en minimizar la distancia $d(x, m_k i)$.

$$\sum_{i=1}^n \min_k d(x_i, m_k i) \quad (2.3)$$

3. Actualización: esta etapa también es conocida como SWAP donde se prueban diferentes candidatos para ser medoides intercambiándolos con los medoides actuales si conducen a una disminución en la suma total de las distancias dentro del clúster.
4. Convergencia: Cuando ha minimizado la distancia después de repetir los pasos 2 y 3.

En el algoritmo de Partitioning Around Medoids (PAM), el criterio de optimalidad se centra en minimizar la suma de las distancias entre cada punto y el medoide de su clúster asignado. A diferencia del k -means, que minimiza las distancias cuadradas al centroide, PAM minimiza la suma de las distancias reales (no cuadradas) hasta el medoide, proporcionando robustez frente a valores atípicos y una medida más directa de la dispersión dentro del clúster.

2.2.3. Jerárquico

El clustering jerárquico es un método de análisis de clúster que busca construir una jerarquía de clústeres. A diferencia de los métodos de particionamiento, que requieren la especificación del número de clústeres de antemano, el clustering jerárquico crea un dendrograma que permite visualizar y decidir el número de clústeres mediante la interpretación de la altura de los cortes. Existen dos tipos de clústeres jerárquicos:

- Aglomerativo: los métodos de clustering jerárquico aglomerativo comienzan con la partición más fina posible (un observación por clúster) y progresivamente combinan clústeres.
- Disociativo: comienza de manera descendente, siendo un proceso inverso al aglomerativo, donde se comienza con un conglomerado con todos los casos y se realizan sucesivas divisiones hasta formar grupos más pequeños.

En el caso de estudio se procede con el método aglomerativo. Esta elección se fundamenta en la flexibilidad en cuanto a la elección de la métrica de enlace, lo que permite adaptar el análisis a las peculiaridades específicas del conjunto de datos, como la sensibilidad a los valores atípicos y la forma de los clústeres. Esta metodología es especialmente adecuada para conjuntos de datos donde las relaciones entre observaciones pueden ser complejas o donde se desconoce la cantidad óptima de clústeres a priori que es el caso de los datos de los que disponemos, proporcionando un marco intuitivo y gradual para la exploración de los datos y la validación de supuestos subyacentes en el agrupamiento.

El criterio de optimalidad en la agrupación jerárquica aglomerativa se basa en la minimización o maximización de una función de distancia entre clústeres, elegida según la métrica de enlace. La métrica de enlace determina cómo se calcula esta distancia entre clústeres formados por vectores de n variables. Dependiendo de la métrica seleccionada: enlace mínimo, máximo o promedio, la distancia refleja diferentes aspectos de la relación entre los clústeres en el espacio n -dimensional.

¹El medoide es el punto de datos dentro de un clúster que minimiza la suma de las distancias a todos los demás puntos en el clúster, ofreciendo una representación robusta y resistente a valores atípicos.

- **Single (Mínimo Enlace):** este método une los clústeres basándose en la distancia más corta entre cualquier par de puntos en los clústeres que se están considerando para unir. Es sensible a los valores atípicos y puede producir clústeres en forma de cadena.
- **Complete (Máximo Enlace):** une los clústeres según la distancia más larga entre cualquier par de puntos en los clústeres que se están considerando. Produce clústeres más compactos y bien separados, siendo menos susceptible a los valores atípicos en comparación con el enlace simple.
- **Average (Enlace Promedio):** calcula el promedio de las distancias entre todos los pares de puntos en los dos clústeres. Este enlace es menos susceptible a los valores atípicos en comparación con el enlace simple y tiende a crear clústeres más balanceados.
- **Ward (Enlace de Ward):** Este método es algo diferente; no se basa directamente en las distancias entre los puntos de datos. En lugar de ello, el método de Ward busca minimizar el aumento en la suma total del cuadrado de las distancias dentro de todos los clústeres. Esto significa que se prefiere unir los clústeres que incrementarán lo menos posible la varianza total dentro del clúster resultante. Es particularmente útil para identificar clústeres de aproximadamente el mismo tamaño.

Para el análisis de agrupación de compradores, se ha seleccionado la distancia de Ward como la métrica de enlace debido a su coeficiente aglomerativo superiormente alto. Este resultado sugiere que el método de Ward genera una estructura de clústeres notablemente homogénea, en la que los clústeres internos son consistentemente cohesivos, y al mismo tiempo, están claramente diferenciados de otros clústeres. Esta característica es esencial para asegurar que los segmentos de compradores identificados sean no solo internamente similares, sino también distintivamente separados unos de otros, facilitando así análisis subsecuentes y la toma de decisiones basada en la agrupación ² resultante.

Algoritmo Jerárquico

1. Inicialización: se inicia con cada observación como un clúster individual, comenzando con n clústeres cada uno conteniendo una sola observación. En este contexto, cada observación es un vector x con 5 variables.
2. Asignación: en cada etapa, une los dos clusters con la distancia más corta, $d(C_k, C_l)$.
3. Actualización: recalcula la matriz de distancias entre el nuevo clúster formado y los demás existentes.
4. Convergencia: cuando todas las observaciones n están agrupadas en un sólo clúster.

Una manera gráfica y para presentar los resultados del método de jerárquico aglomerativo es mediante un dendrograma. Estos muestran la disposición de los clusters formados en cada etapa. La altura de cada nodo en el gráfico es proporcional al valor de la distancia intergrupala entre sus dos "hijas". Alternativamente se puede cortar el árbol según el número de clusters requeridos especificando la altura a la que se debe cortar el árbol.

2.3. Evaluación y validación de los métodos de clúster

Para asegurar una comparación rigurosa y exhaustiva de los algoritmos de clúster empleados en este estudio (k -means, jerárquico y PAM), se adopta un enfoque de validación interna centrado en la evaluación de la compactación, conectividad, y separación de las particiones de clúster. Este enfoque (véase Brock (2008)) se complementa con un análisis de estabilidad, permitiendo una valoración detallada de la calidad y robustez de las agrupaciones obtenidas.

²<https://rich-d-wilkinson.github.io/MATH3030/9.3-hierarchical-clustering-methods.html>

2.3.1. Validación interna

- **Conectividad:** refleja la proximidad entre los puntos y sus vecinos más cercanos. Se define de tal manera que $x_{i,nn_i(j)}$ es cero si la observación i y su j -ésimo vecino están en el mismo clúster, y $\frac{1}{j}$ en caso contrario. La conectividad para el conjunto $C = \{C_1, \dots, C_K\}$ de las N observaciones en K clústeres se calcula como:

$$\text{Conectividad}(C) = \sum_{i=1}^N \sum_{j=1}^L x_{i,nn_i(j)}$$

Donde L es el número de vecinos que contribuyen a la conectividad siendo es el número de datos que hay en el mismo clúster que i . Esta medida debe ser minimizada para lograr una agrupación efectiva.

- **Índice de Silueta:** proporciona una medida compuesta que evalúa tanto la cohesión dentro de los clústeres como la separación entre ellos, facilitando la evaluación de la pertinencia de la asignación de puntos a los clústeres. Para la observación i , el índice de silueta se define como:

$$S(i) = \frac{b_i - a_i}{\max(b_i, a_i)}$$

donde a_i es la distancia media entre i y todas las demás observaciones en el mismo clúster, y b_i es la distancia media entre i y todas las observaciones en el clúster vecino más cercano. Estas distancias se calculan de la siguiente manera:

$$a_i = \frac{1}{n(C(i))} \sum_{j \in C(i)} d(i, j)$$

$$b_i = \min_{C_k \in C \setminus C(i)} \left(\frac{\sum_{j \in C_k} d(i, j)}{n(C_k)} \right)$$

donde $C(i)$ es el clúster que contiene la observación i , $d(i, j)$ representa la distancia entre las observaciones i y j , y $n(C)$ es el número de elementos que forman el clúster C . La amplitud de silueta, que varía entre $[-1, 1]$, debería maximizarse para reflejar una asignación de clúster óptima.

- **Índice de Dunn:** se enfoca en maximizar la relación entre la mínima distancia interclúster y la máxima distancia intraclúster, buscando clústeres que sean compactos y bien separados. Siendo el cociente entre la menor distancia entre observaciones que no están en el mismo clúster y la mayor distancia dentro de un clúster. Se calcula como:

$$D(C) = \frac{\min_{C_k, C_l \in C, C_k \neq C_l} (\min_{i \in C_k, j \in C_l} d(i, j))}{\max_{C_m \in C} \text{diam}(C_m)}$$

donde $\text{diam}(C_m)$ es la máxima distancia entre observaciones en el clúster C_m . El índice de Dunn tiene un valor entre cero e infinito, y debe ser maximizado.

2.3.2. Estabilidad

En las medidas de estabilidad se comparan los resultados del clustering ejecutado con todos los datos con el clustering basado en la eliminación de cada variable, una a la vez (Datta y Datta 2003; Yeung et al. 2001). Para evaluar la estabilidad de las agrupaciones generadas, se consideran las siguientes medidas de estabilidad: proporción media de no solapamiento (APN: Average proportion of non-overlap), la distancia media (AD: Average Distance), la distancia media entre medias (ADM: Average Distance between Means), y la figura de mérito (FOM: Figure of merit). En todos los casos, el promedio se

toma sobre $n-1$ ³, y todas las medidas deben minimizarse. Donde N el número total de observaciones en el conjunto de datos y M el número total de variables numéricas.

- **Proporción media de no solapamiento (APN)**: evalúa la proporción media de no solapamiento entre clústeres en diferentes subconjuntos de datos, indicando la consistencia de las agrupaciones. Sea $C_{i,0}$ el clúster que contiene la observación i usando el clustering original (basado en todos los datos disponibles), y $C_{i,\ell}$ el clúster que contiene la observación i donde el clustering se basa en el conjunto de datos con la columna ℓ eliminada. Entonces, la medida APN se define como:

$$APN(C) = \frac{1}{MN} \sum_{i=1}^N \sum_{\ell=1}^M \left(1 - \frac{n(C_{i,\ell} \cap C_{i,0})}{n(C_{i,0})} \right) \quad (2.4)$$

El APN está en el intervalo $[0, 1]$, con valores cercanos a cero correspondientes a resultados de clustering altamente consistentes.

- **Distancia Media (AD) y Distancia Media entre Medias (ADM)** : estas métricas miden la distancia promedio dentro de los clústeres y la distancia promedio entre los centroides de clúster, proporcionando una visión de la cohesión interna y la diferenciación entre clústeres.

La medida AD calcula la distancia promedio entre las observaciones colocadas en el mismo clúster, basado tanto en los datos completos como en los datos con una variable eliminada. Se define como:

$$AD(C) = \frac{1}{MN} \sum_{i=1}^N \sum_{\ell=1}^M \left(\frac{1}{n(C_{i,0})n(C_{i,\ell})} \sum_{i \in C_{i,0}, j \in C_{i,\ell}} d(i, j) \right) \quad (2.5)$$

donde $C_{i,0}$ representa el clúster que contiene la observación i usando el clustering original y $C_{i,\ell}$ es el clúster que contiene la observación i donde el clustering se basa en el conjunto de datos con la variable ℓ eliminada. AD tiene valores entre cero e infinito, y se prefieren valores menores.

La medida ADM calcula la distancia promedio entre los centros de clúster para observaciones colocadas en el mismo clúster, basado tanto en los datos completos como en los datos con una variable eliminada. Se define como:

$$ADM(C) = \frac{1}{MN} \sum_{i=1}^N \sum_{\ell=1}^M d(\bar{x}_{C_{i,\ell}}, \bar{x}_{C_{i,0}}) \quad (2.6)$$

donde $\bar{x}_{C_{i,0}}$ es la media de las observaciones en el clúster que contiene la observación i , cuando el clustering se basa en los datos completos, y $\bar{x}_{C_{i,\ell}}$ es definido de manera similar para los datos con la variable ℓ eliminada. Actualmente, ADM sólo utiliza la distancia euclidiana. ADM también tiene valores entre cero e infinito, y se prefieren valores menores.

- **Figura de mérito (FOM)** : evalúa la coherencia de la compactación mediante la varianza promedio de las distancias de los puntos de un clúster a su centroide a través de diferentes particiones de datos. El FOM mide la varianza intraclúster promedio de las observaciones en la variable eliminada, donde el clustering se basa en las muestras restantes (no eliminadas). Esto estima el error medio utilizando predicciones basadas en los promedios del clúster. Para una columna específica eliminada ℓ , el FOM se define como:

$$FOM(\ell, C) = \sqrt{\frac{1}{N} \sum_{k=1}^K \sum_{i \in C_k(\ell)} d(x_{i,\ell}, \bar{x}_{C_k(\ell)})} \quad (2.7)$$

donde $x_{i,\ell}$ es el valor de la i -ésima observación en la variable ℓ , y $\bar{x}_{C_k(\ell)}$ es el promedio del clúster $C_k(\ell)$. El FOM se multiplica por un factor de ajuste $\sqrt{\frac{N}{N-K}}$, para aliviar la tendencia

³Donde n representa el número de variables en cada vector.

a disminuir a medida que aumenta el número de clústeres. La puntuación final se promedia sobre todas las variables eliminadas y tiene un valor entre cero e infinito, siendo los valores más pequeños indicativos de un mejor rendimiento.

La aplicación de estas medidas en un marco comparativo se lleva a cabo utilizando el software estadístico R [2] mediante la librería `clValid` [6], facilitando un análisis detallado y reproducible de los resultados.

2.4. Objetivo

Debido a la necesidad de una clasificación por el tipo de comportamiento en cada comprador se procede a realizar una agrupación vía clústeres con el objetivo de identificar las necesidades de cada comprador que predominan en cada tienda de Zara Señora. Este análisis permite segmentar los tipos de compradores en función de las variables de interés y poder clasificar en la tienda a cada comprador. Lo que resulta interesante para identificar que tiendas vendrán marcadas por el destocaje de cliente, los otros tipos de destocaje, las ventas o las devoluciones y así poder gestionar de una manera eficaz la operativa en caja de esas tiendas.

2.5. Procesamiento y validación

El procesamiento y análisis de los datos se realiza utilizando RStudio versión 4.2.1, con la base de datos de unidades acumuladas de la campaña de Invierno 2023, sin contar con las rebajas. Este análisis comprende varias etapas clave, comenzando con la selección del método de clúster más adecuado para el conjunto de datos. Para esta tarea, se utiliza la librería `clValid` en R, que facilita la validación y comparación de diferentes métodos de agrupación mediante métricas de validación internas.

Una vez identificado el método óptimo, se procede con el análisis detallado de los resultados del agrupamiento. La fase final del estudio consiste en la preparación y suministro de una base de datos en formato Excel, la cual contiene los resultados específicos de la asignación de los distintos clústeres a cada tienda, clasificados por comprador y mercado. Este enfoque metódico asegura no solo la precisión en la segmentación de los datos, sino también una adecuada interpretación y aplicación de los resultados en estrategias comerciales y de marketing.

El análisis comienza con la estandarización de las variables de interés para garantizar la uniformidad de la escala y comparabilidad entre ellas. Las variables incluidas en el análisis son: Unidades enviadas, Unidades vendidas, Ratio devolución, Ratio destocaje cliente y Ratio destocaje sin cliente. Este proceso de normalización es esencial para la aplicación efectiva de las técnicas de agrupación.

Posteriormente, se lleva a cabo una evaluación exhaustiva de los algoritmos de clúster k -means, jerárquico y PAM utilizando la librería `clValid` en RStudio. El análisis se enfoca en la validación interna de estos algoritmos para determinar su estabilidad y efectividad en la formación de clústeres. Basándonos en los resultados de esta validación, se elige el método de agrupación jerárquica para analizar el mercado Core Business. Las tablas obtenidas se agrupan en el [Anexo B](#) donde los resultados de validación interna se muestran en la Tabla B.1 y los datos de estabilidad en la Tabla B.2.

La elección de dividir el conjunto de datos en cuatro clústeres se justifica observando la altura significativa del dendrograma que se muestra en la Figura 2.1 a la que ocurren las últimas fusiones importantes antes de alcanzar un nivel de disimilitud demasiado elevado, lo cual podría comprometer la cohesión interna de los clústeres. Un corte horizontal en la altura de aproximadamente 500 unidades en este dendrograma proporciona una clara demarcación entre clústeres relativamente homogéneos. Este nivel se selecciona basándose en el incremento dramático en la altura de fusión que ocurre después de este punto, indicando que cualquier fusión adicional resultaría en clústeres menos cohesivos.

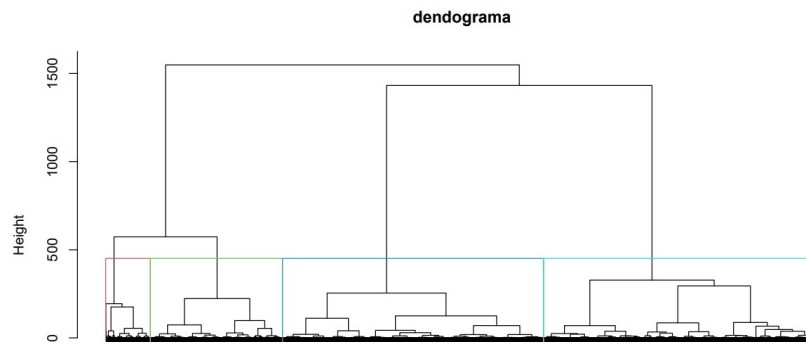


Figura 2.1: Dendrograma representando la jerarquía de agrupación en 4 grupos.

Por ello, utilizar cuatro clústeres permite una segmentación más detallada, lo que es crucial para desarrollar estrategias de marketing y operativas diferenciadas. Cada clúster representa un segmento de mercado con características y necesidades únicas, facilitando la implementación de tácticas específicas diseñadas para abordar y capitalizar esas diferencias.

Los valores de silueta proporcionan una métrica cuantitativa para evaluar la calidad de los clústeres en términos de cohesión interna y separación de otros clústeres. Tal como se observa en la Tabla 2.1 el clúster 1 tiene un valor relativamente bajo comparado con el resto, lo que indica que los puntos dentro de este clúster están bastante cerca de los puntos de otros clústeres, sugiriendo que hay cierta superposición con otros clústeres. El clúster 2 tiene un valor de silueta moderado, sugiriendo que el clúster está razonablemente bien definido. El Clúster 3 tiene un valor similar al clúster 2, mostrando una separación aceptable. Y en el clúster 4 es claramente el más definido y compacto de los cuatro, indicando que los puntos dentro de este clúster están bien separados de los puntos en otros clústeres y bien agrupados entre sí.

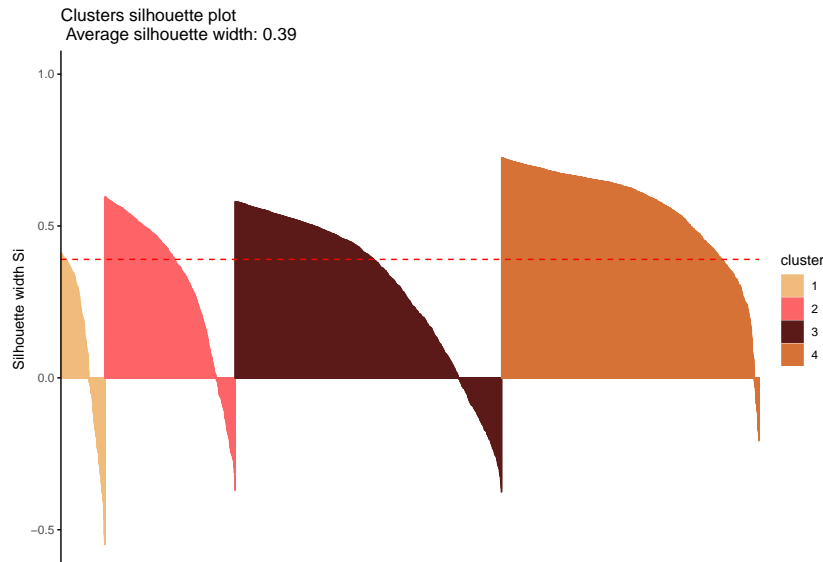


Figura 2.2: Diagrama de silueta de los 4 clústeres identificados, mostrando el coeficiente de silueta para cada punto dentro de su clúster correspondiente.

| Clúster | Tamaño del Clúster | Anchura Media de Silueta |
|---------|--------------------|--------------------------|
| 1 | 248 | 0.10 |
| 2 | 733 | 0.33 |
| 3 | 1499 | 0.31 |
| 4 | 1449 | 0.56 |

Tabla 2.1: Medidas de tamaño de clúster y anchura media de silueta.

Una vez escogido el método jerárquico y los 4 grupos que formaran las agrupaciones se realiza un estudio de bootstrap y subset para conocer la estabilidad de cada uno de estos grupos haciendo referencia al estudio de Hennig (2007), para testear la robustez de los clústeres. Para la validación de la estabilidad de las agrupaciones obtenidas se recurre al índice de Jaccard (J) que por definición para dos conjuntos de clústeres C y D , el cual se representa mediante: $\gamma(C, D) = \frac{|C \cap D|}{|C \cup D|}$ donde $C, D \subseteq X^n$. El coeficiente de Jaccard, (véase Jaccard, 1901) es la proporción de puntos que pertenecen a ambos conjuntos de todos los puntos involucrados en al menos uno de los conjuntos, y por resultando interpretable de una manera directa. Este coeficiente es independiente del número de puntos que no pertenecen a ninguno de los dos conjuntos $1 - \gamma$.

Se obtiene el coeficiente Jaccard y la inestabilidad de los clústeres obtenidos mediante bootstrap naive y subsetting:

- Bootstrap naive: con el objetivo de evaluar la estabilidad de los clusters formados al recrear múltiples conjuntos de datos simulados mediante el remuestreo con reemplazamiento del conjunto original. Evaluándose la estabilidad del clúster de la siguiente manera: suponiendo que existe una distribución de mezcla subyacente $P = \sum_{i=1}^s \varepsilon_i P_i$ donde $P_i, i = 1, \dots, s$, son las distribuciones que generan s los clusters verdaderos, y π_i es la probabilidad de que un punto de P_i sea extraído. Para un conjunto de datos dado con n puntos, el clustering verdadero consistiría entonces en s clusters cada uno de los cuales contiene exactamente los puntos generados por $P_i, i = 1, \dots, s$.

Cuando un conjunto de datos generado desde P es agrupado, los clusters encontrados difieren de los clusters verdaderos, porque el método de clustering introduce un cierto sesgo y variación. Esto puede depender del clúster P_i , por ejemplo, si dos clusters diferentes están débilmente separados o si P_i se desvía fuertemente del modelo de clúster asumido por el método de clustering. Sesgo y variación pueden ser expresados por el coeficiente de Jaccard máximo entre el conjunto de todos los puntos generados por P_i y el clúster más similar en el clustering realmente obtenido.

El bootstrap se realiza con el fin de representar el sesgo y la varianza, porque en la realidad no se conoce ninguna distribución subyacente verdadera ni ningún clustering verdadero. La distribución empírica del conjunto de datos observado se toma entonces para simular P . Los puntos pueden ser extraídos del conjunto de datos y los clusters originalmente encontrados pueden ser tratados como los "verdaderos". El coeficiente de Jaccard máximo medio puede interpretarse como indicativo de la estabilidad de los clusters originales. El cual variará de 0 a 1, donde 1 representará la máxima estabilidad.

Dado un número B de replicas bootstrap y un clúster C del clustering original $E_n(x)$, donde $i = 1, \dots, B$, el algoritmo es el siguiente:

1. Se extrae una muestra bootstrap x_n^i de n puntos con reemplazo del conjunto de datos original x_n .
2. Se ejecuta el clúster $E_n(x_n^i)$.
3. Sea $x_*^i = x_n \cap x_n^i$ los puntos del conjunto de datos original que también están en la muestra de bootstrap. Sea $C_*^i = C \cap x_*^i$, $\Delta = E_n(x_n^i) \cap x_*^i$.
4. Si $C_*^i \neq \emptyset$, la similitud de Jaccard es máxima entre el clúster inducido C_*^i y el nuevo clustering inducido Δ sobre x_*^i : $C_i = \max_{D \in \Delta} \gamma(C_*^i, D)$ (lo cual, D maximiza a $\gamma(C_*^i, D)$); de lo contrario, $\gamma_{C,i} = 0$.

Esto genera una secuencia $C_i, i = 1, \dots, B$. Que obteniendo la media

$$\bar{\gamma}_C = \frac{1}{B^*} \sum_{i=1}^B C_i$$

como medida de estabilidad (siendo B^* con el número de replicas de bootstrap para las cuales $C_*^i \neq \emptyset$ y se usa aquí porque en todos los otros casos $C_i = 0$).

- **Subsetting:** en este caso se extraer un subconjunto de x_n sin reemplazo evitando puntos múltiples y reduciendo el tiempo de computación, lo cual podría ser un problema con conjuntos de datos grandes. Sin embargo con $B=50$ tal como recomienda [Hennig \(2007\)](#) el con el conjunto de datos a analizar apenas se nota diferencia en el tiempo de ejecución con respecto al método bootstrap. De tal manera que siendo x_*^i el subconjunto extraído de x_n . Se calcula la media del número de réplicas para obtener el parámetro Jaccard que indica la estabilidad de las agrupaciones.

Este método requiere la elección del tamaño $m < n$ del subconjunto. Si m es demasiado grande, no generará suficiente variación para la obtención de un valor fiable. Si m es demasiado pequeño, se puede esperar que los resultados del clustering sean mucho peores que los obtenidos del conjunto de datos original. Con $m = \lfloor n/2 \rfloor$, donde $\lfloor x \rfloor$ denota la parte entera.

Los datos obtenidos del estudio bootstrap se muestran a la izquierda de la Tabla 2.2, donde evaluando la estabilidad de los clústeres, los clústeres 3 y 4 tienen alta consistencia y robustez, con valores de Jaccard promedio altos mostrando mínima inestabilidad. A pesar de que los clústeres 1 y 2 presentan una estabilidad menor comparados con los clústeres 3 y 4, sus valores de Jaccard caen cerca del valor crítico de 0.5 según la referencia [Hennig \(2007\)](#) por lo que pueden considerarse estables.

Tabla 2.2: Comparación de Medidas de Jaccard promedio e inestabilidad por clúster.

| Bootstrap | | | Subsetting | | |
|-----------|------------|---------------|------------|------------|---------------|
| Clúster | AvgJaccard | Inestabilidad | Clúster | AvgJaccard | Inestabilidad |
| 1 | 0.63 | 0.32 | 1 | 0.64 | 0.32 |
| 2 | 0.50 | 0.48 | 2 | 0.54 | 0.42 |
| 3 | 0.73 | 0.02 | 3 | 0.73 | 0.04 |
| 4 | 0.79 | 0.00 | 4 | 0.78 | 0.02 |

Mediante la Tabla 2.2 derecha, se obtienen los resultados obtenidos al realizar subsetting para la validación de los clústeres, similares a los obtenidos mediante bootstrap. Los clústeres 3 y 4 resaltan

por su alta estabilidad y coherencia, mientras que los clústeres 1 y 2, aunque en el límite del valor crítico, muestran una consistencia razonable.

2.6. Resultados

A continuación se comentan los resultados obtenidos de la agrupación jerárquica en 4 clústeres:

Porcentaje de tiendas por clúster

El histograma que muestra la Figura 2.3 representa la distribución porcentual de tiendas en los cuatro clústeres definidos a través del análisis jerárquico. Se observa una distribución desigual de las tiendas entre las agrupaciones, lo cual es indicativo de variabilidad en las características de los departamentos de compra predominantes y en la demografía de los clientes que frecuentan estas tiendas. Donde las agrupaciones con mayor representación contienen una mayor proporción de tiendas que podrían ser considerados como mercados objetivo para iniciativas distintas, dada su significativa representación. Sin embargo, los clústeres con menor representación representan segmentos de mercado que, ofreciendo oportunidades para estrategias diferenciadas y enfocadas se podrían explotar características únicas o demandas específicas.

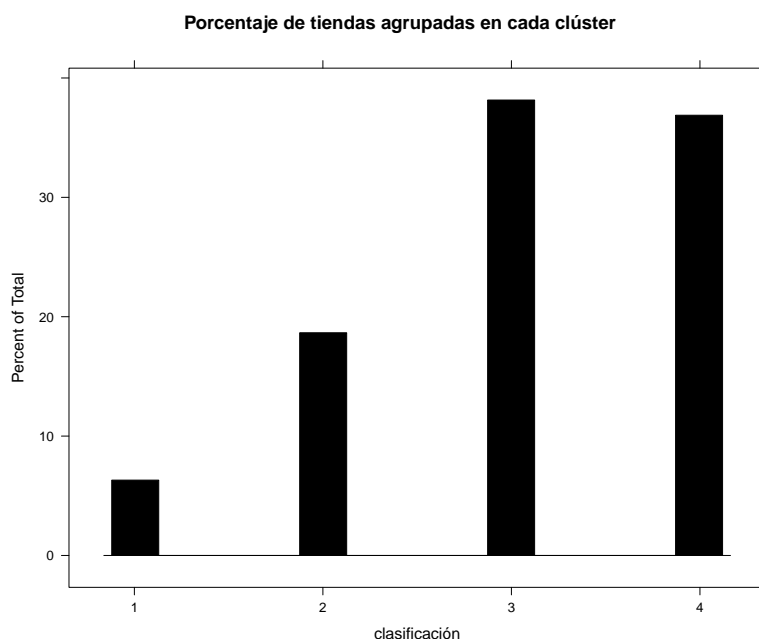


Figura 2.3: Histograma de porcentaje de tiendas clasificadas en cada clúster.

Análisis de los clústeres según las densidades de variables

La Figura 2.4 muestra los gráficos de densidad de las cinco variables a analizar: unidades enviadas, unidades vendidas, ratio de devolución, ratio destocaje sin cliente, y ratio destocaje cliente. Estos gráficos son esenciales para entender las características operativas y de mercado que definen cada uno de los cuatro clústeres identificados mediante el análisis jerárquico de agrupación.

- **Clúster 1 y 2:** muestran densidades significativas para las variables de unidades enviadas y unidades vendidas. Esto indica que estos clústeres podrían estar compuestos por compradores

con alto volumen de ventas y envíos, lo cual sugiere una alta actividad comercial y posiblemente una buena gestión de inventario o popularidad de los artículos.

- **Clúster 3:** destaca en la variable ratio de devolución. Esta variable sugiere que los compradores en este clúster podrían estar enfrentando desafíos con la satisfacción del cliente o la adecuación de los artículos, lo que resulta en una mayor tasa de devoluciones. Esto podría requerir una revisión de los envíos por comprador o de las políticas de devolución para mejorar la satisfacción del cliente.
- **Clúster 4:** los gráficos indican un pico notable en las variables ratio destocaje sin cliente y Ratio destocaje cliente. Esto sugiere que este clúster podría estar caracterizado por una mayor gestión o movimiento de stock, ya sea debido al destocaje producido por los clientes debido a las devoluciones no reaprovechadas en la tienda o debido a los destocajes producidos por decisiones de la central. Los compradores en este clúster podrían estar más orientadas a estrategias de liquidación o manejo eficiente del inventario para evitar un exceso de stock.

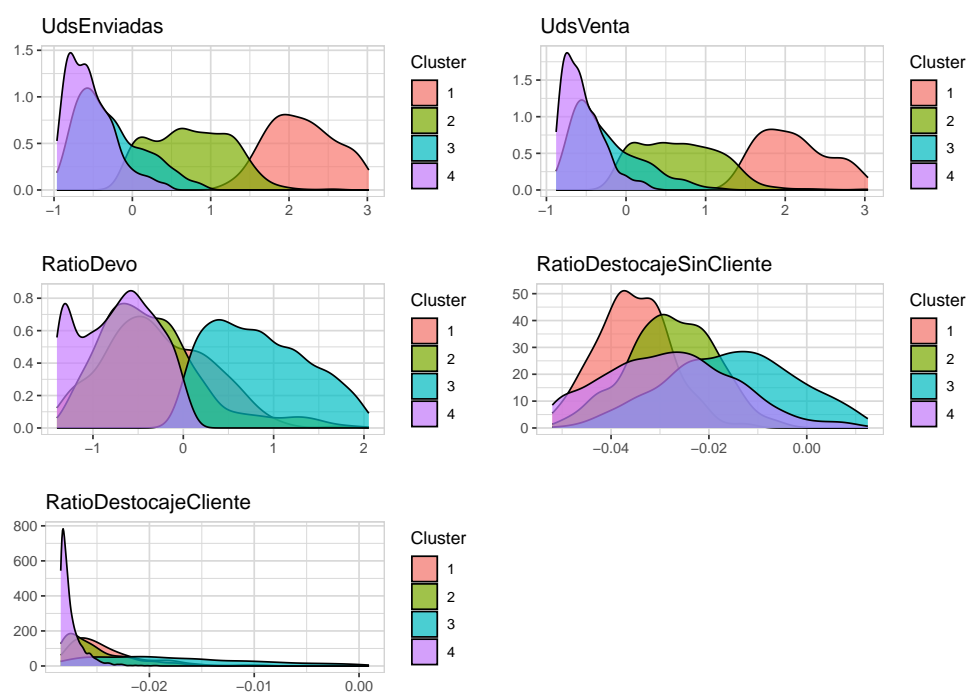


Figura 2.4: Distribución de los clúster en cada variable.

Clústeres por comprador

Analizando los resultados obtenidos en la clasificación por comprador en el mercado Core Business, tal como representa la Tabla 2.1 cada clúster tiene un comprador mayoritario y minoritario. En el clúster 1 se observa una alta representación del comprador CIRCULAR, sin embargo PUNTO, TRF y WOMAN apenas tienen representación en el clúster 1. En el clúster 2 BASIC y CIRCULAR son los compradores mayoritarios. Por lo que sabiendo que los clústeres 1 y 2 son definidos por las variables unidades enviadas y vendidas se concluye que los compradores BASIC y CIRCULAR son los compradores que envían y venden más artículos en la tienda. El clúster 3 es notable por los compradores TRF y WOMAN, que son los que marcan la tendencia de la tienda, sin embargo sugiere que son los

compradores que más devoluciones experimentan en comparación con los demás al ser los mayoritarios en este clúster. Mientras que el clúster 4 que viene definido por los ratios de destocaje el comprador mayoritario es PUNTO, se recuerda que es un comprador que complementa las 3 secciones.

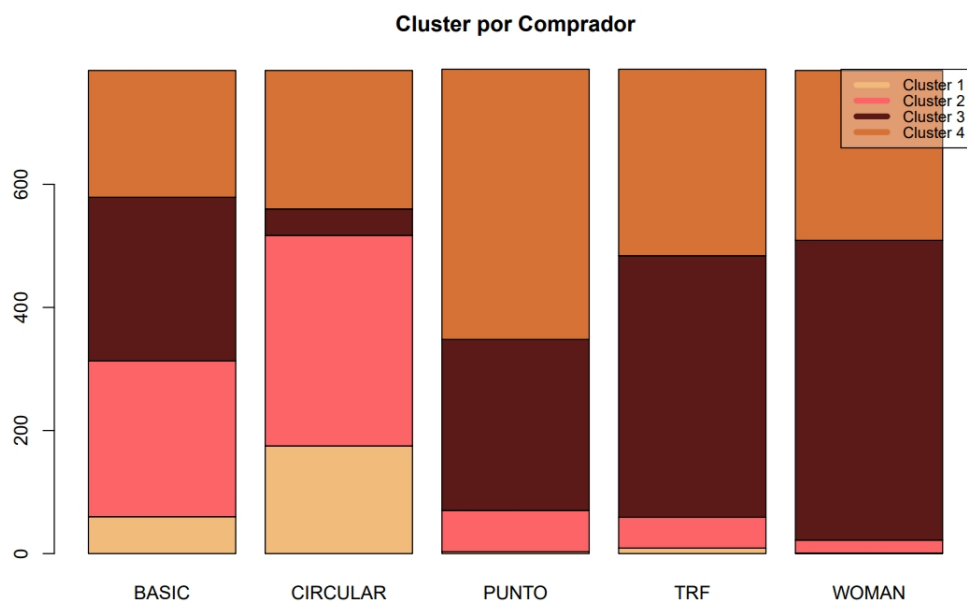


Figura 2.5: Representación de los 4 clústeres por Comprador.

Tabla 2.3: Distribución condicionada de los clústeres por compradores.

| Clúster | BASIC | CIRCULAR | PUNTO | TRF | WOMAN |
|---------|--------|----------|--------|--------|--------|
| 1 | 24.2 % | 70.6 % | 1.2 % | 3.6 % | 0.4 % |
| 2 | 34.5 % | 46.7 % | 9.1 % | 6.8 % | 2.9 % |
| 3 | 17.7 % | 2.9 % | 18.5 % | 28.4 % | 32.5 % |
| 4 | 14.2 % | 15.5 % | 30.3 % | 20.9 % | 19.1 % |

Clústeres por mercado

Se muestra en la Tabla 2.4 y se representa en la Figura 2.6 los resultados del análisis de los clústeres en los diferentes mercados que conforman el Core Business. Esta segmentación permite adaptar las estrategias de mercado de manera más efectiva a las características únicas de cada región.

- Clúster 1:** los mercados que tienen más representación son España, Francia y Reino Unido, se recuerda que este clúster viene marcado por las unidades enviadas y vendidas donde el comprador predominante es CIRCULAR, que fabrica para las 3 secciones principales y por tanto tiene representación en toda la tienda. Es destacable que ni Japón ni Mainland China tengan ninguna representación en el clúster 1.

- **Clúster 2:** en este clúster los mercados con mayor representación son Francia, Italia y España, cabe destacar que el clúster 2 como el 1 viene marcado por las unidades enviadas y vendidas, donde los compradores mayoritarios son BASIC y CIRCULAR.
- **Clúster 3:** el mercado mayoritario de este clúster es España, por lo que indicaría que es el país que más devoluciones recibe en función de las unidades vendidas, también es el mercado con mayor número de tiendas. Seguido de Estados Unidos que es el cuarto país con más tiendas del Core Business. Los mercados minoritarios en este clúster son Japón e Italia, lo que indica que debido a cuestiones culturales y políticas de devolución son mercados caracterizados con muy poco ratio de devolución. Señalando posibles áreas de mejora en satisfacción del cliente o ajustes en la política de devoluciones. Este clúster es definido por los compradores TRF y WOMAN.
- **Clúster 4:** la alta representación en Mainland China y Japón destaca, sugiriendo que las preferencias o necesidades en estos mercados asiáticos son significativamente diferentes de las de otros países. Esto podría estar relacionado con diferencias culturales, económicas o de consumo que influyen la popularidad de ciertos productos, provocando ratios de destocaje más elevados.

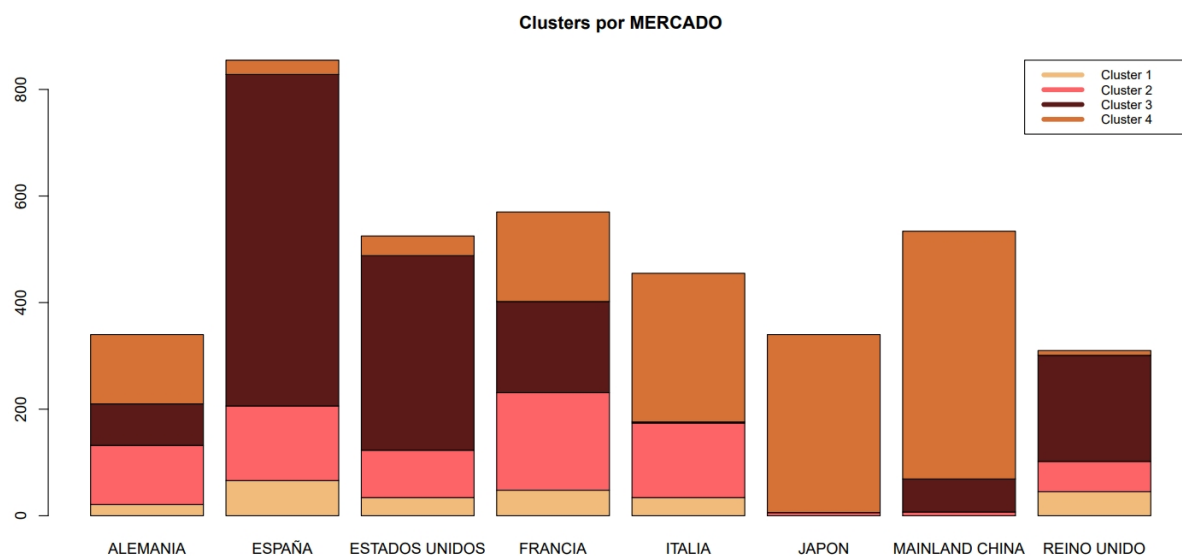


Figura 2.6: Representación de los 4 clúster por los países que conforman el Core Business.

Tabla 2.4: Distribución condicionada de los clústeres por país.

| Clúster | Alemania | España | EEUU | Francia | Italia | Japón | China | UK |
|---------|----------|--------|--------|---------|--------|--------|--------|--------|
| 1 | 8.5 % | 26.6 % | 13.7 % | 19.4 % | 13.7 % | 0.0 % | 0.0 % | 18.2 % |
| 2 | 15.2 % | 19.1 % | 12.1 % | 25.0 % | 19.1 % | 0.8 % | 1.0 % | 7.8 % |
| 3 | 5.2 % | 41.5 % | 24.4 % | 11.4 % | 0.1 % | 0.0 % | 4.1 % | 13.3 % |
| 4 | 9.0 % | 1.9 % | 2.6 % | 11.6 % | 19.3 % | 23.1 % | 32.1 % | 0.6 % |

El análisis jerárquico de clústeres ofrece una visión detallada del comportamiento de los departamentos de compra, las operaciones de tienda y las preferencias del mercado en diferentes regiones, tal como se muestra en la Tabla 2.5. Al segmentar los mercados y compradores en clústeres distintos, el departamento de Gestión del Retail puede utilizar esta herramienta para desarrollar estrategias ajustadas que aborden específicamente las necesidades y oportunidades de cada comprador en cada mercado. Esto no solo mejora la eficiencia y efectividad de las iniciativas de marketing y gestión de inventarios, sino que también asegura una mayor satisfacción del cliente y un enfoque de mercado más personalizado y estratégico.

Tabla 2.5: Resumen de resultados finales por clúster.

| Grupo | Comprador Mayoritario | VARIABLES | Mercado | Comprador minoritario |
|-------|-----------------------|--------------------------------|-------------------------|-----------------------|
| 1 | CIRCULAR | Unidades enviadas y vendidas | España, Francia, UK | Japón, China |
| 2 | BASIC, CIRCULAR | Unidades enviadas y vendidas | España, Francia, Italia | Japón, China |
| 3 | TRF, WOMAN | Ratio devoluciones, Destocajes | España, EEUU | Italia, Japón |
| 4 | PUNTO | Ratio destocajes | Japón, China | UK, España, EEUU |

2.7. Conclusiones y consideraciones para la aplicación práctica

Mediante los resultados obtenidos se observan diferencias a nivel comprador en los mercados que conforman el Core Business, lo que viene determinado también con la cultura y la tipología de cliente, por lo que marcaría la tendencia de las tiendas dependiendo del mercado al que pertenecen.

Se proporciona al equipo de negocio el código del algoritmo como herramienta para la agrupación en clústeres, diseñado para analizar los 5 principales departamentos de compra, con la capacidad de adaptarse para incluir más departamentos de compra y tipo de producto, según sea necesario. También puede extenderse a otras secciones como Caballero y Niño e incluso adaptarse para otras cadenas del grupo. Este algoritmo permite clasificar el comportamiento de los compradores a nivel de mercado y de tienda. Como resultado final, se genera un archivo Excel donde cada tienda de Zara señora está definida por sus principales compradores, clasificados según el clúster correspondiente. Esta clasificación facilita una visualización rápida del comportamiento de cada comprador en todas las tiendas. Dicha clasificación puede ser de gran utilidad al final de cada campaña, ayudando en la predicción de comportamientos de los compradores en los diferentes mercados y tiendas para ayudar a definir la estrategia de venta en futuras campañas; así como para establecer un seguimiento semanal del comportamiento de estos compradores en las tiendas y poder actuar de una manera eficaz y rápida.

Capítulo 3

Modelo Ciclo de tienda

Con el objetivo de optimizar la gestión de los envíos y reducir los destocajes de unidades de vuelta al centro de distribución, este capítulo presenta el desarrollo de un modelo de ciclo de tienda. Este modelo simple pretende reflejar la realidad operativa de las tiendas, para facilitar la planificación estratégica de los envíos en futuras campañas. La relevancia de este modelo radica en su capacidad para adaptarse a las dinámicas cambiantes del retail, donde la precisión en la gestión de los envíos puede significar la diferencia entre el éxito y el exceso de stock no vendido. En un entorno comercial tan competitivo y de rápido movimiento como el de Zara Señora, la eficiencia en la gestión de envíos se convierte en una necesidad. Donde los destocajes no solo representan un coste directo en términos de logística y almacenamiento, sino que también reflejan oportunidades perdidas de venta. Al reducir estos destocajes, la tienda puede asegurar una mejor disponibilidad de artículos para el cliente, a la vez que mejora su rentabilidad general.

3.1. Problema y objetivo

3.1.1. Descripción del problema

Durante toda la campaña, los destocajes representan un desafío operativo significativo en la gestión de inventarios de las tiendas. Estos destocajes ocurren cuando hay un exceso de unidades enviadas a las tiendas que, eventualmente, no se venden. Este exceso de stock no solo ocupa un espacio valioso en los almacenes de las tiendas, sino que también requiere que los artículos sean devueltos a los centros de distribución para liberar espacio, lo que conlleva diversas operativas adicionales. Estas operaciones incluyen la selección de los artículos a destocar, la preparación de estos para el transporte, y la gestión de su recepción y posterior reaprovechamiento en el centro de distribución.

Como se detalla en el [Capítulo 1](#), los destocajes se clasifican en función de las decisiones de negocio desde la central en dos categorías principales:

- **Destocaje de cliente:** originado por las devoluciones de clientes en la tienda física. Estos artículos, que ya no forman parte de la exposición en tienda, deben ser enviados de vuelta al centro de distribución.
- **Destocaje sin cliente:** se produce por operativas internas de la tienda y está directamente relacionado con el exceso de unidades enviadas, lo que representa una oportunidad de mejora en la eficiencia operativa.

El destocaje de cliente es un destocaje producido por las devoluciones recibidas en la tienda física, de tal manera que no se puede controlar, sino hacer un seguimiento. Por ello en el [Capítulo 2](#) se realiza una agrupación por comprador y tienda para identificar cuales son las que experimentan o pueden

llegar a experimentar, al tener un ratio de devoluciones alto, este tipo de operativa. Sin embargo, el destocaje sin cliente, sí que se puede controlar mediante las unidades enviadas, ya que un exceso de unidades enviadas a tienda provoca un aumento de unidades a destocar.

Desde la compañía se propone que mediante un incremento de venta respecto al año anterior, para mejorar un éxito dado, se optimice el envío de unidades para destocar lo mínimo, consiguiendo maximizar la venta con respecto al año anterior.

3.1.2. Objetivo del modelo

El objetivo principal de este modelo es optimizar el envío de unidades a las tiendas con el fin de disminuir los destocajes sin cliente maximizando las ventas, lo cual es un factor controlable que puede ser ajustado proactivamente. La estrategia propuesta busca incrementar las ventas respecto al año anterior, mejorando así el ratio de éxito de las unidades enviadas. Este enfoque no solo pretende maximizar las ventas en comparación con el año anterior, sino también optimizar la gestión del inventario para reducir los costes operativos y mejorar la rentabilidad de la tienda.

Este modelo operativo está diseñado para facilitar la planificación estratégica de futuras campañas y adaptarse a las necesidades cambiantes del mercado y las condiciones de la tienda. Al lograr un equilibrio más eficiente entre las unidades enviadas y las ventas, se puede alcanzar una reducción significativa en los destocajes, consiguiendo una operación más sostenible y rentable.

3.2. Descripción del modelo

Se desarrolla un modelo de optimización matemática para representar la situación por mercado, con este modelo se intenta simular el escenario real de la tienda de una manera más sencilla y simplificada para facilitar la construcción del modelo desde cero.

A continuación se adjunta un diagrama representativo de los flujos de unidades de entrada: unidades enviadas, unidades reaprovechadas de cliente y unidades reaprovechadas de destocaje; y de salida: venta y destocajes.

En la Figura 3.1 se describe el movimiento de unidades del ciclo de tienda, la tienda recibe las unidades enviadas desde un centro de distribución CD1, las cuales pueden salir de la tienda de 2 maneras diferentes, como unidades vendidas o como unidades destocadas.

- **Unidades vendidas:** el cliente completa la compra y se lleva el artículo de la tienda. Sin embargo, en un plazo de un mes se puede producir la devolución de estos artículos volviendo a entrar de nuevo al ciclo de tienda. Si ocurre una devolución de los artículos comprados, estas unidades pueden entrar de nuevo a la tienda física si están presentes en la exposición de la tienda entrando como reaprovechamiento de cliente, o volver a salir de la tienda a un centro de distribución como destocaje de cliente.
- **Unidades destocadas:** son unidades que han sido enviadas a la tienda pero no han sido vendidas y salen de la tienda mediante destocaje sin cliente a un centro de distribución CD1 ó CD2, esto es debido a problemas de capacidad de almacenaje, por unificación de producto en otras tiendas u otros motivos. Estas unidades pueden volver a entrar en la tienda mediante reaprovechamiento destocaje.

Para el desarrollo de este modelo, no se considera la capacidad del almacén de la tienda física, que aunque sea mínima comparada con la zona de exposición de la tienda, parte del producto se queda en el almacén de la tienda como stock.

Diagrama de flujo de las unidades que ocurren en la tienda física (TF)

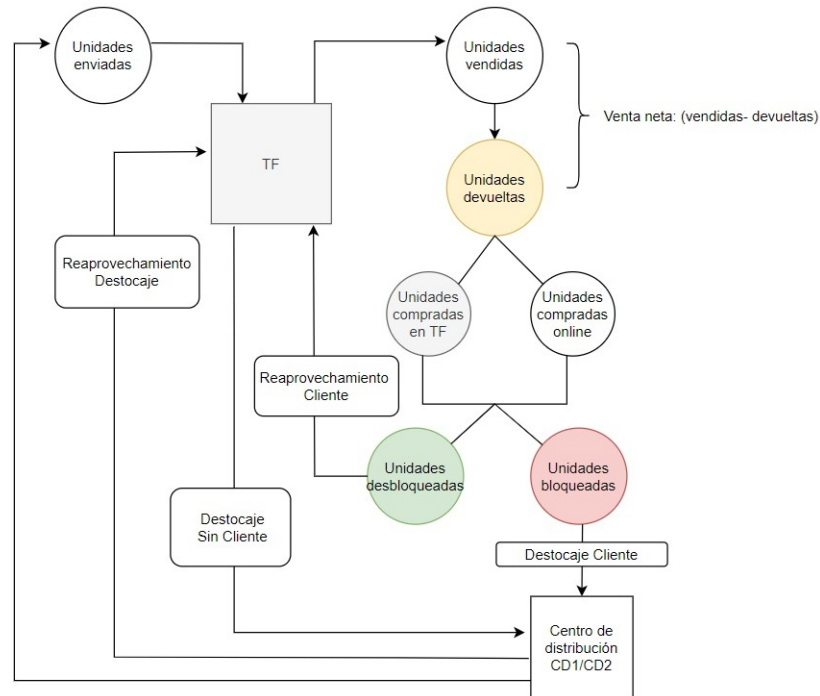


Figura 3.1: Diagrama de flujo de las unidades que ocurren en la tienda física.

Por lo que en el modelo propuesto de ciclo de tienda, en el cual no se considera la capacidad de almacén de la tienda se llamará Exceso de Envío al destocaje sin cliente, ya que nada se quedaría en la tienda, representando que todas las unidades enviadas que no han sido vendidas se destocarán de vuelta al centro de distribución.

3.3. Desarrollo del modelo

A continuación se describe el conjunto, parámetros, variables, función objetivo y restricciones del modelo propuesto para representar la situación del ciclo de tienda, incluyendo todos los procesos de entrada y salida de las unidades de artículos en la tienda física.

3.3.1. Conjunto

Se define el conjunto I como el conjunto de los compradores principales de Zara Señora: BASIC, CIRCULAR, PUNTO, TRF y WOMAN.

- $I = \{1, 2, \dots, N\}$. Para cada $i \in I$, donde $i = 1, 2, \dots, N$.

3.3.2. Parámetros

- $VENTA_i$: unidades vendidas la campaña del año anterior.
- $EXITO_i$: unidades de venta bruta en función de las unidades enviadas.

Éxito = unidades venta bruta / (unidades enviadas + unidades de [continuidad](#)) ¹

- $RATIODEVO_i$: unidades devueltas en tienda por el cliente en función de las unidades vendidas.
Ratio devolución = unidades devolución / unidades venta bruta
- K : penalización para los destocajes, en función de lo agresiva que se quiera ser. Este parámetro varía de 0 a 1, cuanto más pequeño sea este parámetro menos penalización habrá en los destocajes y se enviará hasta el máximo permitido, por el contrario, cuanto más grande sea K se penalizarán más los destocajes, siendo más agresivo, hasta llegar al mínimo envío de unidades requeridas.
- IV : incremento Venta neta, el driver a escoger en función de incrementar/aumentar X puntos de venta.
- R_{C_i} : porcentaje de reaprovechamiento de las unidades recibidas en tienda devueltas por el cliente. Se aprovecharán las prendas que cumplan ciertos requisitos (que estén en exposición) y las que no se destocarán directamente como destocaje de cliente, en este caso de estudio se marca un mismo porcentaje para todos los compradores.
- G_{C_i} : porcentaje de gestión de reaprovechamiento de cliente, en este caso de estudio se marca un 1, considerando que no tiene coste reaprovechar al llegar ya a la tienda. Variará entre 0 y 1 (donde igual a 0 significaría que el coste es tan elevado que no merece la pena reaprovechar la venta de la prenda en tienda). Pudiéndose considerar a futuro que tuviese un coste por llegar tarde.
- R_{D_i} : porcentaje de reaprovechamiento de las unidades de exceso de envío (destocaje sin cliente). En este caso de estudio se marca un un mismo porcentaje para todos los compradores.
- G_{D_i} : porcentaje de gestión de reaprovechamiento del exceso de envío (destocaje). Variará entre 0 y 1 (donde igual a 0 significaría que el coste es tan elevado que no merece la pena reaprovechar la venta del artículo en tienda). Se fija con el precio medio del artículo reaprovechado en función del comprador.

3.3.3. Variables

Teniendo en cuenta que el objetivo del modelo es optimizar los envíos penalizando los destocajes maximizando las ventas, se definen las siguientes variables:

- $Envio_i$: cantidad de unidades a enviar.
- $ExcesoEnvio_i$: cantidad de unidades a destocar sin considerar la capacidad.
- $VentasNetasR_i$: beneficio de las ventas netas reaprovechadas.

3.3.4. Función objetivo y restricciones

A continuación, se define la función objetivo para maximizar las ventas penalizando las unidades destocadas mediante el parámetro K .

$$\text{máx} \quad \sum_{i \in I} \left(VentasNetasR_i + ExcesoEnvio_i \cdot R_{D_i} \cdot G_{D_i} \right) \quad (3.1)$$

$$- \frac{K}{100000} \cdot (ExcesoEnvio_i \cdot (1 - R_{D_i}))^2 \quad (3.2)$$

¹Se tendrá en cuenta las unidades que formen parte de la continuidad, si el modelo se desarrolla al principio de campaña, debido a que muchas unidades que forman parte de la campaña anterior se pasan a la nueva campaña, justo antes del comienzo de las rebajas, lo que significa que ya forman parte del stock de la tienda.

Mediante las siguientes restricciones:

s.a.

$$Envio_i \geq \frac{VENTA_i}{EXITO_i} \quad \forall i \in I \quad (3.3)$$

$$\sum_{i \in I} Envio_i \leq IV \sum_{i \in I} \left(\frac{VENTA_i}{EXITO_i} \right) \quad (3.4)$$

$$VentasNetasR_i = Envio_i \cdot EXITO_i \cdot (1 - RATIODEVO_i + G_{C_i} \cdot R_{C_i} \cdot RATIODEVO_i) \quad \forall i \in I \quad (3.5)$$

$$ExcesoEnvio_i = Envio_i - Envio_i \cdot EXITO_i \cdot (1 - RATIODEVO_i + R_{C_i} \cdot RATIODEVO_i) \quad \forall i \in I \quad (3.6)$$

Donde la **función objetivo** se compone de 2 términos:

- El término (3.1) representa las ventas netas obtenidas de las unidades reaprovechadas junto con el reaprovechamiento de los excesos de envío por los gastos de gestión de cada comprador.
- El término (3.2) representa la penalización al cuadrado de los excesos de envío que no se reaprovechan, donde K representa la agresividad que se quiera optar en la penalización.

Las **restricciones** del modelo son las siguientes:

1. La ecuación de Envío Mínimo (3.3): asegura que las unidades a enviar no sean inferiores a las necesarias para satisfacer la demanda de la tienda, basada en los históricos de la campaña anterior.
2. La ecuación de Incremento de Venta Neta (3.4): limita el total de envíos a un incremento especificado por el parámetro IV, que refleja los objetivos estratégicos de crecimiento en ventas netas de la compañía.
3. La ecuación Ventas Netas Reaprovechadas (3.5): define las ventas netas reaprovechadas ajustando las ventas brutas por las devoluciones, el reaprovechamiento de cliente y la gestión de cliente.
4. La ecuación de Exceso de Envío (3.6): define las unidades que quedan sin vender después del ciclo de envío-devolución-reaprovechamiento de cliente.

3.3.5. Características del problema de optimización

El problema de optimización que se aborda en este modelo incorpora características fundamentales que optimizan la estrategia de envío y gestión para los compradores principales de Zara Señora. La función objetivo y las restricciones del modelo están diseñadas para maximizar la eficiencia operativa asociados con el exceso de envío.

Conjunto convexo

Las restricciones lineales impuestas garantizan que:

- Los envíos no sean menores que las ventas de la campaña anterior, garantizando un envío mínimo que cubra la demanda impuesta por la tienda.
- Los envíos no sean mayores que la limitación impuesta en función del incremento de ventas deseado.
- Las ventas netas y los excesos de envío están ajustados por los ratios de éxito y devolución de la tienda.

Función objetivo cóncava

La función objetivo es cuadrática, de tal manera que se introduce una penalización que crece cuadráticamente conforme aumentan los excesos de envío. Lo que permite controlar y limitar estos de una manera agresiva, promoviendo la eficiencia de los envíos a la tienda. Además el término cuadrático negativo del parámetro K , asegura que la función objetivo sea cóncava para maximizar. Por lo tanto, esto garantiza que cualquier solución óptima local será también una solución óptima global sobre el conjunto convexo definido por las restricciones lineales. Tal como se exploró en las asignaturas de Programación Lineal y Optimización Aplicada, cursadas durante el programa del máster, esta propiedad resulta crucial para la aplicación práctica de los modelos de optimización. De tal manera que asegura una eficiencia significativa en el proceso de optimización, un aspecto fundamental para el éxito del modelo.

3.4. Estudio computacional

El principal objetivo del estudio computacional es comprobar cómo el modelo diseñado refleja correctamente las distintas operativas de la tienda. El lenguaje de programación elegido para la implementación y resolución del modelo es AMPL, mediante el solver Gurobi 10.0.2. Los análisis se realizaron en un ordenador personal con una CPU Intel(R)Core(TM) i7-8550UCPU@1.80GHz de 64 bits y 16 GB de RAM, donde los tiempos de ejecución han sido instantáneos debido a la sencillez del problema.

La base de datos utilizada para el modelo de ciclo de tienda, se consigue mediante el acumulado en unidades de la campaña Invierno 2023 y 2022 para el mercado España. Junto con el flujograma de negocio y los datos de gestión del flujo de coste unitario del destocaje por prenda.

Los parámetros escogidos para alimentar al modelo e los distintos experimentos son los siguientes:

- **Venta** del año anterior por comprador.
- **Éxito** del año anterior por comprador contando con la continuidad, ya que los datos acumulados para el análisis son de fecha de inicio de campaña. Para ello se procede a realizar un cálculo de las unidades que ya estaban en tienda y formaban parte de la anterior campaña.
- **Ratio de devolución** por comprador.
- R_C , tasa de reaprovechamiento de cliente por comprador.
- R_D , tasa de reaprovechamiento del exceso de envío por comprador.
- G_C , tasa de la gestión de reaprovechamiento del cliente, el cual se considerará uno, asumiendo que no hay un gasto de gestión asociado debido a que es el cliente el que lleva el artículo a la tienda. Sin embargo, figura en el modelo por si en un futuro se considera un gasto asociado.
- G_D , gestión de reaprovechamiento del exceso de envío el cual es calculado por el coste unitario del destocaje por unidad de prenda por comprador.
- K , penalización por excesos de envío, donde se realizará un estudio variando la influencia de variar este parámetro.
- **IV**, incremento de venta esperado que se repartirá entre todos los compradores.

Para realizar el estudio se considerarán los siguientes escenarios en el mercado España:

1. A nivel agregado por comprador: verificación con los datos de la campaña 2022 y proyección para la campaña Invierno 2023.
2. A nivel desagregado por comprador para la campaña de Invierno 2023: análisis de las unidades de envío en función del parámetro K con un incremento de venta del 13 % mediante los datos de la campaña Invierno 2022.
3. Con datos de final de campaña Invierno 2023: proyección para alcanzar el mismo incremento de venta del 13 % en la campaña Invierno 2024 variando el parámetro K.

3.5. Validación y resultados

La validación del modelo computacional constituye un paso esencial para asegurar su fiabilidad y precisión antes de su aplicación práctica. En este estudio, se evaluó el comportamiento del modelo bajo diferentes condiciones para garantizar su robustez y capacidad de replicar resultados realistas y útiles para la toma de decisiones. A continuación, se detalla la validación del modelo utilizando datos reales de la campaña de Invierno 2022, para el mercado España, uno de los más significativos por su volumen de operaciones y ventas.

Esta sección se centra en tres escenarios específicos diseñados para probar la adaptabilidad del modelo a variaciones en los parámetros de entrada y para observar el impacto de dichas variaciones en los resultados de la simulación. Estos escenarios son críticos para entender cómo diferentes factores, como los cambios en las políticas de envío por comprador y las condiciones de venta y devolución del mercado, pueden influir en el desempeño del modelo. Cada escenario es descrito y analizado a continuación:

1. **Validación Agregada:** se verifica la capacidad del modelo para reproducir los resultados de ventas basándose en datos históricos acumulados hasta la fecha 02.10.2022, ajustando posteriormente para proyectar las ventas de la campaña de Invierno 2023.

Tal como se observa en la tabla 3.1, que son los datos acumulados de la campaña de Invierno 2022 e Invierno 2023 en la misma semana correspondiente al 2 de Octubre. Se obtiene un incremento de venta neta del 13 % en 2023 respecto a la misma semana del 2022. Por lo que el análisis de validación del modelo se usará un incremento de venta del 13 % en todos los casos de estudio.

Tabla 3.1: Datos de unidades acumuladas $\cdot 10^6$ correspondientes a la campaña Invierno 2022 y 2023.

| Año | Facturado | Venta Bruta | Devoluciones | Venta Neta | Destocajes | Destocaje Cliente |
|------|-----------|-------------|--------------|------------|------------|-------------------|
| 2022 | 15.7 | 12.8 | 2.00 | 10.8 | 3.90 | 0.10 |
| 2023 | 17.9 | 14.3 | 2.10 | 12.2 | 5.36 | 1.32 |

Variando el parámetro K, para la penalización en función de lo agresiva que se quiera ser penalizando los excesos de envío. Se obtiene los siguientes resultados que muestra la Tabla 3.2, a nivel agregado para la sección de Zara Señora en función del parámetro K:

- K=0.1 siendo agresiva, penalizando al máximo los excesos de envío.
- K=0.001 siendo lo menos agresiva posible, relajando al máximo los excesos de envío.

Tabla 3.2: Resultados de las unidades $\cdot 10^6$ que se enviarían en campaña 2023 en función de la K con los datos de campaña 2022, mediante el modelo propuesto.

| IV: 1.13 | Incremento de envío % | Facturado | Venta Bruta | Devoluciones | Venta Neta | Exceso de Envío | Destocaje Cliente |
|----------|-----------------------|-----------|-------------|--------------|------------|-----------------|-------------------|
| K=0.1 | 1.00 | 15.7 | 12.8 | 2.0 | 10.8 | 6.6 | 1.1 |
| K=0.001 | 1.13 | 18.0 | 14.5 | 2.3 | 12.2 | 7.4 | 1.3 |

Por lo que utilizando los datos acumulados del 2022 penalizando al máximo los excesos de envío mediante el parámetro $K=0.1$, el modelo envía lo mínimo para cumplir con la demanda del año anterior. Sin embargo al relajar la penalización se observa que incrementando el envío un 13%, se conseguirían las ventas esperadas. Teniendo en cuenta que el exceso de envío obtenido es el destocaje sin cliente sin contar con la capacidad del almacén de la tienda.

2. **Análisis por Comprador:** este escenario evalúa cómo el modelo distribuye las unidades de envío entre los principales compradores de Zara Señora: BASIC, CIRCULAR, PUNTO, TRF y WOMAN, en función del parámetro K. Se analiza cómo esta distribución afecta el rendimiento global en términos de ventas y logística. Alimentando al modelo con los datos del acumulado de campaña de Invierno 2022 hasta la fecha 02.10.2022.

Considerando un incremento de venta del 13%, se presentan a continuación los resultados por comprador en función de la variación del parámetro K. En el extremo más agresivo ($K=0.7$), el modelo penaliza al máximo los excesos de envío, enviando las unidades mínimas necesarias para cumplir con la restricción de mantener los niveles de envío de la campaña anterior de Invierno 2022. En el otro extremo, con un enfoque más relajado hacia los excesos de envío ($K=0.01$), se ajusta la cantidad enviada a tienda tomando en cuenta los artículos que ya estaban en tienda de la campaña anterior de Verano 2023 y que pasaron a la nueva campaña de Invierno 2023.

Se examinan las variaciones del parámetro K y cómo estas influyen en la priorización de los envíos para cada comprador. Los resultados detallados se presentan en las tablas del [Anexo B](#) desde la Tabla B.3 hasta la Tabla B.8.

Se considera un incremento de venta del 13%, y mediante la variación del parámetro K, se ajusta la agresividad en la penalización de los excesos de envío. Esta variación influye en cómo se priorizan los envíos para cada comprador, lo cual se refleja claramente en la Tabla 3.3. Los resultados muestran cómo la elección de K afecta directamente a las unidades enviadas en cada comprador, evidenciando una estrategia más o menos conservadora dependiendo del valor de K seleccionado.

En función de la agresividad deseada en la penalización de los excesos de envíos, el modelo ajusta la priorización entre los compradores. Con un $K=0.7$, donde la penalización es máxima, el modelo no prioriza a ningún comprador específico y opta por el envío mínimo. A medida que se reduce el valor de K, comienza a aumentar la priorización hacia los compradores WOMAN y PUNTO para alcanzar el incremento de ventas propuesto del 13%. A medida que la penalización se relaja, los envíos para el comprador PUNTO disminuyen, mientras que aumentan los asignados a WOMAN, hasta llegar a la situación de $K=0.01$ donde se minimiza la preocupación por los excesos de envío y se da prioridad máxima a WOMAN. Tal como se muestra en la Figura 3.2.

Es importante destacar que los compradores BASIC, CIRCULAR y TRF no se incluyen en las estrategias prioritarias del modelo para alcanzar el incremento de ventas propuesto. La exclusión

Tabla 3.3: Incremento de Envío en función del parámetro K para la campaña Invierno 2023 por comprador.

| Comprador | K=0.7 | K=0.4 | K=0.2 | K=0.08 | K=0.05 | K=0.01 |
|-----------|-------|-------|-------|--------|--------|--------|
| BASIC | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| CIRCULAR | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| PUNTO | 1.00 | 1.41 | 1.43 | 1.31 | 1.19 | 1.00 |
| TRF | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| WOMAN | 1.00 | 1.58 | 1.68 | 1.77 | 1.85 | 1.99 |
| TOTAL | 1.00 | 1.12 | 1.13 | 1.13 | 1.13 | 1.13 |

de estos compradores se debe a que su incorporación llevaría a un aumento significativo en los excesos de envío, lo que sería penalizado severamente bajo la estructura actual de la función objetivo del modelo. Esta función no busca minimizar los destocajes de manera directa, sino que penaliza los excesos de envío de forma que promueve una reducción en los mismos como una estrategia óptima. Por lo tanto, el modelo se centra en aquellos compradores cuya participación en la estrategia de distribución contribuye a alcanzar los objetivos de ventas con las menores penalizaciones posibles por excesos de envío.

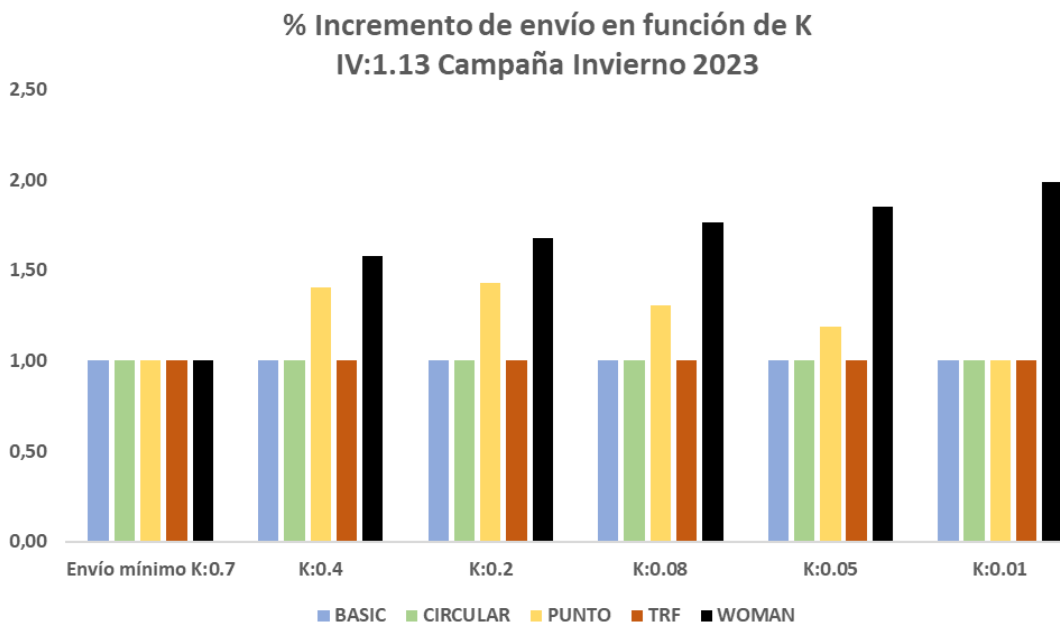


Figura 3.2: Incremento de envíos por comprador variando el parámetro K de 0.7 a 0.01, para el análisis por comprador de la campaña Invierno 2023.

3. **Proyección Futura:** utilizando los datos de finalización de la campaña de Invierno 2023, se simula la campaña de Invierno 2024 para predecir los envíos con un incremento esperado del 13%, variando el parámetro K para observar diferentes resultados y optimizar las decisiones de envío.

Utilizando los datos de la campaña de Invierno 2023 y asumiendo un incremento de ventas del 13%, se analizan los resultados en función del parámetro K. Con K=1, se configura la condición de envío mínimo, y a medida que se relaja la penalización por excesos de envíos, se empieza a priorizar a los diferentes compradores con el objetivo de alcanzar el incremento propuesto en las ventas, penalizando los excesos en los envíos. Este proceso termina priorizando principalmente al comprador BASIC. La secuencia de priorización observada es WOMAN, PUNTO, TRF y finalmente BASIC, tal como se detalla en la Tabla 3.4 y se muestra en la Figura 3.3.

Tabla 3.4: Incremento de Envío en función del parámetro K para la campaña Invierno 2024 por comprador.

| Comprador | K=1 | K=0.7 | K=0.4 | K=0.2 | K=0.08 | K=0.05 | K=0.01 |
|-----------|------|-------|-------|-------|--------|--------|--------|
| BASIC | 1.00 | 1.00 | 1.02 | 1.045 | 1.12 | 1.17 | 1.49 |
| CIRCULAR | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| PUNTO | 1.00 | 1.00 | 1.09 | 1.06 | 1.00 | 1.00 | 1.00 |
| TRF | 1.00 | 1.00 | 1.18 | 1.15 | 1.04 | 1.00 | 1.00 |
| WOMAN | 1.00 | 1.38 | 1.92 | 1.93 | 1.93 | 1.86 | 1.00 |
| TOTAL | 1.00 | 1.04 | 1.13 | 1.13 | 1.13 | 1.13 | 1.13 |

Los resultados detallados se presentan en las tablas del [Anexo B](#) desde la Tabla B.9 hasta la Tabla B.15.

3.6. Interpretación del modelo de ciclo de tienda

Con los resultados obtenidos en el apartado anterior, se comprueba que el modelo se explicaría mediante la siguiente fórmula:

$$\frac{\text{EXITO}}{\text{VENTA} \cdot (1 - \text{EXITO} \cdot (1 - \text{RATIODEVO} + G \cdot R \cdot \text{RATIODEVO}))^2} \quad (3.7)$$

Indicando que:

- Cuanto mayor sea EXITO mejor, pues menos envíos acaban teniendo que destacarse convirtiéndose en ventas.
- El término $(1 - \text{EXITO} \cdot (1 - \text{RATIODEVO}))$ representa el porcentaje de las unidades enviadas que acabará siendo exceso de envío. Cuanto más grande sea en peor situación de exceso de envío se encontraría la tienda, ya que al elevarlo al cuadrado penalizaría aún más esta generación de exceso de envío.

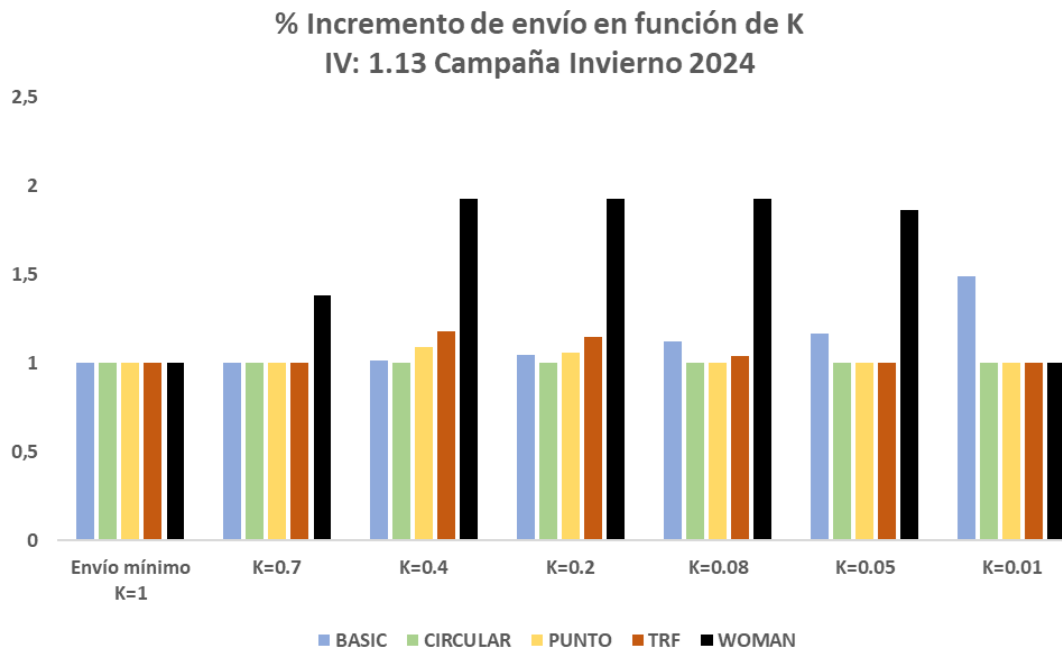


Figura 3.3: Incremento de envíos por comprador variando el parámetro K de 1 a 0.01, para la proyección de la campaña de Invierno 2024.

- El término $VENTA \cdot (1 - EXITO \cdot (1 - RATIODEVO + G \cdot R \cdot RATIODEVO))^2$ representa que la penalización anterior debe tener en cuenta que a mayor VENTA más exceso de envío generará si se satisface íntegramente (a más venta mayor capacidad de generar beneficio, pero también mayor stock se genera por lo que más exceso se producirán).
- Cuanto mayor sea el reaprovechamiento R y el parámetro de gestión, mejor, pues disminuirá la cantidad de porcentaje de exceso e envío.

3.7. Conclusiones y Consideraciones para su Aplicación Práctica

Este modelo es una representación simple del ciclo de la tienda que funciona como una herramienta de optimización para ayudar a tomar decisiones más informadas sobre cuánto producto enviar a cada comprador y cuanto exceso de envío se produciría en función de la agresividad del parámetro K. Contando con el ajuste de la estrategia de incremento de venta, reduciendo los excesos de envío basándose en datos de la campaña anterior del éxito de las ventas, la tasa de devoluciones, el reaprovechamiento y gestión.

El modelo capta la interacción entre las ventas esperadas, el historial de éxito de ventas, y la eficacia en el reaprovechamiento de las devoluciones para optimizar los envíos. Tiene en cuenta cómo se espera aumentar las ventas en la nueva campaña por comprador, ajustando los envíos según lo sucedido en la campaña del año anterior. Permitiendo capturar las diferencias entre los principales compradores para determinar dónde podría ser más beneficioso intensificar los envíos y en cuales reducirlos.

Sin embargo, el modelo asume que las relaciones y tendencias pasadas continuarán en el futuro, lo cual no siempre puede ser cierto. Esto sucede en algunos compradores que no han conseguido la venta esperada comparado con campañas anteriores. Por lo que se propone realizar un estudio a

principio de campaña y ajustarlo a medida que la campaña avance dependiendo de la recepción del cliente a la campaña actual. Además hay que tener en cuenta que al no disponer de las capacidades a nivel comprador la capacidad no se considera en el desarrollo del modelo. Por lo que en el desarrollo del modelo en un futuro se aconseja agregar este parámetro al modelo, para ser más precisos en la estimación de los destocajes.

Hay que considerar que es un modelo que necesita un futuro desarrollo para aproximar el ciclo de tienda a la situación real de la tienda, en el cual se debe incorporar las capacidades y el canal online. Además de implementar una serie de ajustes de negocio. Para que el modelo se pueda desarrollar por los equipos de la compañía, se crea una aplicación en Excel a partir de Solver, donde se incluyen los distintos parámetros, drivers, ajustes y penalizaciones de interés para el equipo de Gestión de Control de Zara para el cálculo de estrategia de la campaña Invierno 2024, con los datos de campaña de invierno 2023 y 2022. La creación de esta aplicación en Excel, facilita su futuro desarrollo al ajustarse al software de la compañía, con la posibilidad de ampliar el número de compradores y aplicarlo a diferentes secciones y cadenas.

Capítulo 4

Conclusiones del Análisis y Consideraciones para su Aplicación Práctica

Este trabajo ha abarcado dos enfoques fundamentales en el análisis de datos aplicados al contexto de Zara: la agrupación de los compradores mediante clústeres para entender su comportamiento en cada una de las distintas tiendas y el modelo de ciclo de tienda para optimizar la logística de envíos y gestión de inventarios. Cada capítulo ha aportado perspectivas que se resumen a continuación:

- **Agrupación por clústeres:** el análisis reveló diferencias en los comportamientos de compra que varían según la cultura y tipología del cliente, influyendo directamente las estrategias de venta y marketing. La implementación del algoritmo de clústeres ha demostrado ser una herramienta eficaz para segmentar a los compradores, permitiendo una visualización y análisis detallado del comportamiento en diversas tiendas. Esto facilita una planificación más estratégica y personalizada para las futuras campañas de marketing. Un caso ilustrativo es el de los mercados de Italia, Japón y China, donde el algoritmo de clúster diseñado ha identificado diferencias en el comportamiento de devoluciones, atribuibles a las políticas de devolución específicas de cada país. Este conocimiento permite adaptar las estrategias locales de manera más efectiva, optimizando tanto la experiencia del cliente como la gestión del inventario.
- **Modelo de ciclo de tienda:** el modelo proporcionó un marco para simular y optimizar los envíos de productos a las tiendas, considerando variables críticas como las ventas previas, las devoluciones y el reaprovechamiento de los artículos. A pesar de su naturaleza simplificada, el modelo capturó eficazmente la dinámica entre las expectativas de venta y la logística de distribución, ofreciendo una base sólida para tomar decisiones informadas y reducir los costes operativos producidos por los destocajes. Un aspecto destacable de su aplicación es la planificación para la próxima campaña de invierno 2024, donde se priorizan segmentos de compradores como WOMAN, TRF y BASIC. Esta priorización se basa en análisis de datos históricos que proyectan altas expectativas de venta sin el riesgo de generar un exceso de stock. Además, se consideran los gastos asociados con la gestión del reaprovechamiento, asegurando una operación más eficiente y rentable.

La implementación de estos modelos proporciona al equipo de negocio de Zara Gestión herramientas robustas para mejorar la precisión en la planificación y ejecución de estrategias comerciales. Estas herramientas no solo mejoran la eficiencia operativa sino que también enriquecen la toma de decisiones estratégicas, permitiendo ajustes ágiles en respuesta a las dinámicas del mercado.

De cara al futuro, existen diversas áreas en las que es posible expandir y perfeccionar estos modelos:

- **Expansión del modelo de clústeres:** explorar la aplicación de este modelo a otras secciones, departamentos de compra y regiones geográficas para generalizar su utilidad en diversas condiciones de mercado.
- **Integración del canal online:** incorporar análisis de ventas online para ofrecer una visión más realista del comportamiento del consumidor y optimizar aún más los niveles de stock y distribución.
- **Incorporación de las capacidades de almacén:** integrar las capacidades de almacén de la tienda al modelo del ciclo de tienda como parámetro, beneficiará la gestión del inventario ya que las decisiones sobre cuánto producto y qué artículos enviar será más precisa.
- **Adaptación continua:** continuar adaptando y ajustando los modelos con datos en tiempo real para mejorar la precisión y la eficacia operativa.

Estas recomendaciones están dirigidas hacia una utilización más integral y eficiente de los recursos y capacidades analíticas disponibles, proyectando hacia un futuro en el que la adaptabilidad y la innovación se mantengan como pilares fundamentales de la estrategia comercial en Zara, contribuyendo significativamente a la sostenibilidad y el crecimiento estratégico de la empresa.

Apéndice A

Glosario de Variables de negocio

- **AGRUPACIÓN DE MERCADOS:** Agrupación comercial de mercados en función de su participación en la venta o del área geográfica donde estén ubicados.
- **ALMACÉN TIENDA:** Espacio o conjunto de espacios dentro o fuera del local o locales comerciales de una tienda física dedicados a almacenar la mercancía de cada sección. El stock del almacén tienda es único y se considera stock de tienda física, independientemente de los espacios físicos donde se encuentre almacenado.
- **CADENA:** Marca comercial dentro del grupo Inditex que aglutina un conjunto de tiendas, productos, instalaciones y recursos humanos gestionados de manera independiente.
- **CAMBIO DE CAMPAÑA:** Decisión comercial por la que se decide traspasar el artículo de una campaña a otra diferente a la actual. Por ejemplo: los artículos que no se deciden saldar, se traspasan a la siguiente campaña. Esta decisión no se toma a nivel artículo, se decide a nivel artículo-país, pudiendo estar en saldo un artículo en un país pero en otro no-(ver CONTINUIDAD).
- **CAMPAÑA :** Periodo de tiempo en el que se comercializan los productos asociados a verano o invierno. Se identifica por el año y por verano e invierno o sus abreviaturas V e I (I19, V20), siendo el año en el que se finaliza la campaña. Cada campaña tiene un único presupuesto de compra. A efectos de venta, en cada campaña hay dos periodos: Temporada y Rebajas. También a efectos de venta pueden convivir una campaña de verano con una de invierno, pero no dos de verano o dos de invierno. Se denominan Campañas Contiguas a aquellas con distinto tipo de campaña (V19 e I19) y Campañas Análogas a aquellas con el mismo tipo campaña (V19 y V20).
- **CAPACIDAD:** Unidades máximas que puede tener una tienda en función a sus dimensiones.
- **CAPACIDAD ALMACÉN:** Unidades máximas que puede tener una tienda en función a sus dimensiones de almacén.
- **CAPACIDAD EXPOSICIÓN:** Capacidad de la zona de exposición de tienda.
- **CD1:** Centro de Distribución Primario: CD de producto terminado que reciben mercancía de proveedor y envían a tiendas o a otros almacenes. La característica que los define sobre el resto es que reciben mercancía de proveedor como operación recurrente.
- **CD2:** Centro de Distribución Secundario: CD de producto terminado que reciben mercancía desde almacenes primarios para su redistribución a cliente online o tiendas.

- **CLASIFICACIÓN DE TIENDAS:** Agrupación de tiendas en base a diferentes criterios de clasificación. Existen tres clasificaciones de tiendas:

Sección: relativa a la clasificación general de la tienda en Señora, Caballero y Niño. No tiene impacto en el volumen de unidades que se envían a tienda, pero de ella se extraen las tiendas top y top comerciales que suelen tener prioridad en el envío de novedades si no hay stock suficiente para todas las tiendas seleccionadas. Se suele recalcular una vez al año.

Comprador : utilizada ocasionalmente por algún departamento de Compras para configurar grupos de reparto de tiendas. Se suele recalcular cada campaña o a petición del comprador.

Subfamilia : es la que tiene impacto en el volumen de unidades de reposición que se enviarán a las tiendas, pues los días de cobertura que se apliquen van a depender del tipo de clasificación que tenga la tienda en cada una de las subfamilias (con independencia de la clasificación que tenga la tienda en la sección y comprador). Se recalcula al menos una vez cada campaña.

- **CLASIFICACIÓN TIENDAS POR COMPRADOR:** Agrupación de tiendas en base a su comportamiento en ventas (Venta de la demanda sin saldo) en cada comprador. Los clústeres y percentiles serán configurables por el usuario de cadena/sección y serán, por ejemplo, del tipo Top A,B,C y D.
- **CLIENTE:** Usuario que ha realizado alguna compra online o que se ha registrado en alguna de las tiendas.
- **COMPRADOR:** Departamento de compra dentro de cada sección. También se denomina Compradores a las personas que realizan las tareas de compra dentro de cada Comprador (ver definición).
- **CONTINUIDAD:** Cambio de campaña de un Mocaco (ver definición) a la campaña contigua siguiente en base a una decisión comercial. Suele realizarse en artículos básicos de gran volumen de venta que no se quieren incluir en Rebajas, aunque puede haber excepciones como continuidades en artículos que no hayan tenido venta en la campaña. Existen dos momentos en los que se realizan grandes continuidades cada año, aproximadamente un mes antes de que empiecen las rebajas de las campañas de Verano e Invierno. Durante las semanas previas a ese momento cada comprador selecciona los artículos candidatos a continuidad que son revisados y aprobados por Dirección Comercial para su ejecución durante la gran continuidad. Además, se realizan otras continuidades puntuales durante cualquier otro periodo en base a decisiones comerciales específicas u otras casuísticas. Finalmente se realizan también continuidades conocidas como ficticias, normalmente en artículos que funcionan bien durante el saldo y se quieren seguir vendiendo como de temporada en la campaña siguiente. La Continuidad también puede hacerse a nivel Moca en el caso de que se quieran cambiar todos los Mocacos del mismo (traspaso). También puede haber retrotraspaso (traspaso de la campaña nueva a la antigua).
- **DESTOCAJES:** Unidades totales devueltas desde localizaciones del tipo tienda (tiendas físicas, almacenes locales, showrooms o CD2 SGE) a CD1 y CD2 AS400. Recoge principalmente el destocaje de tiendas a almacenes. Hay varios tipos de destocaje:

D1 Guiado: producido para agrupar producto entre las tiendas.

D1 Cliente: por las devoluciones producidas por el cliente en la TF que no se pueden reaprovechar en la tienda debido que son prendas que no están ya en exposición, Se debe considerar que las devoluciones de clientes en TF tienen varias procedencias: misma TF, otra TF del país y online.

D2: destocaje producido por problemas de capacidad.

Otros: retirada de la venta, especiales, etc

- **DEVOLUCIÓN VENTA:** Devolución realizada por un cliente en una tienda física o en .com. Se imputan siempre a la tienda a la que se le imputó la venta, independientemente de donde se haya realizado físicamente la devolución. Se valoran a su precio de venta en el momento de la venta.
- **ENVIADO:** Unidades enviadas a tienda considerando no solo la mercancía facturada desde los centros de distribución si no también el resto de los movimientos y ajustes de entrada y salida de mercancía en la tienda distintos a la venta.
- **EXCESO DE ENVÍO:** Destocaje producido sin contar con la capacidad de almacén de la tienda.
- **ÉXITO ENVIADO:** Unidades de venta bruta entre unidades enviadas:

$$\text{Éxito} = \frac{\text{unidades vendidas}}{\text{unidades enviadas}}$$

- **FACTURADO TOTAL:** Unidades facturadas desde CD1 y CD2 AS400 a localizaciones del tipo tienda (tiendas físicas, almacenes locales, showrooms o CD2 SGE). Las unidades facturadas generan un albarán.
- **MERCADO:** Se considera un agrupador de tiendas de cadenas y almacenes y se corresponde con divisiones geográficas de nación, región, provincia o territorio. Se prefiere al término “país” por ser una división determinada libremente, que no coincide de manera necesaria con los límites políticos.
- **MOCACO:** Producto potencialmente comercializable referenciado por MOdelo, CALidad y COlor, incluyendo todas sus variaciones a nivel talla.
- **PERIODO CAMPAÑA:** Indica el momento temporal dentro de una campaña: Temporada o Rebajas.
- **RATIO DESTOCAJE CLIENTE:** Unidades destocadas por D1 Cliente entre unidades enviadas.
Ratio destocaje cliente = unidades destocadas D1 Cliente/unidades enviadas
- **RATIO DESTOCAJE SIN CLIENTE:** Unidades destocadas por D Guiado, D2 y Otros entre las unidades enviadas.
Ratio destocaje sin cliente = (unidades destocadas D1 Guiado + D2 + Otros)/unidades enviadas
- **RATIO DEVOLUCIÓN:** Unidades devolución venta entre unidades de venta bruta:
Ratio devolución = unidades devolución/unidades venta bruta
- **REAPROVECHAMIENTO:** Porcentaje de unidades vendidas que proceden de unidades destocadas (ver definición de “Devoluciones a Almacén: Destocaje”) y unidades devueltas por el cliente. Existen varios tipos de reaprovechamiento:
 - **Reaprovechamiento Cliente:** Porcentaje de unidades vendidas en la tienda que provienen de unidades devueltas por el cliente en la TF. Estas unidades pueden proceder de la misma TF, de otra TF del país o de la compra online.
 - **Reaprovechamiento Cliente Online:** Porcentaje de unidades vendidas en online que provienen de unidades devueltas por el cliente, que fueron adquiridas y devueltas a través de la plataforma online.
 - **Reaprovechamiento destocaje:** Porcentaje de unidades vendidas que provienen de unidades destocadas que se redistribuyen nuevamente para su puesta en venta.

- **REBAJAS:** Espacio temporal dentro de una campaña correspondiente al periodo de Rebajas. También se denomina Saldo.
- **SECCIÓN:** Agrupación en que se divide cada una de las cadenas en base al público objetivo de los productos que comercializa. Las secciones son: señora, caballero y niño.
- **TEMPORADA:** Espacio temporal dentro de una campaña correspondiente al periodo de Temporada.
- **TIENDA ACTIVA:** Tienda física o virtual que ha tenido venta y/o enviado > 0 durante la campaña. Se calcula a nivel sección-tienda, tipo de producto y campaña.
- **TIENDA FÍSICA:** Local donde se realiza la venta directa al público de los artículos del grupo (TF).
- **TIPO PRODUCTO:** Tipología de artículo según su naturaleza, e.g. ropa, calzado, perfumería, etc
- **VENTA ACUMULADO:** Venta descontando las devoluciones de clientes También se conoce como venta neta y también como acumulado.
- **VENTA BRUTA:** Venta total sin descontar las devoluciones de clientes.
$$\text{Venta bruta} = \text{VentaBrutaPresencial.unidades} + \text{VentaBrutaSINT.unidades} \\ + \text{VentaBrutaIPOD.unidades} + \text{VentaBrutaCLICK.unidades}$$
- **VENTA NETA:** Venta descontando las devoluciones de clientes. También se conoce como venta neta:
$$\text{Venta neta} = \text{VentaPresencial.unidades} + \text{VentaSint.unidades} + \text{VentaIPOD.unidades} + \text{VentaCLICK.unidades}$$

Apéndice B

Tablas

| Método | Medida | clústeres | 2 | 3 | 4 | 5 | 6 |
|-----------------|--------------|-----------|--------|--------|--------|--------|--------|
| Jerárquico | Conectividad | | 4.61 | 5.86 | 10.24 | 16.15 | 19.08 |
| | Dunn | | 0.88 | 1.77 | 0.28 | 0.07 | 0.07 |
| | Silueta | | 0.97 | 0.95 | 0.83 | 0.79 | 0.66 |
| <i>k</i> -means | Conectividad | | 4.61 | 5.86 | 10.24 | 85.26 | 163.94 |
| | Dunn | | 0.88 | 1.77 | 0.28 | 0.00 | 0.00 |
| | Silueta | | 0.97 | 0.95 | 0.83 | 0.52 | 0.42 |
| PAM | Conectividad | | 108.23 | 162.25 | 264.30 | 337.25 | 329.93 |
| | Dunn | | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | Silueta | | 0.44 | 0.42 | 0.30 | 0.33 | 0.33 |

Tabla B.1: Medidas de validación para métodos de agrupación con diferentes números de clústeres.

| Método | Medida | clústeres | 2 | 3 | 4 | 5 | 6 |
|------------|--------|-----------|------|------|------|------|------|
| Jerárquico | APN | | 0.00 | 0.00 | 0.00 | 0.01 | 0.04 |
| | AD | | 1.97 | 1.96 | 1.94 | 1.92 | 1.91 |
| | ADM | | 0.01 | 0.00 | 0.01 | 0.06 | 0.22 |
| | FOM | | 0.76 | 0.63 | 0.63 | 0.63 | 0.54 |
| K-means | APN | | 0.00 | 0.00 | 0.05 | 0.02 | 0.08 |
| | AD | | 1.97 | 1.96 | 1.94 | 1.53 | 1.28 |
| | ADM | | 0.00 | 0.00 | 0.17 | 0.04 | 0.27 |
| | FOM | | 0.73 | 0.63 | 0.63 | 0.48 | 0.48 |
| PAM | APN | | 0.21 | 0.10 | 0.15 | 0.17 | 0.19 |
| | AD | | 1.76 | 1.32 | 1.24 | 1.11 | 1.06 |
| | ADM | | 0.56 | 0.24 | 0.29 | 0.28 | 0.33 |
| | FOM | | 1.00 | 0.86 | 0.85 | 0.80 | 0.78 |

Tabla B.2: Medidas de validación para métodos de agrupación con diferentes números de clústeres.

Tabla B.3: Resultados del incremento del 13% en ventas en función del parámetro K=0.7 para la campaña Invierno 2023.

| Comprador | Envío | Incremento de envío % | Envío_C | Ventas Brutas | Devo | Ventas Netas | Exceso Envío | Destocaje Cliente |
|-----------|-------|-----------------------|---------|---------------|------|--------------|--------------|-------------------|
| BASIC | 4.94 | 1.00 | 4.27 | 3.40 | 0.63 | 3.05 | 1.89 | 0.36 |
| CIRCULAR | 6.17 | 1.00 | 5.21 | 4.32 | 0.49 | 4.04 | 2.12 | 0.28 |
| PUNTO | 1.68 | 1.00 | 1.51 | 1.14 | 0.14 | 1.05 | 0.62 | 0.08 |
| TRF | 2.45 | 1.00 | 2.03 | 1.63 | 0.30 | 1.45 | 1.00 | 0.17 |
| WOMAN | 2.30 | 1.00 | 1.98 | 1.82 | 0.39 | 1.60 | 0.71 | 0.22 |
| TOTAL | 17.54 | 1.00 | 15.00 | 12.31 | 1.96 | 11.20 | 6.35 | 1.12 |

Tabla B.4: Resultados del incremento del 13% en ventas en función del parámetro $K=0.4$ para la campaña Invierno 2023.

| Comprador | Envío | Incremento de envío % | Envío_C | Ventas Brutas | Devo | Ventas Netas | Exceso Envío | Destocaje Cliente |
|-----------|-------|-----------------------|---------|---------------|------|--------------|--------------|-------------------|
| BASIC | 4.94 | 1.00 | 4.27 | 3.40 | 0.63 | 3.05 | 1.89 | 0.36 |
| CIRCULAR | 6.17 | 1.00 | 5.21 | 4.32 | 0.49 | 4.04 | 2.12 | 0.28 |
| PUNTO | 2.36 | 1.41 | 2.19 | 1.60 | 0.20 | 1.48 | 0.88 | 0.12 |
| TRF | 2.45 | 1.00 | 2.03 | 1.63 | 0.30 | 1.45 | 1.00 | 0.17 |
| WOMAN | 3.64 | 1.58 | 3.32 | 2.88 | 0.62 | 2.53 | 1.12 | 0.35 |
| TOTAL | 19.56 | 1.12 | 17.02 | 13.83 | 2.24 | 12.55 | 7.01 | 1.28 |

Tabla B.5: Resultados del incremento del 13% en ventas en función del parámetro $K=0.2$ para la campaña Invierno 2023.

| Comprador | Envío | Incremento de envío % | Envío_C | Ventas Brutas | Devo | Ventas Netas | Exceso Envío | Destocaje Cliente |
|-----------|-------|-----------------------|---------|---------------|------|--------------|--------------|-------------------|
| BASIC | 4.94 | 1.00 | 4.27 | 3.40 | 0.63 | 3.05 | 1.89 | 0.36 |
| CIRCULAR | 6.17 | 1.00 | 5.21 | 4.32 | 0.49 | 4.04 | 2.12 | 0.28 |
| PUNTO | 2.40 | 1.43 | 2.23 | 1.62 | 0.21 | 1.51 | 0.89 | 0.12 |
| TRF | 2.45 | 1.00 | 2.03 | 1.63 | 0.30 | 1.45 | 1.00 | 0.17 |
| WOMAN | 3.87 | 1.68 | 3.54 | 3.06 | 0.66 | 2.68 | 1.18 | 0.38 |
| TOTAL | 19.82 | 1.13 | 17.28 | 14.03 | 2.29 | 12.73 | 7.09 | 1.30 |

Tabla B.6: Resultados del incremento del 13% en ventas en función del parámetro $K=0.08$ para la campaña Invierno 2023.

| Comprador | Envío | Incremento de envío % | Envío_C | Ventas Brutas | Devo | Ventas Netas | Exceso Envío | Destocaje Cliente |
|-----------|-------|-----------------------|---------|---------------|------|--------------|--------------|-------------------|
| BASIC | 4.94 | 1.00 | 4.27 | 3.40 | 0.63 | 3.05 | 1.89 | 0.36 |
| CIRCULAR | 6.17 | 1.00 | 5.21 | 4.32 | 0.49 | 4.04 | 2.12 | 0.28 |
| PUNTO | 2.20 | 1.31 | 2.03 | 1.49 | 0.19 | 1.38 | 0.89 | 0.11 |
| TRF | 2.45 | 1.00 | 2.03 | 1.63 | 0.30 | 1.45 | 1.00 | 0.17 |
| WOMAN | 4.07 | 1.77 | 3.74 | 3.22 | 0.69 | 2.82 | 1.18 | 0.40 |
| TOTAL | 19.82 | 1.13 | 17.28 | 14.06 | 2.30 | 12.74 | 7.09 | 1.31 |

Tabla B.7: Resultados del incremento del 13% en ventas en función del parámetro $K=0.05$ para la campaña Invierno 2023.

| Comprador | Envío | Incremento de envío % | Envío_C | Ventas Brutas | Devo | Ventas Netas | Exceso Envío | Destocaje Cliente |
|-----------|-------|-----------------------|---------|---------------|------|--------------|--------------|-------------------|
| BASIC | 4.94 | 1.00 | 4.27 | 3.40 | 0.63 | 3.05 | 1.89 | 0.36 |
| CIRCULAR | 6.17 | 1.00 | 5.21 | 4.32 | 0.49 | 4.04 | 2.12 | 0.28 |
| PUNTO | 1.99 | 1.19 | 1.83 | 1.35 | 0.17 | 1.25 | 0.74 | 0.10 |
| TRF | 2.45 | 1.00 | 2.03 | 1.63 | 0.30 | 1.45 | 1.00 | 0.17 |
| WOMAN | 4.27 | 1.85 | 3.95 | 3.38 | 0.73 | 2.96 | 1.31 | 0.42 |
| TOTAL | 19.82 | 1.12 | 17.28 | 14.08 | 2.32 | 12.76 | 7.07 | 1.32 |

Tabla B.8: Resultados del incremento del 13 % en ventas en función del parámetro $K=0.01$ para la campaña Invierno 2023.

| Comprador | Envío | Incremento de envío % | Envío_C | Ventas Brutas | Devo | Ventas Netas | Exceso Envío | Destocaje Cliente |
|-----------|-------|-----------------------|---------|---------------|------|--------------|--------------|-------------------|
| BASIC | 4.94 | 1.00 | 4.27 | 3.40 | 0.63 | 3.05 | 1.89 | 0.36 |
| CIRCULAR | 6.17 | 1.00 | 5.21 | 4.32 | 0.49 | 4.04 | 2.12 | 0.28 |
| PUNTO | 1.68 | 1.00 | 1.51 | 1.14 | 0.14 | 1.05 | 0.62 | 0.08 |
| TRF | 2.45 | 1.00 | 2.03 | 1.63 | 0.30 | 1.45 | 1.00 | 0.17 |
| WOMAN | 4.59 | 1.99 | 4.26 | 3.63 | 0.78 | 3.18 | 1.40 | 0.45 |
| TOTAL | 19.82 | 1.13 | 17.28 | 14.12 | 2.35 | 12.78 | 7.04 | 1.34 |

Tabla B.9: Resultados del incremento del 13 % en ventas en función del parámetro $K=1$ para la campaña Invierno 2024.

| Comprador | Envío | Incremento de envío % | Ventas Brutas | Devo | Ventas Netas | Exceso Envío | Destocaje Cliente |
|-----------|-------|-----------------------|---------------|------|--------------|--------------|-------------------|
| BASIC | 7.93 | 1.00 | 7.19 | 1.37 | 5.82 | 1.52 | 0.78 |
| CIRCULAR | 11.47 | 1.00 | 9.45 | 1.13 | 8.31 | 2.66 | 0.65 |
| PUNTO | 3.96 | 1.00 | 3.25 | 0.52 | 2.73 | 1.01 | 0.30 |
| TRF | 3.55 | 1.00 | 2.98 | 0.61 | 2.37 | 0.92 | 0.35 |
| WOMAN | 2.99 | 1.00 | 2.66 | 0.60 | 2.06 | 0.67 | 0.34 |
| TOTAL | 29.89 | 1.00 | 25.53 | 4.23 | 21.30 | 6.77 | 2.41 |

Tabla B.10: Resultados del incremento del 13% en ventas en función del parámetro $K=0.7$ para la campaña Invierno 2024.

| Comprador | Envío | Incremento de envío % | Ventas Brutas | Devo | Ventas Netas | Exceso Envío | Destocaje Cliente |
|-----------|-------|-----------------------|---------------|------|--------------|--------------|-------------------|
| BASIC | 7.93 | 1.00 | 7.19 | 1.37 | 5.82 | 1.52 | 0.78 |
| CIRCULAR | 11.47 | 1.00 | 9.45 | 1.13 | 8.31 | 2.66 | 0.65 |
| PUNTO | 3.96 | 1.00 | 3.25 | 0.52 | 2.73 | 1.01 | 0.30 |
| TRF | 3.55 | 1.00 | 2.98 | 0.61 | 2.37 | 0.92 | 0.35 |
| WOMAN | 4.13 | 1.38 | 3.67 | 0.82 | 2.85 | 0.92 | 0.47 |
| TOTAL | 31.03 | 1.04 | 26.54 | 4.45 | 22.09 | 7.03 | 2.54 |

Tabla B.11: Resultados del incremento del 13% en ventas en función del parámetro $K=0.4$ para la campaña Invierno 2024.

| Comprador | Envío | Incremento de envío % | Ventas Brutas | Devo | Ventas Netas | Exceso Envío | Destocaje Cliente |
|-----------|-------|-----------------------|---------------|------|--------------|--------------|-------------------|
| BASIC | 8.06 | 1.02 | 7.31 | 1.39 | 5.92 | 1.54 | 0.79 |
| CIRCULAR | 11.47 | 1.00 | 9.45 | 1.13 | 8.31 | 2.66 | 0.65 |
| PUNTO | 4.31 | 1.09 | 3.54 | 0.57 | 2.97 | 1.10 | 0.32 |
| TRF | 4.19 | 1.18 | 3.52 | 0.72 | 2.79 | 1.08 | 0.41 |
| WOMAN | 5.75 | 1.92 | 5.12 | 1.15 | 3.97 | 1.29 | 0.65 |
| TOTAL | 33.78 | 1.13 | 28.93 | 4.96 | 23.97 | 7.67 | 2.82 |

Tabla B.12: Resultados del incremento del 13% en ventas en función del parámetro $K=0.2$ para la campaña Invierno 2024.

| Comprador | Envío | Incremento de envío % | Ventas Brutas | Devo | Ventas Netas | Exceso Envío | Destocaje Cliente |
|-----------|-------|-----------------------|---------------|------|--------------|--------------|-------------------|
| BASIC | 8.28 | 1.05 | 7.51 | 1.43 | 6.08 | 1.58 | 0.81 |
| CIRCULAR | 11.47 | 1.00 | 9.45 | 1.13 | 8.31 | 2.66 | 0.65 |
| PUNTO | 4.21 | 1.06 | 3.45 | 0.55 | 2.90 | 1.07 | 0.31 |
| TRF | 4.07 | 1.15 | 3.41 | 0.70 | 2.71 | 1.05 | 0.40 |
| WOMAN | 5.76 | 1.93 | 5.12 | 1.15 | 3.98 | 1.29 | 0.65 |
| TOTAL | 33.78 | 1.13 | 28.95 | 4.96 | 23.99 | 7.66 | 2.83 |

Tabla B.13: Resultados del incremento del 13% en ventas en función del parámetro $K=0.08$ para la campaña Invierno 2024.

| Comprador | Envío | Incremento de envío % | Ventas Brutas | Devo | Ventas Netas | Exceso Envío | Destocaje Cliente |
|-----------|-------|-----------------------|---------------|------|--------------|--------------|-------------------|
| BASIC | 8.90 | 1.12 | 8.07 | 1.53 | 6.54 | 1.70 | 0.87 |
| CIRCULAR | 11.47 | 1.00 | 9.45 | 1.13 | 8.31 | 2.66 | 0.65 |
| PUNTO | 3.96 | 1.00 | 3.25 | 0.52 | 2.73 | 1.01 | 0.30 |
| TRF | 3.69 | 1.04 | 3.10 | 0.63 | 2.46 | 0.95 | 0.36 |
| WOMAN | 5.77 | 1.93 | 5.13 | 1.15 | 3.98 | 1.29 | 0.66 |
| TOTAL | 33.78 | 1.13 | 28.99 | 4.97 | 24.02 | 7.62 | 2.83 |

Tabla B.14: Resultados del incremento del 13% en ventas en función del parámetro $K=0.05$

| Comprador | Envío | Incremento de envío % | Ventas Brutas | Devo | Ventas Netas | Exceso Envío | Destocaje Cliente |
|-----------|-------|-----------------------|---------------|------|--------------|--------------|-------------------|
| BASIC | 9.23 | 1.17 | 8.37 | 1.59 | 6.78 | 1.77 | 0.91 |
| CIRCULAR | 11.47 | 1.00 | 9.45 | 1.13 | 8.31 | 2.66 | 0.65 |
| PUNTO | 3.96 | 1.00 | 3.25 | 0.52 | 2.73 | 1.01 | 0.30 |
| TRF | 3.55 | 1.00 | 2.98 | 0.61 | 2.37 | 0.92 | 0.35 |
| WOMAN | 5.57 | 1.86 | 4.96 | 1.11 | 3.85 | 1.25 | 0.63 |
| TOTAL | 33.78 | 1.13 | 29.01 | 4.97 | 24.04 | 7.60 | 2.83 |

Tabla B.15: Resultados del incremento del 13% en ventas en función del parámetro $K=0.01$ para la campaña Invierno 2024.

| Comprador | Envío | Incremento de envío % | Ventas Brutas | Devo | Ventas Netas | Exceso Envío | Destocaje Cliente |
|-----------|-------|-----------------------|---------------|------|--------------|--------------|-------------------|
| BASIC | 11.81 | 1.49 | 10.71 | 2.04 | 8.68 | 2.26 | 1.16 |
| CIRCULAR | 11.47 | 1.00 | 9.45 | 1.13 | 8.31 | 2.66 | 0.65 |
| PUNTO | 3.96 | 1.00 | 3.25 | 0.52 | 2.73 | 1.01 | 0.30 |
| TRF | 3.55 | 1.00 | 2.98 | 0.61 | 2.37 | 0.92 | 0.35 |
| WOMAN | 2.99 | 1.00 | 2.66 | 0.60 | 2.06 | 0.67 | 0.34 |
| TOTAL | 33.78 | 1.13 | 29.05 | 4.90 | 24.15 | 7.52 | 2.79 |

Bibliografía

- [1] A mathematical language program. AMPL IDE Version: 3.6.10.202208131811
- [2] R studio version 4.2.1
- [3] Pateiro-López, B. y Sánchez-Sellero, C. (2022) Análisis Multivariante. Dpto. de Estadística, Análisis Matemático y Optimización. Universidad de Santiago de Compostela. 25 mayo, 2022
- [4] Cuadras, C. (1996). Nuevos Métodos de Análisis Multivariante. CMC Edicions.
- [5] McQueen, J. (1967) *Some Methods for Classification and Analysis of Multivariate Observations*. Computer and Chemistry, 4, 257-272.
- [6] Brock, G., Pihur, V., Datta, S., & Datta, S. (2008). *clValid: An R Package for cluster Validation*. Journal of Statistical Software, 25(4), 1-22.
- [7] Hennig, C. (2007) Clusterwise assessment of cluster stability. *Computational Statistics and Data Analysis*, 52, 258-271.
- [8] Jaccard, P., 1901. *Distribution de la florine alpine dans la Bassin de Dranses et dans quelques regions voisines*. Bull. Soc. Vaud. Sci. Nat. 37, 241-272.
- [9] Kassambara, A. (2017). Practical Guide to Cluster Analysis in R: Unsupervised machine learning (Vol. 1). Sthda.
- [10] Casas-Méndez, B.V., Mosquera-Rodríguez, M.A. (2023) Programación lineal y Entera. Dpto. de Estadística, Análisis Matemático y Optimización. Universidad de Santiago de Compostela.
- [11] González-Díaz, J. (2023) Optimización Aplicada. Dpto. de Estadística, Análisis Matemático y Optimización. Universidad de Santiago de Compostela.