

UNIVERSIDAD SAN FRANCISCO DE QUITO USFQ

Colegio de Ciencias e Ingenierías

**Análisis y Optimización de la Gestión de Retornos dentro de la
Cadena de Suministro**

Said Alejandro Berrezueta Pumacuro

Sofía Valentina Vargas Díaz

Joseph Adrian Galarza Ochoa

Ingeniería Industrial

Trabajo de fin de carrera presentado como requisito
para la obtención del título de

INGENIERO INDUSTRIAL

Quito, 08 de mayo de 2025

UNIVERSIDAD SAN FRANCISCO DE QUITO USFQ

Colegio de Ciencias e Ingenierías

HOJA DE CALIFICACIÓN DE TRABAJO DE FIN DE CARRERA

**Análisis y Optimización de la Gestión de Retornos dentro de la
Cadena de Suministro**

**Said Alejandro Berrezueta Pumacuro
Sofía Valentina Vargas Díaz
Joseph Adrian Galarza Ochoa**

Nombre del profesor, Título académico

Carlos Suárez, PhD

Quito, 08 de mayo de 2025

© DERECHOS DE AUTOR

Por medio del presente documento certifico que he leído todas las Políticas y Manuales de la Universidad San Francisco de Quito USFQ, incluyendo la Política de Propiedad Intelectual USFQ, y estoy de acuerdo con su contenido, por lo que los derechos de propiedad intelectual del presente trabajo quedan sujetos a lo dispuesto en esas Políticas.

Asimismo, autorizo a la USFQ para que realice la digitalización y publicación de este trabajo en el repositorio virtual, de conformidad a lo dispuesto en la Ley Orgánica de Educación Superior del Ecuador.

Nombres y apellidos: Said Alejandro Berrezueta Pumacuro

Código: 00321711

Cédula de identidad: 1754160883

Nombres y apellidos: Sofia Valentina Vargas Díaz

Código: 00217204

Cédula de identidad: 1724564388

Nombres y apellidos: Joseph Adrian Galarza Ochoa

Código: 00324551

Cédula de identidad: 1750837880

Lugar y fecha: Quito, 08 de mayo de 2025

ACLARACIÓN PARA PUBLICACIÓN

Nota: El presente trabajo, en su totalidad o cualquiera de sus partes, no debe ser considerado como una publicación, incluso a pesar de estar disponible sin restricciones a través de un repositorio institucional. Esta declaración se alinea con las prácticas y recomendaciones presentadas por el Committee on Publication Ethics COPE descritas por Barbour et al. (2017) Discussion document on best practice for issues around theses publishing, disponible en <http://bit.ly/COPETheses>.

UNPUBLISHED DOCUMENT

Note: The following capstone project is available through Universidad San Francisco de Quito USFQ institutional repository. Nonetheless, this project – in whole or in part – should not be considered a publication. This statement follows the recommendations presented by the Committee on Publication Ethics COPE described by Barbour et al. (2017) Discussion document on best practice for issues around theses publishing available on <http://bit.ly/COPETheses>.

Resumen

La logística inversa es una etapa crítica en la cadena de suministro, encargada de planificar, transportar y gestionar devoluciones. En los últimos años, la importancia de la logística inversa ha crecido notablemente debido al creciente interés de los consumidores por adquirir productos que puedan ser devueltos, a la presión social y normativa que obliga a las empresas a responsabilizarse por la correcta gestión de sus desechos, a las oportunidades económicas en mercados secundarios o de descuento y a las limitaciones de espacio en los vertederos (Hurtado, 2019).

Este proyecto se desarrolla en el área de Logística Inversa de Provefarma Ecuador, donde se identificaron oportunidades de mejora en los procesos documentados por la empresa, así como limitaciones en la trazabilidad documental y la gestión manual de datos. Para abordar estos problemas, se planteó un doble enfoque: por un lado, el análisis y validación de los procesos operativos levantados por la empresa, y por otro, la automatización de la digitalización de documentos empleando un código de Python y el desarrollo de un modelo de visualización de datos en la herramienta Power BI. Estas soluciones buscan optimizar los tiempos de procesamiento, mejorar la precisión operativa y fortalecer la toma de decisiones estratégicas.

Palabras clave: Logística Inversa, optimización de procesos, Python, Power BI, automatización.

Abstract

Reverse logistics is a critical stage in the supply chain, responsible for planning, transporting, and managing returns. In recent years, the importance of reverse logistics has grown significantly due to the increasing consumer interest in purchasing products that can be returned, social and regulatory pressures that compel companies to take responsibility for the proper management of their waste, economic opportunities in secondary or discount markets, and the limitations of space in landfills (Hurtado, 2019).

This project is developed in the Reverse Logistics area of Provefarma Ecuador, where opportunities for improvement were identified in the processes documented by the company, as well as limitations in document traceability and manual data management. To address these issues, a dual approach was proposed: on one hand, the analysis and validation of the operational processes documented by the company, and on the other, the automation of document digitization using Python code and the development of a data visualization model in the Power BI tool. These solutions aim to optimize processing times, improve operational accuracy, and strengthen strategic decision-making.

Key words: Reverse Logistics, process optimization, Python, Power BI, automation.

Tabla de contenido

1. Introducción	9
2. Revisión literaria	10
3. Metodología	12
4. Definir	14
4.1. Herramientas e información disponible.....	14
4.2. Descripción de la situación actual	16
4.3. Identificación de oportunidades de mejora.....	18
5. Medir	21
5.1. Definición y medición de variables	21
5.2. Resumen de resultados	25
6. Validación de procesos	26
6.1. Analizar	26
6.2. Mejorar	26
6.3. Resultados y discusión.....	27
7. Digitalización de documentos	28
7.1. Analizar	28
7.2. Mejorar	29
7.3. Resultados y discusión.....	31
8. Gestión y visualización de datos	32
8.1. Analizar	33
8.2. Mejorar	35
8.3. Resultados y discusión.....	37
9. Conclusiones	39
10. Limitaciones y recomendaciones.....	40
11. Referencias bibliográficas	42
12. Anexos	46

Índice de Figuras

Figura 1. Etapas identificadas y procesos documentados.....	17
---	----

Índice de Tablas

Tabla 1. Variables de la validación de procesos documentados	22
Tabla 2. Variables de la digitalización de documentos	23
Tabla 3. Variables de la gestión y visualización de datos	24
Tabla 4. Variables de la gestión y visualización de datos	25
Tabla 5. Resultados de la prueba de desempeño	31

1. Introducción

El presente proyecto se desarrolla en la bodega de Logística Inversa (LI) del Centro de Distribución de Provefarma, ubicado en Sangolquí, Pichincha, Ecuador. Provefarma forma parte de la Corporación GPF, empresa que opera a nivel nacional en Ecuador y pertenece a FEMSA Salud, un conglomerado multinacional de origen mexicano. Esta bodega se encarga de recibir y gestionar las devoluciones de más de mil puntos de venta a nivel nacional. Entre estos puntos de venta se incluyen franquicias de la Corporación GPF, como Fybeca, SanaSana y OkiDoki, además de otras farmacias independientes (FEMSA Salud Ecuador, 2023).

El área de Logística Inversa es de gran relevancia debido al alto volumen de mercancía y la cantidad de dinero que maneja. En 2024, la bodega procesó cinco millones de unidades, equivalentes a más de veinte millones de dólares, lo que resalta la magnitud de sus operaciones y la necesidad de optimizar su gestión. Además, esta área se encarga del procesamiento de medicamentos, los cuales están sujetos a fiscalización por parte de la Agencia Nacional de Regulación, Control y Vigilancia Sanitaria (ARCSA, 2025).

A pesar de su importancia, Logística Inversa recibe poca prioridad dentro de la empresa y su cadena de suministro, debido a la falta de visibilidad sobre los problemas y desafíos específicos que enfrenta la operación. Esta falta de visibilidad lleva a plantear la siguiente pregunta de investigación: *¿Cuáles son los problemas existentes en el área de Logística Inversa y cómo pueden optimizarse para mejorar los tiempos y la gestión de datos?*

Para abordar esta problemática se ha seleccionado la metodología DMAIC, ya que proporciona un enfoque estructurado para el desarrollo de proyectos de optimización y

mejora continua (Nandakumar et al., 2020). Además, esta metodología facilita la recopilación de información relevante, lo que es crucial para determinar las causas de los problemas y proponer soluciones efectivas.

Dentro de este contexto, el presente documento aborda de manera ordenada una revisión literaria de casos de estudio dentro de campos similares, la metodología utilizada, y el desarrollo de cada etapa. Adicional a esto, se presentan los resultados obtenidos y su correspondiente discusión, además de una sección posterior de conclusiones que abarca todo el proyecto. Por último, se incluye una sección final que presenta limitaciones que se hicieron presentes a lo largo del desarrollo de este proyecto y recomendaciones que permitan afrontarlas.

2. Revisión literaria

La presente revisión literaria se basa en el análisis de casos de estudio dentro de contextos similares para conocer las metodologías y herramientas tecnológicas aplicadas a la mejora de procesos organizacionales, enfocándose en el análisis de datos, automatización documental y visualización de flujos de trabajo. Dentro de la industria farmacéutica, la metodología DMAIC presenta varios casos de éxito, como el caso de estudio presentado por Vanegas (2018), en el cual se destaca la aplicación de la metodología DMAIC para mantener un formato estructurado y disciplinado a lo largo de la ejecución del proyecto. Además, evidencia el impacto del uso de las herramientas asociadas a esta metodología para identificar oportunidades de mejora y proponer acciones que optimicen la situación actual.

Dentro del área de logística inversa, casos como el que presenta Peralta et al. (2018), evidencian cómo el uso de estrategias alineadas a la metodología DMAIC permitió mejorar

la gestión del área con un mayor control sobre las operaciones y mejor manejo de la información generada. Casos como los mencionados demuestran un fuerte sustento para el uso de la metodología DMAIC para el presente estudio, siendo que su enfoque estructurado permite un adecuado desarrollo del trabajo.

Por su parte, los diagramas de flujo son de gran ayuda para detectar fallos estructurales y mejorar procesos. Zhu et al. (2024) analizan los diagramas de flujo para identificar cuellos de botella, redundancias y otros problemas que afectan la eficiencia operativa, lo que permite rediseñar procesos de forma más efectiva.

En cuanto a la digitalización de documentos, el uso de tecnología OCR con Tesseract, mediante Python, ha facilitado la extracción automática de texto en documentos escaneados. Según Robby et al. (2019), su implementación ha permitido clasificar y almacenar archivos con mínima intervención humana, optimizando procesos como el escaneo rápido y la organización automatizada de documentos.

Power BI ha sido reconocido como una herramienta eficaz para la visualización de datos e informes, permitiendo integrar en un mismo dashboard los indicadores clave del proceso. Nickell et al. (2023) muestran que, al integrarse con Power BI, se facilita el análisis de grandes volúmenes de datos y la detección de patrones atípicos en información financiera, fortaleciendo el control interno y mejorando la toma de decisiones basada en datos confiables.

Esta revisión literaria respalda el enfoque metodológico adoptado en el presente proyecto, al evidenciar que herramientas como los diagramas de flujo, el OCR con Tesseract y Power BI han sido efectivamente aplicadas en otras organizaciones para optimizar

procesos, automatizar tareas y mejorar la toma de decisiones. Por tanto, se justifica su selección e implementación como soluciones viables y alineadas con prácticas validadas en entornos empresariales reales.

3. Metodología

En la industria, la metodología DMAIC se ha consolidado como un pilar esencial en los proyectos de mejora continua. Su aplicación resulta clave para identificar problemas, analizar su impacto y diseñar soluciones eficaces en entornos donde la complejidad de los procesos exige un enfoque metódico. Esta metodología se compone de cinco etapas interdependientes: Definir, Medir, Analizar, Mejorar y Controlar, cada una diseñada para garantizar que las mejoras implementadas sean sostenibles (Nandakumar et al., 2020).

En la etapa Definir, el objetivo es identificar los problemas dentro de la operación y sus causas. En este contexto, la revisión de diagramas de flujo resulta una herramienta clave, ya que permite representar de manera clara y ordenada cada actividad del proceso, facilitando la detección de cuellos de botella, redundancias, actividades innecesarias y puntos críticos (Kulkarni, 2024). Además, estos diagramas ayudan a establecer un lenguaje común entre las áreas involucradas, lo que favorece la comunicación y colaboración.

En la etapa Medir, se recopilan datos de variables clave del proceso, con el objetivo de confirmar que el problema identificado realmente existe y comprender su magnitud (Vidal et al., 2018). Esta validación permite establecer una línea base del desempeño actual, la cual servirá como punto de referencia para evaluar los efectos de las mejoras implementadas más adelante.

En la etapa Analizar, se revisan los datos recopilados con el fin de comprender las causas raíz de los problemas identificados. Según López et al. (2021), a partir de esta comprensión se exploran diversas soluciones, evaluando tanto su viabilidad como su impacto potencial, para finalmente priorizar aquellas que sean factibles de implementar y que ofrezcan el mayor beneficio al proceso.

En la etapa Mejorar, se aplican las soluciones propuestas para abordar las causas raíz de los problemas, lo que puede incluir ajustar procesos, modificar procedimientos operativos u optimizar las herramientas o tecnologías utilizadas. Se realizan pruebas piloto para validar la efectividad de las mejoras antes de su aplicación definitiva y se capacita al personal en los nuevos procedimientos para asegurar su correcta ejecución (Pérez & Caballero, 2024).

En la etapa Controlar, se implementan mecanismos de seguimiento para asegurar que las mejoras aplicadas se mantengan en el tiempo. Esta fase busca prevenir que los problemas reaparezcan y garantizar la estabilidad del proceso, por lo que se establecen indicadores de control, se definen responsables y se diseñan planes de monitoreo continuo (Utama & Abirfatin, 2023).

El presente proyecto se plantea como un estudio del área de LI que emplea la metodología DMAIC para identificar oportunidades de mejora e implementar soluciones viables en la etapa Mejorar. No obstante, la etapa Controlar no se desarrolla, ya que la implementación de mecanismos de control requiere conocimientos técnicos avanzados en tecnología de Reconocimiento Óptico de Caracteres (OCR), programación y bases de datos. Estas condiciones técnicas requieren una exhaustiva validación operativa para garantizar la sostenibilidad de las mejoras, ya que la integración de herramientas de control debe ser validada previamente para asegurar resultados sostenibles y alineados con el rendimiento

esperado, tal como se documenta en aplicaciones de DMAIC (Indrawati & Ridwansyah, 2015).

Con una metodología clara y adecuada definida para el desarrollo de este proyecto, y con numerosos casos de éxito dentro del campo de estudio, se procede a presentar el proyecto realizado en el área de Logística Inversa de Provefarma Ecuador.

4. Definir

En la etapa Definir, el proyecto se orienta a identificar con claridad los problemas que afectan al área de Logística Inversa, esta necesidad surge ante la falta de un diagnóstico preciso por parte de la empresa, lo que ha dificultado reconocer los desafíos específicos que atraviesan. Con base en ello, las siguientes secciones dentro de esta etapa desarrollan un análisis detallado de la situación actual, que servirá como punto de partida para alcanzar el objetivo general del proyecto: *determinar las causas de las ineficiencias en los procesos de Logística Inversa, a partir del análisis de la operación actual, para optimizar tiempos y mejorar la gestión de datos.*

4.1. Herramientas e información disponible

A continuación, se detallarán las herramientas empleadas y la información proporcionada por la empresa para realizar el diagnóstico de Logística Inversa. Estas herramientas fueron fundamentales para comprender el funcionamiento del sistema e identificar los puntos críticos. Además, la información proporcionada por la empresa permitió una integración más fluida del equipo en los procesos, lo que facilitó la identificación directa de oportunidades de mejora.

- **Procesos documentados:** La empresa cuenta actualmente con doce procesos documentados en el área de Logística Inversa, los cuales fueron levantados previamente por el área de Procesos. En la Figura 1 se presenta cada uno de estos procesos (1 al 12) y su ubicación dentro del flujo operativo de la bodega. Esta documentación fue consultada de forma periódica durante el desarrollo del proyecto y revisada junto con el coordinador del área para asegurar una comprensión adecuada del sistema. Además, sirvió como base para la capacitación del equipo antes de participar directamente en la gestión de devoluciones con los operadores.
- **Gemba Walk:** herramienta de la gestión lean que consiste en caminar por el lugar de trabajo para observar directamente las actividades y detectar oportunidades de mejora (Taylor, 2021). Esta herramienta permite identificar discrepancias entre la documentación y la ejecución real de los procesos.
- **Entrevistas:** realizadas a operadores y al coordinador del área para conocer sus experiencias y las limitaciones que perciben para realizar un trabajo eficiente. Estas entrevistas se complementaron con inducciones en otras áreas de la empresa para comprender el flujo de información, interdependencias y dinámica de trabajo.
- **5W-1H:** Esta herramienta se utilizó como guía para estructurar las entrevistas realizadas durante el diagnóstico del área. A través de las preguntas ¿Qué?, ¿Quién?, ¿Cuándo?, ¿Dónde?, ¿Por qué? y ¿Cómo?, se logró desglosar de manera detallada los problemas y situaciones observadas. Cada pregunta cumple una función específica: *¿Qué?* ayuda a identificar la actividad o problema específico, *¿Quién?* señala a los responsables o involucrados, *¿Cuándo?* permite entender la frecuencia o momento en que ocurre, *¿Dónde?* ubica el punto dentro del proceso, *¿Por qué?* busca entender

la causa del evento y *¿Cómo?* detalla la forma en que se lleva a cabo la actividad. Gracias a esta estructura, se facilita la recolección ordenada de información y se garantiza que ningún aspecto relevante quede fuera del análisis (Jinks, 2019).

4.2. Descripción de la situación actual

Al analizar la información disponible y emplear las herramientas mencionadas, se obtuvo información concreta sobre la operación, que se ha dividido en seis etapas consecutivas (de la A a la F), las cuales se presentan a continuación y en la Figura 1. Estas etapas contienen los doce procesos previamente documentados por la empresa, como se mencionó anteriormente.

- A. Generar devolución en el punto de venta (PDV):** en cada PDV se generan las devoluciones que deben ser enviadas a Logística Inversa. Los productos se clasifican según su tipo (farmacia o consumo) y la causal de devolución: vencimiento, recall (retiro solicitado por proveedor o por la Agencia Nacional de Regulación, Control y Vigilancia Sanitaria - ARCSA), sobrestock (unidades excedentes) o promociones (productos de promociones vencidas). Se arman las cajas y se etiquetan con la causal, número único de la caja, tipo de producto, fecha y nombre del PDV.
- B. Transportar la mercadería:** el servicio de transporte está tercerizado a través de tres empresas. Las devoluciones enviadas desde Guayaquil y zonas aledañas pasan por un cross-dock antes de ser trasladadas a Sangolquí. Desde otros puntos del país, las devoluciones se envían directamente a la bodega.

- C. Receptar las devoluciones:** al llegar a la bodega, los operadores de Logística Inversa verifican la cantidad de cajas contra el documento físico enviado por el PDV, revisan el etiquetado y las clasifican en pallets según su causal de devolución.
- D. Abrir las cajas:** los operadores ingresan el número único de la caja en el sistema, lo que les permite visualizar la información de los productos que contiene. Posteriormente, abren las cajas para verificar el estado de los productos, la fecha de caducidad y la cantidad. En caso de discrepancias entre el sistema y el contenido físico, los operadores corrigen la información en el sistema antes de continuar.
- E. Clasificar y perchar:** los productos se clasifican por proveedor y por destino final. En esta etapa se realiza una nueva verificación del estado, la fecha de caducidad y la cantidad antes de reempacar la mercadería en cajas según su disposición final.
- F. Disponer la finalidad de los productos:** las cajas armadas se envían a su destino asignado, que puede ser: destrucción (incineración), donación, envío con un gestor ambiental, eventos promocionales de la empresa o devolución a proveedor.

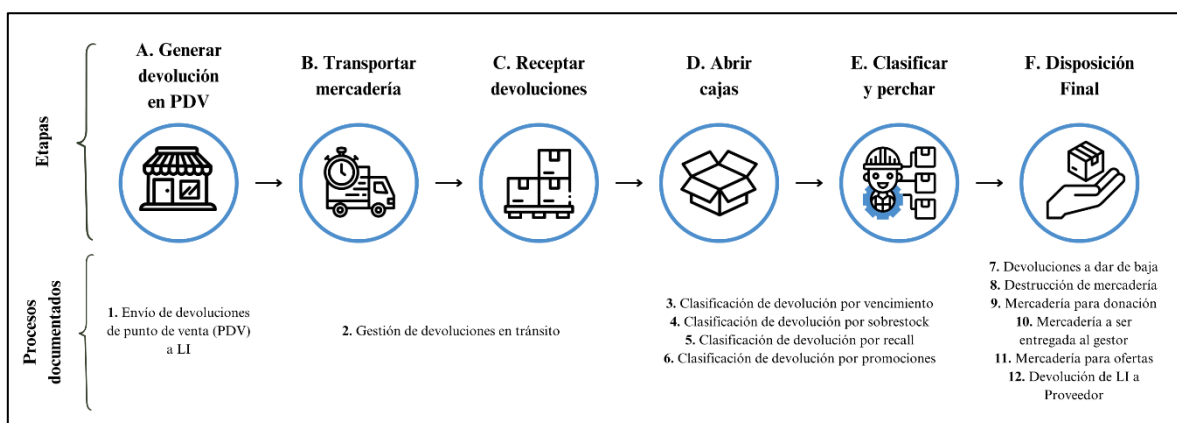


Figura 1. Etapas identificadas y procesos documentados

4.3. Identificación de oportunidades de mejora

Al comparar los procesos documentados con la operación real, se identificaron inconsistencias en la diagramación y la omisión de actividades que sí se llevan a cabo en la práctica. Esto evidencia una brecha entre lo planificado y lo ejecutado, por lo que se plantea como primer objetivo específico: *analizar los procesos documentados de Logística Inversa para identificar discrepancias entre la documentación y la ejecución real*. Además, los colaboradores del área muestran resistencia a admitir fallas y consideran que los procesos son simplemente una cuestión de cumplimiento, restando importancia y relevancia a los mismos. Adicionalmente, con la información levantada sobre la situación actual, se identificaron inconvenientes particulares existentes dentro de las seis etapas definidas anteriormente, los cuales forman parte del diagnóstico que se presenta a continuación.

A. Generar devolución en el punto de venta (PDV): en esta etapa se identificó que tanto LI como PDV desconocen completamente los procesos del otro, esto se debe a que no existe una comunicación activa entre ambos. Además, desde LI se tenía la percepción de que todos los problemas se originaban en PDV, por lo que se realizó una visita a la Fybeca de Cumbayá, Quito, ubicada en la Av. Oswaldo Guayasamín, para conocer su proceso de devoluciones y sus necesidades. De esta visita se concluyó que varios de los inconvenientes que se atribuyen a los PDV se relacionan con el hecho de que sus procesos también se gestionan de manera manual. Además, según lo conversado con el coordinador de este punto de venta, sería valioso contar con información de los lotes y fechas de caducidad de los productos que reciben, ya que actualmente el sistema solo muestra la cantidad enviada desde el centro de logística avanzada (CEDIS). Como el PDV no cuenta con la información de lotes dentro del

sistema, cada operador debe realizar revisiones periódicas de cada uno de los productos, lo que consume varias horas de trabajo semanalmente y es una actividad propensa a errores.

- B. Transportar mercadería:** esta etapa es gestionada por proveedores externos y presenta limitaciones para el levantamiento de información debido a la falta de visibilidad directa y a la reducida posibilidad de intervención en sus procesos. Por lo que no se presenta un diagnóstico específico de esta etapa.
- C. Receptar las devoluciones:** en esta etapa se manejan varios documentos físicos para dar seguimiento a las cajas enviadas desde los PDV. Estos documentos son llenados por los PDV, transportistas, y operadores de LI; posteriormente, deben ser digitalizados para mantener un respaldo digital dentro de LI, lo cual implica una actividad manual y repetitiva.
- D. Abrir las cajas:** las devoluciones aparecen en el sistema 48 horas después de ser generadas por el PDV ya que es el tiempo que tarda el sistema en actualizarse, por lo que, si los operadores ingresan el número único de la caja y no aparece, deben apartar estas cajas hasta que se actualice el sistema. Además, en esta etapa se presentan novedades con frecuencia, como productos en mal estado o números únicos de caja incorrectos. Los operadores gestionan estas novedades con regularidad, pero no mantienen un registro de estas, lo que complica la identificación y cuantificación de problemas comunes.
- E. Clasificar y perchar:** se realizan actividades repetitivas, como revisiones múltiples del estado, fecha de caducidad y cantidad de los productos, lo que consume más tiempo de los operadores, especialmente cuando los lotes son grandes.

F. Disponer la finalidad de los productos: cuando se devuelve mercadería a los proveedores, los operadores deben hacer firmar a los transportistas las actas de entrega. Estas actas se firman en físico, luego deben ser escaneadas y almacenadas digitalmente como respaldo. A partir de octubre de 2024, esta actividad comenzó a generar una acumulación significativa de documentos pendientes por digitalizar. Esta situación se relaciona con el aumento considerable en el volumen de devoluciones que se gestionan en los últimos meses del año, siendo noviembre uno de los periodos de mayor carga operativa (Coordinador de LI, comunicación personal, febrero de 2025).

Una de las oportunidades de mejora más relevantes identificadas en el proceso de Logística Inversa es la digitalización de documentos, ya que se realiza en las etapas C (*Receptar las devoluciones*) y F (*Disponer la finalidad de los productos*). En ambas se generan actas físicas que deben ser escaneadas y archivadas digitalmente como respaldo. Esta tarea, a cargo de un solo operador, conlleva una acumulación de documentos pendientes por digitalizar, especialmente en épocas de alta carga operativa. Dado que en la etapa F se devuelve mercadería a los proveedores, la digitalización resulta crítica para contar con respaldo inmediato en caso de que estos presenten reclamos, por lo que se plantea como segundo objetivo específico *automatizar el flujo de digitalización y clasificación de documentos firmados mediante un sistema basado en tecnologías OCR y NLP, con el fin de reducir las tareas manuales y mejorar la eficiencia del proceso.*

Finalmente, el sistema genera diariamente una gran cantidad de datos relacionados con las devoluciones realizadas en los puntos de venta, la disposición final de los productos y la información de los proveedores y puntos de venta involucrados. Toda esta información

es almacenada por el sistema y ha sido utilizada para la creación de tres dashboard en un archivo de Excel, el cual es consultado diariamente por el coordinador del área. Sin embargo, se han identificado varias limitaciones en la importación y visualización de los datos, como el tiempo prolongado para acceder a la información actualizada, el peso excesivo del archivo y problemas en el diseño del dashboard. Además, se observa una falta de estructura en los datos y ausencia de mantenimiento, ya que el archivo fue creado en 2021 y no ha tenido mantenimiento desde entonces. Estas condiciones reducen la eficiencia del análisis y muestran la necesidad de actualizar el sistema actual. Por estas razones, se plantea como tercer objetivo específico *desarrollar un sistema de visualización en Power BI que permita un análisis más preciso y facilite el manejo de los datos operativos de la bodega de LI.*

5. Medir

Una vez definidas las oportunidades de mejora y los objetivos específicos del proyecto, resulta fundamental establecer variables que permitan validar dichos objetivos y establecer una línea base de la situación actual. Esta línea base servirá como punto de comparación para evaluar el impacto de las mejoras implementadas en etapas posteriores. Las siguientes secciones dentro de esta etapa presentan las variables seleccionadas en relación con cada objetivo específico y un resumen de los resultados obtenidos al realizar la medición de cada una.

5.1. Definición y medición de variables

En relación con el primer objetivo específico, enfocado en la validación de los procesos documentados, se definieron las variables que se detallan en la Tabla 1. Estas fueron

cuantificadas por el equipo al contrastar directamente la documentación con la operación real, señalando dentro de los diagramas los errores de lógica que generaban confusión y marcando los procesos que omitían actividades clave que sí se llevan a cabo en la bodega.

Variables	Definición	Método de medición	Justificación
Cantidad de errores en la diagramación	Número de errores de lógica encontrados dentro de las gráficas de procesos	Revisión de la documentación y cuantificación de errores	Los errores de diagramación generan malentendidos en la ejecución y análisis del proceso
Cantidad de procesos incompletos	Número de procesos que no registran actividades que se ejecutan	Revisión de la documentación y ejecución del proceso y cuantificación de actividades no registradas	Un proceso con actividades faltantes no demuestra la realidad de la operación

Tabla 1. Variables de la validación de procesos documentados

Ahora bien, en la Tabla 2 se presentan las variables definidas para el segundo objetivo específico, relacionado con la actividad de digitalización de documentos. Para su cuantificación, se midió el tiempo que un operador tarda en ejecutar esta actividad, la cual se dividió en dos etapas: el escaneo del documento y la clasificación del archivo digital. La recolección de datos se realizó con el operador que usualmente desempeña esta actividad, quien no la ejecuta en horarios ni días específicos, lo que limitó el número de observaciones posibles. Debido a estas restricciones, se aplicó un muestreo por conveniencia, registrando 20 mediciones distribuidas en dos días (ver Anexo 1). A partir de estos datos, se estimó un tiempo total promedio de 124.85 segundos por documento, de los cuales 34.85 segundos (desviación estándar: 1.57 segundos) corresponden a la etapa de escaneo, y 90 segundos (desviación estándar: 2.32 segundos) a la etapa de clasificación del archivo digital. Por otro

lado, para cuantificar el número de documentos pendientes por digitalizar, se realizó un conteo manual y se presenta el número aproximado.

Variables	Definición	Método de medición	Justificación
Tiempo de digitalización de documentos	Tiempo en segundos que toma digitalizar y clasificar un documento	Cronometraje manual desde que el operario toma el documento para escanearlo hasta que completa su organización	Evaluar el tiempo que toma esta actividad permite justificar su impacto al resto del proceso
Cantidad de documentos pendientes por digitalizar	Número de documentos físicos que aún no han sido digitalizados	Conteo de los documentos pendientes	Permite monitorear la carga de trabajo pendiente

Tabla 2. Variables de la digitalización de documentos

Finalmente, en la Tabla 3 se presentan las variables definidas para el tercer objetivo específico, relacionado con la gestión y visualización de datos. Se seleccionaron variables que impactan directamente en la disponibilidad, calidad y velocidad de acceso a la información (Ghiglione, 2021). En primer lugar, se midió la cantidad de fuentes de datos que alimentan el dashboard actual identificando las conexiones activas en Power Query, herramienta de Microsoft Excel utilizada para importar, transformar y combinar datos de múltiples fuentes de forma automatizada (Ferrão, 2020). La cantidad de preguntas que puede responder los dashboard se evaluó contando las consultas directas que era capaz de resolver sin necesidad de procesamiento manual adicional, las cuales se detallan a continuación:

- ¿Cuántas líneas han ingresado y han sido devueltas al proveedor?
- ¿Cuántas unidades han ingresado y han sido devueltas al proveedor?
- ¿Cuál es el valor monetario de la mercadería que ha ingresado y ha sido devuelta al proveedor?

- ¿Cuántos envíos se han recibido desde los puntos de venta (PDV)?

El Índice de Peso del Archivo (IPA) se calculó dividiendo el peso actual del archivo Excel (432 MB) entre un peso ideal de 100 MB, considerado adecuado para un desempeño eficiente (Microsoft, 2024). Finalmente, el Tiempo de Acceso a la Información (TAI) se midió cronometrando el tiempo desde la apertura del archivo hasta la visualización completa del dashboard. Estas mediciones se realizaron directamente sobre el único archivo operativo disponible, por lo que no fue necesario aplicar muestreo, ya que se analizó la totalidad de la población de datos en condiciones controladas, conforme a los criterios recomendados en proyectos de mejora de procesos (Montgomery, 2020).

Variables	Definición	Método de medición	Justificación
Cantidad de fuentes de datos	Número de orígenes de datos que alimentan el dashboard	Revisión de fuentes conectadas	Entender la complejidad del archivo e identificar el flujo de datos
Cantidad de preguntas a las que responde el dashboard	Número de preguntas clave que el dashboard es capaz de responder de manera directa	Revisión funcional del dashboard	Refleja la utilidad del dashboard actual
Índice del Peso del Archivo (IPA)	Relación entre el peso actual del archivo Excel y un peso ideal de 100 MB	$IPA = \text{Peso actual del archivo (MB)} \div 100 \text{ MB}$	Mantener el peso bajo mejora la velocidad de acceso y la estabilidad del sistema.
Tiempo de Acceso a la Información (TAI)	Tiempo en minutos que tarda en estar disponible todo el dashboard	Cronometrar desde el momento de cargar datos hasta que ya estén visualizados.	Un acceso rápido permite decisiones ágiles. Tiempos de carga elevados afectan la productividad y la capacidad de reacción

Tabla 3. Variables de la gestión y visualización de datos

5.2. Resumen de resultados

Ahora bien, en la Tabla 4 se presentan los resultados obtenidos de la medición de cada variable, los cuales serán analizados a profundidad en la siguiente etapa. Estos resultados son el reflejo de la situación actual y se utilizarán para analizar el impacto de las mejoras propuestas.

Objetivo de enfoque	Variables y Resultados	
Validación de procesos	Cantidad de errores en la diagramación: 3	Cantidad de procesos incompletos: 4
Digitalización de documentos	Tiempo de digitalización de documentos: 124.85 seg	Cantidad de documentos pendientes por digitalizar: +700
Gestión y visualización de datos	Cantidad de fuentes de datos: 6	Cantidad de preguntas a las que responde el dashboard: 4
	Índice del Peso del Archivo (IPA): 4.32	Tiempo de Acceso a la Información (TAI): 40 min

Tabla 4. Variables de la gestión y visualización de datos

Con el fin de facilitar la comprensión y lectura del presente proyecto, a partir de este punto se desarrollarán de manera conjunta las dos etapas siguientes de la metodología DMAIC, Analizar y Mejorar, para cada uno de los objetivos específicos planteados, ya que como se mencionó anteriormente, la etapa Controlar queda fuera del alcance definido para este proyecto. Adicionalmente, para cada objetivo específico se presentan los resultados obtenidos y su respectiva discusión.

La siguiente sección presenta el desarrollo completo para el primer objetivo específico.

6. Validación de procesos

Con los datos obtenidos dentro de la anterior etapa, se procede a un análisis más detallado de cada variable y las causas de los problemas detectados. Esto corresponde a la etapa Analizar y tiene como finalidad identificar oportunidades de mejora dentro del sistema evaluado, permitiendo proponer acciones que optimicen su desempeño. Posterior a esto, se presenta la etapa Mejorar en la que se detallan las acciones tomadas para optimar la situación que refleja la etapa de Analizar. Por último, se desarrollan los resultados y discusión.

6.1. Analizar

- **Cantidad de errores en la diagramación (3):** los errores encontrados dentro de los diagramas corresponden a un error de lógica, específicamente en el flujo de actividades. Se diagraman procesos de manera consecutiva, cuando realmente corresponden a procesos mutuamente excluyentes.
- **Cantidad de procesos incompletos (4):** los procesos de clasificación de mercadería no documentan lo que ocurre cuando el número de documento de la caja está incorrecto, a pesar de que es una novedad que se presenta con frecuencia.

6.2. Mejorar

Considerando la información obtenida durante la etapa Analizar, se procede con la etapa Mejorar. Durante esta etapa se corrigieron errores de diagramación, ya que actividades que debían ser excluyentes se encontraban representadas como consecutivas. La corrección de este flujo se puede evidenciar en el Anexo 2 resaltado mediante una zona roja. La imagen presentada en el Anexo 2 corresponde a un extracto del proceso de clasificación de devoluciones por recall. Debido al acuerdo de confidencialidad no se muestra el proceso

completo, sin embargo, esta misma corrección fue realizada para el resto de los procesos que presentaban este error.

Por otro lado, se incorporó un nuevo tipo de novedad en los procesos de clasificación de mercadería, relacionado con la gestión de cajas con el número único o ID incorrecto. Esta situación se presentaba con frecuencia en la operación, pero no se encontraba documentada en los procesos oficiales. La actualización correspondiente se muestra en el Anexo 3, en el proceso de clasificación de devoluciones por recall, solo se incluye el extracto corregido para precautelar la información de la empresa.

6.3. Resultados y discusión

Como resultado de la validación y corrección de los procesos, se logró eliminar los errores de diagramación identificados inicialmente. Se pasó de registrar tres errores de diagramación a cero, facilitando la lectura y comprensión de los diagramas de flujo. Por otro lado, se corrigieron los cuatro procesos incompletos mediante la inclusión de las actividades que no habían sido documentadas, logrando así que todos los procesos reflejen de manera íntegra la operación real.

A partir de estos resultados, es posible discutir los efectos de la validación y corrección de los procesos documentados. La corrección de los diagramas de procesos y la inclusión de actividades faltantes generan varios beneficios para el área. En primer lugar, se fortalece la transparencia de la información, ya que los diagramas actualizados reflejan de manera precisa las actividades realizadas en la operación. Además, la nueva documentación sirve como respaldo ante auditorías internas o externas.

En segundo lugar, se facilita la identificación de ineficiencias operativas, al reflejar actividades innecesarias o repetitivas que pueden ser optimizadas. Finalmente, las correcciones realizadas permiten que personas ajenas a la operación, como nuevos colaboradores o personal de otras áreas, puedan comprender el funcionamiento del sistema únicamente a través de la documentación.

Con esto concluye el análisis correspondiente al primer objetivo específico. En la siguiente sección se presenta el desarrollo del segundo objetivo específico.

7. Digitalización de documentos

A partir de los datos recolectados, se realiza un análisis de las variables, con el propósito de identificar factores que inciden en la eficiencia operativa y determinar posibles oportunidades de mejora. El enfoque dentro de la etapa Analizar se orienta a comprender la distribución del tiempo en las actividades involucradas, así como las implicaciones que esto tiene en la gestión documental y en la optimización de los recursos disponibles. Posteriormente, se presentan las acciones tomadas dentro de la etapa Mejorar, y, por último, se presenta una sección de resultados y discusión.

7.1. Analizar

- **Tiempo de digitalización de documentos (124.85 seg):** la fase de organizar los documentos consiste en renombrar el archivo escaneado según el proveedor de la mercadería y la fecha en la que se realiza la devolución. Posteriormente, el archivo se almacena en carpetas organizadas por proveedor, y por fecha de entrega. Este conjunto de pasos toma un 72% del tiempo de la actividad, lo cual evidencia una

oportunidad de automatización en esta fase. Se puede eliminar la intervención del operador para realizar la organización de los documentos.

- **Cantidad de documentos pendientes por digitalizar (+700):** Los documentos pendientes por digitalizar se han ido acumulando desde octubre, lo que evidencia un abandono progresivo de esta actividad. Esta situación refleja una carga de trabajo considerable que aún está pendiente de ser gestionada, representando un claro incumplimiento del proceso de digitalización. Como consecuencia, ante un eventual reclamo de un proveedor, es necesario buscar el documento físico, lo cual implica una mayor inversión de tiempo y recursos, afectando la eficiencia del servicio.

7.2. Mejorar

Gracias a la información obtenida dentro de la etapa Analizar, es posible tomar decisiones concretas para optimizar la situación. Durante esta etapa se diseñó y desarrolló un sistema de clasificación automática de documentos mediante el uso de Python que puede encontrarse en el Anexo 4. Debido a la naturaleza de los archivos a clasificar, (imágenes en formato PDF) se implementó el uso de Tesseract, un sistema de reconocimiento óptico de caracteres (OCR) que extrae el texto contenido dentro de los documentos (Tesseract-OCR, 2024). Posteriormente, el texto extraído es procesado por SpaCy, un sistema de procesamiento de lenguaje natural (NLP) que toma la información necesaria para la clasificación del documento (SpaCy, 2024).

El sistema desarrollado funciona completamente de manera local, por lo que no se vulnera información sensible de la empresa. Además, dentro de la función de extraer texto mediante Tesseract, se implementó una etapa de preprocesamiento de la imagen. En esta

etapa se aumentan el contraste del documento y se aplica un aumento de zoom tres veces mayor al original para facilitar la lectura de este y tener un desempeño más consistente. El sistema procesa todos los archivos presentes en una carpeta de OneDrive conectada directamente al equipo con el cual se escanean los documentos. Los archivos se procesan uno a uno de manera secuencial hasta que la carpeta quede vacía. Una vez que el contenido se extrae mediante el uso de Tesseract, SpaCy toma el código del proveedor y el ID de la devolución. Para obtener estos datos en específico, se realiza una búsqueda de palabras clave como “proveedor” y “devolución”, posteriormente se comprueba que la información que prosigue a estas palabras sea de tipo numérico. Con estos datos se renombra el archivo usando el siguiente formato “*NombreProveedor_IDDevolución_Fecha*”. Mediante este nombre se clasifica en carpetas según el proveedor y el mes de la devolución. Adicionalmente, se agregó la característica de crear carpetas en caso de ser necesario, y, cuando un archivo no puede ser renombrado adecuadamente, se envía a una carpeta separada para que sea revisado por el operador.

Para lograr un desempeño consistente por parte del código, se realiza la lectura del código del proveedor y no de su nombre directamente. Con el código del proveedor, se realiza una consulta en la base de proveedores de FEMSA y se obtiene el nombre. El nombre extraído para cada proveedor es el mismo que se usa para nombrar las carpetas, esto simplifica la clasificación de los documentos al manejar menos variables.

El programa pasó por dos fases de desarrollo. De manera preliminar, se realizaron pruebas únicamente con SpaCy como librería principal. Sin embargo, debido a los resultados obtenidos durante las pruebas preliminares, se decidió integrar Tesseract con el fin de alcanzar un desempeño mayor. La Tabla 5 muestra un resumen de los resultados obtenidos

tanto en las pruebas preliminares como en la versión final del código. Para esta situación, se entiende como tasa de éxito al porcentaje de documentos que fueron renombrados y clasificados adecuadamente.

Versión	Tamaño de muestra	Tiempo de procesamiento	Tasa de éxito
Código preliminar con SpaCy	118 documentos	127 minutos	34%
Integración de SpaCy y Tesseract	281 documentos	7.5 minutos	81%

Tabla 5. Resultados de la prueba de desempeño

La integración de SpaCy y Tesseract supuso una mejora significativa tanto en el tiempo de procesamiento como en la tasa de éxito del sistema. De acuerdo con la Tabla 5, se redujo efectivamente el tiempo de procesamiento un 94.09% con respecto a la versión preliminar y la tasa de éxito se incrementó en 47 puntos porcentuales. Aunque la tasa de éxito en la clasificación de documentos aumentó drásticamente, aún existe un 19% de documentos que no están siendo clasificados con éxito. Esto se debe a errores en la lectura de los documentos. Debido a la calidad de los escaneos, en ocasiones, estos son ligeramente borrosos, lo cual hace que los códigos de proveedor y el ID de la devolución no se lean correctamente. Esto genera que los documentos no sean renombrados correctamente y, por lo tanto, se mueven a la carpeta de errores para que sean revisados por el operador.

7.3. Resultados y discusión

Siguiendo los resultados en el desempeño de la aplicación desarrollada, estos permiten pasar a una futura implementación. Con esta implementación, el tiempo que toma esta actividad pasa de 124.85 segundos a 34.85 segundos por documento. En otras palabras,

se elimina la clasificación manual por parte del operario y la actividad se ejecuta un 72% más rápido. Tomando en cuenta que se tienen aproximadamente 700 documentos pendientes por digitalizar, el tiempo que se ahorra en procesar esta cantidad de documentos es de alrededor de 18 horas. Además, de manera mensual, se estima que se generan 400 actas de devolución a proveedor, por lo tanto, esta implementación permite un ahorro aproximado de 10 horas de trabajo de manera mensual.

En función de los resultados obtenidos, se puede analizar el impacto del programa desarrollado en la operación actual. Mediante el programa desarrollado se facilita cumplir el requerimiento de realizar respaldos digitales de las actas de devolución, además de que el formato de nombre propuesto facilita la búsqueda de documentos específicos en caso de ser necesario. Originalmente no se incluía el ID de la devolución, por lo que, en caso de reclamo, las búsquedas se realizaban únicamente conociendo el nombre del proveedor y la fecha.

Gracias a la versatilidad de esta aplicación, es posible extender su implementación al resto de actividades que requieren un respaldo digital y una posterior clasificación de los documentos que generan.

Con esto concluye el análisis correspondiente al segundo objetivo específico. En la siguiente sección se presenta el desarrollo del tercer objetivo específico.

8. Gestión y visualización de datos

A partir de los resultados obtenidos, en la etapa Analizar se estudia el conjunto de variables asociadas al acceso, estructura y visualización de la información, con el fin de identificar limitaciones técnicas y proponer acciones que serán desarrolladas dentro de la

etapa Mejorar que optimicen el uso del dashboard como herramienta de apoyo a la toma de decisiones. Por último, se incluye una sección de resultados y discusión.

8.1. Analizar

- Cantidad de fuentes de datos (6):** se identificaron seis fuentes conectadas en Power Query. Aunque integrar múltiples fuentes puede enriquecer el análisis, en este caso particular se observó la presencia de duplicidad de bases, exceso de filas innecesarias y falta de depuración de los datos, lo que incrementa la complejidad del proceso de actualización y eleva la probabilidad de errores. Asimismo, el volumen de registros supera el límite máximo de filas permitido por Excel (1,048,576 filas), lo que limita la cantidad de información que puede ser procesada y analizada mediante el dashboard. Estas deficiencias evidencian la necesidad de aplicar mejores prácticas de integración de fuentes y optimización de consultas
- Cantidad de preguntas a las que responde el dashboard (4):** Actualmente, no se aprovechan todos los datos disponibles dentro de Logística Inversa, ya que el dashboard operativo responde únicamente a cuatro cuestionamientos específicos: el valor, las unidades y la causal de la mercadería ingresada en una fecha determinada; el valor, mercadería devuelta al proveedor en una fecha determinada; el ranking de proveedores con mayor cantidad de devoluciones; y el número de devoluciones que ingresaron en una fecha específica. Este resultado revela que el alcance del dashboard está acotado a cuatro consultas de carácter operativo relacionadas con líneas de producto, unidades, valores monetarios y documentos de recepción o devolución. Esta limitación se debe a que no se aprovecha todo el volumen de datos que se carga en el dashboard, restringiendo así el potencial analítico disponible, reduciendo el

soporte a la toma de decisiones oportunas y evidenciando la necesidad de ampliar la variedad y profundidad de las consultas disponibles, de acuerdo con los principios establecidos para el diseño de dashboards eficientes (Ghiglione, 2021).

- **Tiempo de Acceso a la Información (TAI):** Se encontró un tiempo aproximado de 40 minutos desde la carga de datos hasta la visualización de los mismo. Durante el proceso de carga se observaron errores como bloqueos del sistema, cierres imprevistos y demoras al actualizar filtros y tablas dinámicas, afectando directamente la productividad del área y elevando los riesgos de operar con información desactualizada.
- **Índice de Peso del Archivo (IPA):** Se determinó que el archivo de Excel tiene un tamaño de 432 MB, resultando en un IPA de 4.32. Este tamaño elevado dificulta la manipulación del archivo, provoca lentitud en las operaciones, errores frecuentes al aplicar filtros, y limita la posibilidad de compartir el archivo con otras áreas de manera eficiente. Adicionalmente, incrementa la probabilidad de cierres inesperados del sistema y bloqueos ("No responde"), afectando la disponibilidad de la información

Adicionalmente, se identificaron deficiencias en la experiencia del usuario (UX) y en el diseño de la interfaz del usuario (UI), como se muestra en el Anexo 5. En la zona 1 del anexo, se observa un bajo contraste visual entre los gráficos de líneas y el color de fondo, lo que dificulta la legibilidad. Asimismo, el uso inadecuado de colores dentro de los gráficos compromete la rápida interpretación de los datos. En las zonas 1, 2 y 3, correspondientes a los gráficos, filtros y secciones de visualización de datos, se evidencia una falta de jerarquía informativa que impide distinguir con claridad los elementos más relevantes. Además, en los

espacios intermedios entre cada zona se aprecia un aprovechamiento ineficiente del área disponible en pantalla. Todas estas falencias limitan la comprensión ágil y precisa de la información, afectando negativamente la eficacia del dashboard como herramienta de apoyo a la toma de decisiones (ver Anexo 5; imagen referencial, datos reales omitidos por confidencialidad). Estas falencias reducen la eficacia del dashboard como herramienta de apoyo a la toma de decisiones y limitan el análisis profundo de la información disponible. Se identifica así una oportunidad de mejora tanto en la visualización como en el aprovechamiento del volumen de datos cargados, optimizando la forma en que se presenta y se interpreta la información.

8.2. Mejorar

Con base en los problemas detectados en la etapa Analizar, se implementaron diversas soluciones estructuradas para optimizar el sistema de gestión de datos de la bodega de Logística Inversa. Una de las primeras acciones fue la creación de un diccionario de datos (ver Anexo 6, se muestra una parte de los datos identificados por ser información confidencial), el cual permitió identificar y categorizar todos los datos utilizados en la construcción del dashboard en Excel. Se identificaron un total de 50 campos de datos extraídos de las diversas fuentes integradas en el dashboard. No obstante, para localizar aquellos verdaderamente necesarios para el análisis, se llevó a cabo un proceso de normalización de datos. Este procedimiento fue indispensable debido a que, al descargar los datos desde su base de datos y utilizarlos directamente en el dashboard, se generaba una carga de información mayor a la requerida. Por este motivo, los datos fueron organizados bajo el esquema de la Tercera Forma Normal (3FN), que consiste en estructurar la información de

forma que se eliminen redundancias y se establezcan relaciones lógicas entre las tablas, garantizando así la integridad y consistencia de la base de datos. Como resultado, se logró reducir la redundancia, mejorar la integridad referencial y disminuir de manera significativa el peso del archivo (Hernández & Rodríguez, 2008). Este proceso permitió determinar con mayor claridad qué datos eran necesarios para la visualización y evidenció que muchos campos no aportaban valor directo al análisis operativo. Por lo cual, mediante esta depuración, se seleccionaron únicamente 36 campos relevantes, eliminando 14 campos necesarios, lo cual simplificó la estructura de datos y facilitó su gestión.

Aprovechando esta reestructuración, se recomendó y procedió a la migración hacia Power BI, una plataforma específicamente diseñada para la creación de análisis de datos empresariales que facilita la integración de grandes volúmenes de información de manera más eficiente que Excel tradicional (Microsoft, 2024). El uso de Power BI ofreció ventajas clave, como la capacidad de realizar consultas optimizadas mediante el lenguaje DAX (Data Analysis Expressions), que permite generar medidas, cálculos y análisis dinámicos de forma interna, sin aumentar significativamente el tamaño del modelo (Collie & Singh, 2015). Gracias a la implementación de DAX y al nuevo esquema de datos estructurado, se logró reducir la dependencia de cálculos manuales en Excel y se mejoró significativamente el rendimiento general del sistema.

Adicionalmente, mediante comunicación personal con el coordinador de LI (2025), se amplió la capacidad analítica del dashboard incorporando nuevas consultas y medidas. Se añadieron consultas sobre líneas, unidades, monto, y documentos de ingreso y salida de mercadería en Logística Inversa, específicamente aquellos documentos que registran las recepciones desde los puntos de venta y las devoluciones hacia los proveedores. Así mismo,

se incorporaron nuevas medidas que permiten calcular totales, promedios y porcentajes de devolución respecto a la mercadería ingresada (ver Anexo 7; datos reales omitidos por confidencialidad). Esta mejora permite evaluar el desempeño de los procesos de devolución de forma más precisa, ya que no toda la mercadería enviada a los proveedores logra ser devuelta efectivamente.

Finalmente, para asegurar el correcto uso del nuevo dashboard y facilitar su adopción por parte del personal operativo, se elaboró una Guía de Usuario. Este documento describe de manera clara y estructurada el funcionamiento del sistema, las principales consultas disponibles, los procedimientos de actualización y las mejores prácticas para interpretar correctamente los indicadores.

8.3. Resultados y discusión

Tras el uso de Power BI, el número de campos utilizados en el dashboard se redujo de 50 a 36, eliminándose 14 campos redundantes. El archivo de Excel, cuyo tamaño inicial era de 432 MB, experimentó una reducción del 90% de su peso hasta llegar a 43 MB, alcanzando un Índice de Peso del Archivo (IPA) de 0.43. El Tiempo de Acceso a la Información (TAI) disminuyó de 40 minutos a aproximadamente 10 minutos, representando una reducción del 75%. El número de consultas operativas disponibles se incrementó, incorporando nuevas métricas como líneas, unidades, monto y documentos de ingreso y salida de mercadería, así como medidas adicionales de total, promedio y porcentaje.

Asimismo, según comunicación personal con el coordinador de LI (2025), la experiencia de usuario y su interfaz (UI/UX) le pareció atractiva e intuitiva para la búsqueda y relación de datos, permitiendo obtener información de forma más rápida y autónoma.

A partir de estos resultados, se pueden discutir los impactos de la implementación de Power BI en la gestión de datos de Logística Inversa. La implementación de Power BI permitió mejorar sustancialmente la eficiencia de la gestión de datos en Logística Inversa, solucionando los principales problemas de estabilidad, velocidad de acceso y limitaciones analíticas que presentaba el sistema anterior basado en Excel. La reestructuración del modelo de datos y la incorporación de medidas más específicas optimizaron la visualización de la información y facilitaron el análisis de los procesos de devolución, permitiendo identificar oportunidades de mejora con mayor rapidez y precisión.

Además, el incremento en la cantidad y calidad de las consultas disponibles fortaleció la capacidad de respuesta operativa y la toma de decisiones estratégicas. El uso de métricas como los porcentajes de devolución introdujo un enfoque más detallado para evaluar el desempeño de la Logística Inversa, contribuyendo a un manejo más eficiente de los inventarios y a una mejor gestión de las relaciones con los proveedores. Estas mejoras consolidan el valor de la visualización de datos como herramienta crítica para el control y la optimización de las operaciones.

Una vez desarrolladas y analizadas todas las etapas del proyecto, es posible identificar los principales hallazgos y valorar el cumplimiento de los objetivos propuestos. En este sentido, a continuación, se presentan las conclusiones más relevantes que permiten sintetizar los resultados alcanzados y plantear posibles líneas de mejora o continuidad para futuros trabajos.

9. Conclusiones

La corrección de la documentación de procesos representó un avance significativo para el área de Logística Inversa, ya que permitió eliminar errores de representación y omisiones de actividades recurrentes en la ejecución. Esta mejora no solo mejora la claridad de los procedimientos, sino que también fortalece la transparencia y la confiabilidad de las operaciones. Esto garantiza que todas las acciones estén debidamente registradas y alineadas con la práctica real. Además, contar con una documentación precisa facilita la capacitación de nuevos colaboradores y permite que el personal de otras áreas pueda integrarse temporalmente o colaborar con Logística Inversa de forma más efectiva, ya que disponen de una guía estructurada y coherente sobre cómo deben ejecutarse las actividades.

Por otra parte, la implementación del sistema de clasificación automática de documentos supuso una mejora significativa en la eficiencia operativa. Este cambio redujo drásticamente el tiempo requerido para ejecutar esta actividad, liberando el tiempo del operador, que ahora puede ser reasignado a funciones de mayor prioridad dentro del área. Al automatizar este proceso, no solo se agiliza el flujo de trabajo, sino que también se minimiza la posibilidad de errores humanos asociados a la clasificación manual. Además, con el formato de nombre propuesto, se facilita la búsqueda de documentos específicos en caso de ser necesario.

Finalmente, la creación de un nuevo modelo de datos, junto con la migración del sistema a Power BI, representó una transformación profunda en términos de eficiencia, estabilidad y capacidad analítica del área. La reducción del 75% en el tiempo de acceso a la información se traduce en una mejora directa en la productividad del coordinador del área, quien ahora puede dedicar menos tiempo a la gestión de datos y más tiempo al análisis

estratégico. Este nuevo enfoque mejora la velocidad con la que se accede a la información y también incrementa la precisión de los análisis. El preprocesamiento de datos realizado por el modelo de datos elimina los errores de carga y fortalece la calidad de los cálculos. Esto resulta en un sistema sólido y fácil de usar para la toma de decisiones basadas en datos.

De manera adicional a las conclusiones presentadas, es importante reconocer las limitaciones que tuvieron cierta influencia sobre el proyecto desarrollado. La identificación de estas permite contextualizar mejor los resultados obtenidos y brindar una base sólida para futuros proyectos. A continuación, se presentan las principales limitaciones encontradas, así como recomendaciones para superar estos desafíos.

10. Limitaciones y recomendaciones

La falta de conocimiento sobre los procesos de devolución entre los puntos de venta y el área de Logística Inversa representa una barrera significativa para una comunicación efectiva entre ambas partes. Dado que los puntos de venta y el área de Logística Inversa son interdependientes, esta desconexión dificulta la identificación de oportunidades de mejora para ambos y complica la implementación de estrategias orientadas a optimizar la gestión de devoluciones. A esta situación se suma una cierta actitud de reserva por parte del personal de LI al momento de compartir información relacionada con el desempeño o errores dentro de la operación. Esto limitó parcialmente el análisis de la situación actual; sin embargo, esta resistencia se redujo al demostrar los beneficios del proyecto realizado.

Otro inconveniente identificado es la pérdida de trazabilidad de la mercadería una vez que esta sale del centro de distribución. Cuando la mercadería se despacha a los puntos de venta, no se comparte con ellos información relacionada al lote de cada producto, solo la

cantidad enviada. En consecuencia, los operadores en los puntos de venta deben revisar uno a uno los productos para identificar la mercadería que debe ser devuelta. Por otro lado, desde Logística Inversa tampoco se conoce si un producto en concreto ha sido vendido, está en mal estado o si está próximo a vencer. Lo que genera incertidumbre sobre el volumen y el tipo de devoluciones que se tendrán que gestionar en los siguientes meses.

Ante estos desafíos, se establecen las siguientes recomendaciones. En primer lugar, es fundamental establecer canales de comunicación recurrentes entre PDV y LI, de modo que exista un flujo constante y oportuno de información. Esto facilitaría la resolución de inconvenientes que puedan presentarse en las devoluciones y se promovería una comunicación activa entre ambas partes. Asimismo, se recomienda compartir información relevante entre ambas partes, como los lotes de los productos y el estado de los productos en punto de venta. Esta medida simplificaría el proceso de devolución en punto de venta y permitiría a LI realizar pronósticos más certeros sobre la mercadería que recibirá, mejorando así su capacidad de respuesta.

Por otro lado, antes de implementar mejoras a nivel de código en la aplicación desarrollada para la digitalización de documentos, se recomienda comenzar por ajustes operativos simples. Por ejemplo, aumentar el tamaño de letra en las actas de devolución y mejorar la resolución de los escaneos no implica un costo adicional y puede generar un impacto inmediato en el desempeño del sistema. Finalmente, se recomienda continuar con la migración de los sistemas de visualización y gestión de datos hacia Power BI, dado que los resultados obtenidos en este proyecto evidencian las ventajas de utilizar una herramienta más robusta y adecuada para el volumen de datos generado en la empresa, asegurando que la información es confiable y accesible.

11. Referencias bibliográficas

- ARCSA. (2025). *Normativa – Instructivos – Agencia Nacional de Regulación, Control y Vigilancia Sanitaria. Gob.ec.* <https://www.controlsanitario.gob.ec/documentos-vigentes/>
- Collie, R., & Singh, A. (2015). *Power Pivot and Power BI: The Excel User's Guide to DAX, Power Query, Power BI & Power Pivot in Excel 2010-2016.* Tickling Keys, Inc.
- FEMSA Salud Ecuador. (2023, August 8). *La Corporación - CorporacionGPF.* CorporacionGPF. <https://www.corporaciongpf.com/la-corporacion/>
- Ferrão, P. (2020). *Mastering Power Query in Power BI and Excel: A Comprehensive Guide to Data Transformation.* Packt Publishing.
- Ghiglione, R. (2021). Indicadores de gestión en mejora continua: claves para definir métricas relevantes. *Revista Latinoamericana de Ingeniería Industrial*, 12(1), 45-57.
- Hernández, C., & Rodríguez, J. E. R. (2008). Preprocesamiento de datos estructurados. *Revista vínculos*, 4(2), 27-48.
- Hurtado García, Ketty del Rocío. (2019). Responsabilidad social empresarial, logística inversa y desarrollo de la contabilidad de costos. *Cooperativismo y Desarrollo*, 7(3), 333-340. Epub 02 de diciembre de 2019. Recuperado en 05 de mayo de 2025, de http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S2310-340X2019000300333&lng=es&tlng=es

- Indrawati, S., & Ridwansyah, M. (2015). Manufacturing continuous improvement using Lean Six Sigma: An iron ores industry case application. *Procedia Manufacturing*, 4, 528–534. <https://doi.org/10.1016/j.promfg.2015.11.072>
- Jinks, T. (2019). Psychological perspectives on reality, consciousness and paranormal experience. Palgrave Macmillan. <https://link.springer.com/book/10.1007/978-3-030-28902-7>
- Kulkarni, S. (2024). *Mold qualification flowchart, production release, and troubleshooting*. En S. Kulkarni (Ed.), *Robust process development and scientific molding* (3ra ed., pp. 311-333). Hanser. <https://doi.org/10.1016/B978-1-56990-908-9.50013-2>
- López, F. A. G., Monsalve, L. L. H., & Coronado, M. H. V. (2021). Mejora de la productividad empleando la metodología DMAIC. *INGENIERÍA: Ciencia, Tecnología e Innovación*, 8(2), 77-91. <https://doi.org/10.26495/icti.v8i2.1907>
- Microsoft. (2024). Best practices for optimizing Excel workbooks. Recuperado el 6 de abril de <https://support.microsoft.com/es-es/office/especificaciones-y-l%C3%ADmites-de-excel-1672b34d-7043-467e-8e27-269d656771c3>
- Montgomery, D. C. (2020). *Introduction to Statistical Quality Control* (8th ed.). Wiley.
- Nandakumar, N., Saleeshya, P. G., & Harikumar, P. (2020). *Bottleneck Identification and Process Improvement by Lean Six Sigma DMAIC Methodology*. *Materials Today: Proceedings*, 24, 1217–1224. <https://doi.org/10.1016/j.matpr.2020.04.436>
- Nickell, E. B., Schwebke, J., & Goldwater, P. (2023). An introductory audit data analytics case study: Using Microsoft Power BI and Benford's Law to detect accounting

irregularities. *Journal of Accounting Education*, 64, 100855.

<https://doi.org/10.1016/j.jaccedu.2023.100855>

Peralta, C., Abarca, J. L. P., & Luis José. (2018). Reducción de costos de distribución de productos farmacéuticos aplicando el Sistema Kanban Estudio de Caso. 6° Cilog, 6(1), 12–19.

Pérez-Balboa, I. C., & Caballero-Morales, S. O. (2024). Define, Measure, Analyze, Improve, Control (DMAIC). In *Lean Manufacturing in Latin America: Concepts, Methodologies and Applications* (pp. 333-352). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-70984-5_15

Robby, G. A., Tandra, A., Susanto, I., Harefa, J., & Chowanda, A. (2019). Implementation of Optical Character Recognition using Tesseract with the Javanese Script Target in Android Application. *Procedia Computer Science*, 157, 499–505.
<https://doi.org/10.1016/j.procs.2019.09.006>

SpaCy (2024). *Industrial-strength Natural language processing in Python*. <https://spacy.io/>

Taylor, Jake Donald, "The Impact of Gemba Walks On Preventative Maintenance Productivity" (2021). *Masters Theses & Specialist Projects*. Paper 3545.
<https://digitalcommons.wku.edu/theses/3545>

Tesseract-OCR. (2024). *GitHub - tesseract-ocr/tesseract: Tesseract Open Source OCR Engine (main repository)*. GitHub. <https://github.com/tesseract-ocr/tesseract>

Utama, D. M., & Abirfatin, M. (2023). *Sustainable Lean Six-Sigma: A new framework for improving sustainable manufacturing performance*. *Cleaner Engineering and Technology*, 17, 100700. <https://doi.org/10.1016/j.clet.2023.100700>

Venegas, A. (2018). Metodología DMAIC para el control de merma conocida en productos farmacéuticos. Universidad Militar Nueva Granada.

<https://repository.umng.edu.co/server/api/core/bitstreams/e6ff80ce-f39d-48ce-ab50-72128821a17b/content>

Vidal, B. P., Soler, V. G., & Molina, A. I. P. (2018). Metodología Six Sigma. Comparación entre ciclo PDCA y DMAIC. In Cuadernos de investigación aplicada (pp. 27-34).

3ciencias. <https://dialnet.unirioja.es/servlet/articulo?codigo=7173589>

Zhu, Q., Facco, P., Zhao, Z., & Barolo, M. (2024). Capturing connectivity information from process flow diagrams by sequential-orthogonalized PLS to improve soft-sensor performance. *Chemometrics and Intelligent Laboratory Systems*, 252, 105192.

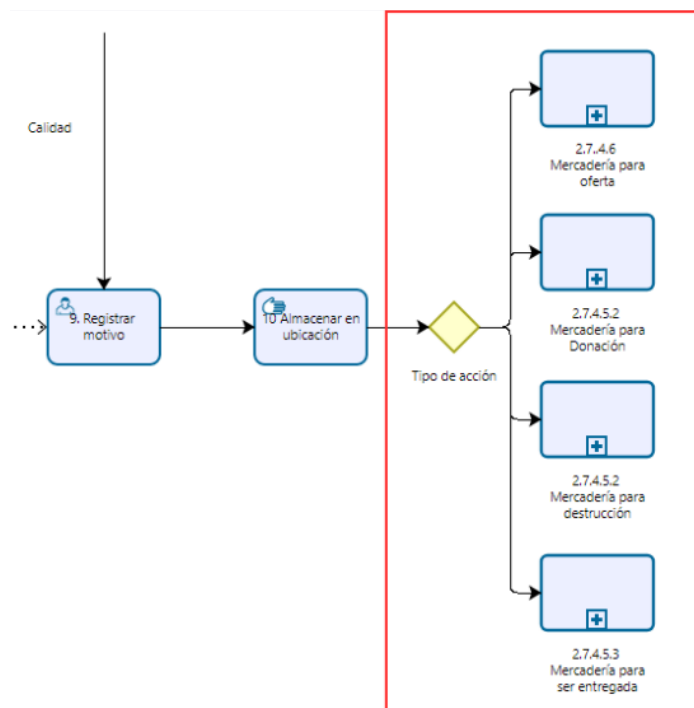
<https://doi.org/10.1016/j.chemolab.2024.105192>

12. Anexos

Anexo 1. Toma de datos para el tiempo de escaneo y clasificación

Observación	Tiempo de Escaneo (seg)	Tiempo de Clasificación (seg)
1	34	91
2	36	89
3	35	92
4	33	88
5	36	92
6	32	90
7	37	87
8	35	93
9	36	86
10	34	90
11	33	91
12	37	88
13	35	92
14	34	87
15	36	93
16	32	90
17	37	89
18	35	94
19	36	87
20	34	91
Tiempo promedio	34.85	90
Desviación estandar	1.57	2.32

Anexo 2. Corrección de errores de diagramación



Anexo 3. Inclusión de actividades faltantes en los procesos

P2	¿Tipo de novedad?	<ul style="list-style-type: none"> - Calidad: ir a la actividad 14. - Cantidad: ir a la actividad 16. - Parámetros del sistema (PVP): ir a la actividad 18. - ID de caja incorrecto: separar la caja en el área designada.
----	-------------------	--

Anexo 4. Aplicación desarrollada en Python

```

import pytesseract
from pytesseract import Output
import pypdfium2 as pdfium
import spacy
import pandas as pd
import os
from datetime import datetime
from PIL import Image

```

```

# Configuración Tesseract para Windows
pytesseract.pytesseract.tesseract_cmd = r"C:\Program
Files\Tesseract-OCR\tesseract.exe"

# Cargar spaCy
nlp = spacy.load("es_core_news_lg")

# Leer base de proveedores
proveedores = pd.read_excel("Base Proveedores Actualizada.xlsx",
sheet_name="Hoja1")

# Carpeta con los escaneos
carpeta_escaneos = "Escaneos/Escaneos de prueba"
archivos = [
    os.path.join(carpeta_escaneos, f)
    for f in os.listdir(carpeta_escaneos)
    if os.path.isfile(os.path.join(carpeta_escaneos, f)) and
f.lower().endswith(".pdf")
]

# Carpetas para documentos que no se pudieron procesar
carpeta_errores = os.path.join(carpeta_escaneos, "Documentos que
no se pudieron procesar")
os.makedirs(carpeta_errores, exist_ok=True)

# Función para aplicar OCR a cada página y concatenar el texto
def extraer_texto_ocr(pdf_path):
    pdf = pdfium.PdfDocument(pdf_path)
    texto_total = ""

    for i in range(len(pdf)):
        page = pdf[i]
        image = page.render(scale=3).to_pil()
        imagen_bn = image.convert('L').point(lambda x: 0 if x <
150 else 255, '1')
        texto = pytesseract.image_to_string(imagen_bn, lang='spa',
config='--psm 6')
        texto_total += texto + "\n"

```



```

pdf.close()
return texto_total

# Función para extraer datos con spaCy
def extraer_info_spacy(texto):
    doc = nlp(texto)
    cod_prov = None
    cod_dev = None

    for i, token in enumerate(doc):
        if token.lower_ == "proveedor" and i + 1 < len(doc):
            next_token = doc[i + 2]
            if next_token.text.isdigit():
                cod_prov = next_token.text

        if token.lower_ in ["devolucion", "devolución"] and i + 2
< len(doc):
            if doc[i + 1].text == ":":
                cod_dev = doc[i + 2].text
            else:
                cod_dev = doc[i + 1].text

        if cod_prov and cod_dev:
            break

    return cod_prov, cod_dev

# Contadores
ok_count = 0
warning_count = 0
error_count = 0

# Procesar archivos
for ruta_origen in archivos:
    try:
        texto = extraer_texto_ocr(ruta_origen)
        cod_prov, cod_dev = extraer_info_spacy(texto)

        # CASO WARNING: no extrajo códigos
        if not cod_prov or not cod_dev:

```

```

        print(f"[!] No se extrajo código prov o devolución:
{ruta_origen}")
        warning_count += 1
        # Mover a carpeta Warnings
        dest = os.path.join(carpeta_errores,
os.path.basename(ruta_origen))
        os.rename(ruta_origen, dest)
        continue

    fila_prov = proveedores.loc[proveedores["Cod Prov"] ==
int(cod_prov)]
    # CASO WARNING: proveedor no encontrado en Excel
    if fila_prov.empty:
        print(f"[!] Código proveedor {cod_prov} no encontrado en Excel:
{ruta_origen}")
        warning_count += 1
        # Mover a carpeta Warnings
        dest = os.path.join(carpeta_errores,
os.path.basename(ruta_origen))
        os.rename(ruta_origen, dest)
        continue

    # CASO OK: renombrar y mover a carpeta de proveedor
    nom_proveedor = fila_prov['nombre_carpeta'].values[0]
    carpeta_destino = os.path.join(carpeta_escaneos, nom_proveedor)
    os.makedirs(carpeta_destino, exist_ok=True)

    fecha_actual = datetime.now().strftime("%d-%m-%Y")
    nuevo_nombre = f"{nom_proveedor}_{cod_dev}_{fecha_actual}.pdf"
    ruta_nueva = os.path.join(carpeta_destino, nuevo_nombre)

    os.rename(ruta_origen, ruta_nueva)
    print(f"[✓] Archivo renombrado y movido: {ruta_nueva}")
    ok_count += 1

except Exception as e:
    # CASO ERROR
    print(f"[X] Error procesando {ruta_origen}: {e}")
    error_count += 1
    # Mover a carpeta Errors
    dest = os.path.join(carpeta_errores, os.path.basename(ruta_origen))
    try:
        os.rename(ruta_origen, dest)
    except Exception as mv_err:

```

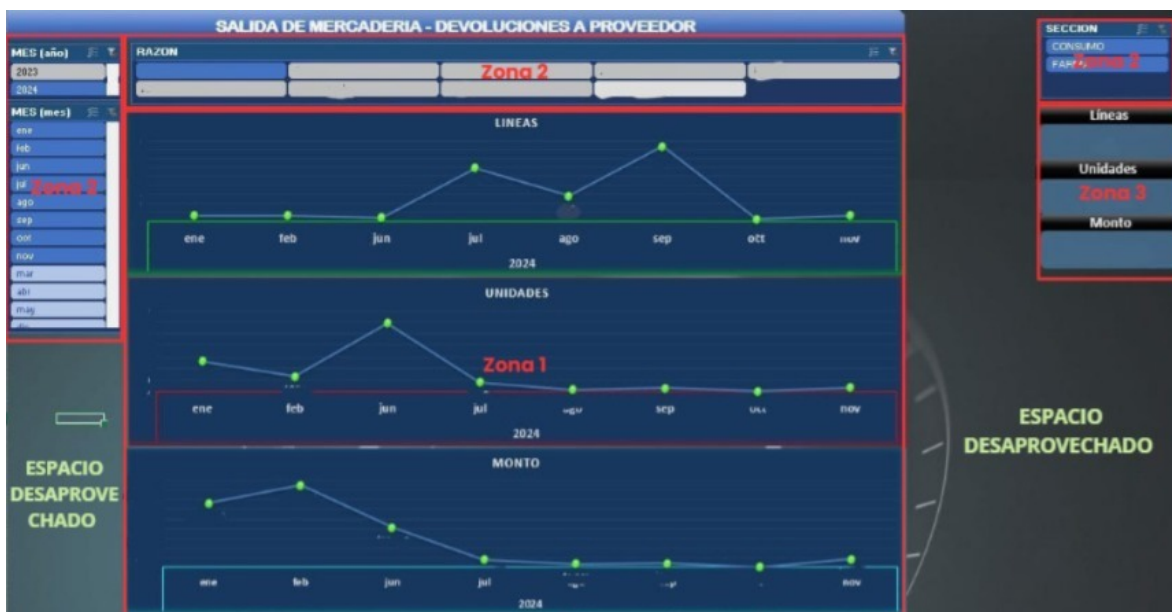
```

        print(f"[X] No se pudo mover a Errors: {mv_err}")
        continue

# Reporte final
print("\n=== RESUMEN FINAL ===")
print(f"✓ Correctos      : {ok_count}")
print(f"! Advertencias   : {warning_count}")
print(f"X Errores        : {error_count}")
print("=====")

```

Anexo 5. Dashboard actual en Excel



Anexo 6. Extracto del diccionario de datos

Campo	Campo	Tipo de dato	Descripción	Requiere valor (no null)	Valores permitidos
ITEM	item_id	INT	Código único del producto	Sí	≥ 1
DESCRIPCION	descripcion	VARCHAR	Descripción del producto	Sí	Texto
COSTO	costo	DECIMAL	Costo por cada producto	Sí	≥ 0
SECCIÓN CONTABLE	unidad_comercial	VARCHAR	Categoría contable del producto	Sí	Texto

CATEGORIA	cod_proveedor	VARCHAR	Categoría de la sección contable	No	Texto
SUBCATEGORIA	subcategoria	VARCHAR	Subcategoría de la sección contable	No	Texto
REFRIGERADO	refrigerado	VARCHAR	El producto es o no refrigerado	No	Sí / No
PSICOTROPICO	psicotropico	VARCHAR	El producto es o no psicotrópico	No	Sí / No

Anexo 7. Dashboard optimizado en Power BI

