



**Estudio de tiempos de clasificación en equipos Put-to-Light y segmentación de
productividad de operarios: un enfoque basado en pruebas no paramétricas**

Andrés Esteban Marín Manco

Trabajo de grado para optar al título de Ingeniero Industrial

Asesores

Juan G. Villegas, Ph.D.

Angie Paola Vargas, M.Eng. (c)

Universidad de Antioquia

Facultad de Ingeniería

Ingeniería Industrial

Medellín, Colombia

2024

Cita	(Marín-Manco, 2024)
Referencia	Marín-Manco, A. E. (2024). <i>Estudio de tiempos de clasificación en equipos Put-to-Light y segmentación de productividad de operarios: un enfoque basado en pruebas no paramétricas</i> [Trabajo de grado]. Universidad de Antioquia, Medellín, Colombia.
Estilo APA 7 (2020)	



Grupo de Investigación Analítica e Investigación para la Toma de Decisiones – ALIADO



Centro de Documentación Ingeniería (CENDOI)

Repositorio Institucional: <http://bibliotecadigital.udea.edu.co>

Universidad de Antioquia – www.udea.edu.co

El contenido de esta obra corresponde al derecho de expresión de los autores y no compromete el pensamiento institucional de la Universidad de Antioquia ni desata su responsabilidad frente a terceros. Los autores asumen la responsabilidad por los derechos de autor y conexos.

Dedicatoria

A mi madre, Esneida, por su eterno amor, apoyo y acompañamiento a lo largo de mi vida.

A mi abuela, Dora Lina, quien me cuidó con amor y aprendí el bonito don de enseñar con dedicación.

A mi abuelo, Reinaldo Antonio, a quien admiré desde pequeño por su alegría, tenacidad y ver lo lindo de la vida.

A mis hermanos y familia, por confiar y apoyarme en cada uno de los caminos que he tomado y darme el impulso necesario cuando he fallado.

A mis amigos, compañeros y cada una de las personas que han hecho parte de este proceso formativo y me han ayudado a crecer como persona y profesional.

Agradecimientos

Agradecimientos especiales al profesor Juan G. Villegas y a Angie P. Vargas (estudiante de Maestría en Ingeniería), los cuales más allá de su rol de asesores, me brindaron todo su apoyo y estuvieron atentos ante cualquier emergencia o necesidad. A cada uno de los profesores con los cuales pude compartir dentro y fuera de las aulas, por la dedicación y pasión de enseñar y formar.

Finalmente, a la Universidad de Antioquia, por abrirme las puertas y sentir el orgullo de pertenecer a una de las mejores universidades de Colombia y Latinoamérica.

Tabla de contenido

Resumen	10
Abstract	11
Introducción	12
1 Marco Teórico	15
Definiciones	16
2 Estado del arte	18
3 Metodología	21
Fase 1. Identificación del problema y objetivos	22
Fase 2. Adquisición y preparación de los datos	22
Entendimiento del conjunto de datos	22
Preparación de los datos	23
Fase 3. Análisis Exploratorio de los datos e inferencias	26
4 Resultados	28
Análisis exploratorio y descriptivo de los datos	28
Tiempos y cantidades clasificadas por mes	28
Tiempos y Cantidades Clasificadas por Día de la Semana	31
Tiempos y Cantidades Clasificadas por Hora del Día	35
Tiempos de Clasificación según el Turno de Trabajo	38
Eficiencia según el tipo de equipo PTL	40
Tiempos de clasificación por operario	45
Correlación entre variables	47
Pruebas de estadística inferencial y no paramétricas	49
Pruebas de normalidad del tiempo de clasificación	49
Diferencia de medias en el tiempo de clasificación por SKU por tipo de equipo PTL	51

Diferencia de medias en el tiempo de clasificación por SKU por salidas en equipo tipo 1	51
Diferencia de medias en el tiempo de clasificación por SKU por salidas en equipo tipo 0	54
Diferencia de medias en el tiempo de clasificación por SKU por días	58
Diferencia de medias por turnos en el tiempo de clasificación por SKU	59
Diferencia de medias en el tiempo de clasificación por SKU de los operarios	60
Clasificación de operarios por productividad	61
Análisis de productividades por grupo	63
Análisis de productividades por grupo en cada tipo de equipo PTL	65
5 Conclusiones	67
6 Limitaciones y Trabajo Futuro	70
7 Referencias	72

Lista de tablas

Tabla 1 Datos suministrados por Muvum Corp.	22
Tabla 2 Algoritmo para cálculo de time_sort	26
Tabla 3 Planteamiento de hipótesis y pruebas estadísticas para clasificación de tiempos ajustadas a una distribución normal	50
Tabla 4 Métricas de desempeño de los grupos de salidas del equipo PTL tipo 1	53
Tabla 5 Métricas de desempeño de los grupos de salidas del equipo PTL tipo 0	56
Tabla 6 Fragmento del resultado de comparaciones pareadas de los operarios por sus medias de tiempo de clasificación por SKU	61
Tabla 7 Clasificación y cantidad de operarios por categoría de productividad	63
Tabla 8 Métricas de desempeño de cada una de las categorías de operarios por productividad	64
Tabla 9 Métricas de desempeño de cada una de las categorías de operarios en el equipo PTL tipo 1	65
Tabla 10 Métricas de desempeño de cada una de las categorías de operarios en el equipo PTL tipo 0	65

Lista de figuras

Figura 1	Proceso de operación del equipo PTL	15
Figura 2	Partes de un equipo PTL. (a) Zona con un solo módulo de estantería. (b) Zona con dos módulos de estantería	17
Figura 3	Fases del ciclo de vida de proyectos entorno a minería de datos aplicando la metodología CRISP-DM	21
Figura 4	Acción de lectura y confirmación de SKU	25
Figura 5	Acción de confirmación y confirmación de SKU	25
Figura 6	Tiempos de clasificación por Mes	29
Figura 7	Cantidades de SKU clasificadas por mes	30
Figura 8	Línea de tiempo de los operarios asignados a tareas de clasificación en equipos PTL	30
Figura 9	Tiempos de clasificación por día de la semana	32
Figura 10	Cantidades de SKU clasificadas por día de la semana	32
Figura 11	Comportamiento del tiempo de clasificación por SKU según el turno y día de la semana	33
Figura 12	Frecuencia de uso de los equipos PTL por día de la semana	34
Figura 13	Número promedio de operarios por día de la semana	34
Figura 14	Tiempos de clasificación por hora del día	35
Figura 15	Cantidades de SKU clasificadas por hora del día	36
Figura 16	Comportamiento del tiempo de clasificación por SKU según la hora del día y día de la semana	37
Figura 17	Tiempos de Clasificación según el Turno de Trabajo	38
Figura 18	Cantidades de SKU Clasificadas por Turno de Trabajo	39
Figura 19	Asignación de salidas por tipo de equipo	41
Figura 20	Tiempos medio de clasificación por SKU en cada salida según el tipo de equipo PTL	42

Figura 21 Confirmaciones totales por salida en cada tipo de equipo PTL	43
Figura 22 Cantidad de confirmaciones de cada salida posteriores a la lectura del SKU	44
Figura 23 Tiempos de clasificación por SKU de cada operario	46
Figura 24 Matriz de correlaciones de variables	47
Figura 25 Diagrama de dispersión entre las cantidades de SKU clasificadas y el tiempo de clasificación	48
Figura 26 Diagrama de dispersión 3D entre el tiempo total de clasificación, la cantidad de SKU clasificados y el peso del SKU	49
Figura 27 Comportamiento en la distribución de los tiempos de clasificación (1) y tiempos de clasificación por SKU (2)	50
Figura 28 Agrupación de los salidas del equipo PTL tipo 1 basado en las comparaciones de medias	52
Figura 29 Salidas del equipo PTL tipo 1 categorizadas por productividad en el tiempo de clasificación	54
Figura 30 Agrupación de los salidas del equipo PTL tipo 0 basado en las comparaciones de medias	55
Figura 31 Salidas del equipo PTL tipo 0 categorizadas por productividad en el tiempo de clasificación	58
Figura 32 Agrupación de los operarios basado en las comparaciones de medias	62

Siglas, acrónimos y abreviaturas

ANN	Artificial Neural Networks
Bagging	Bootstrap Aggregation
CEDI	Centro de Distribución
CRISP-DM	Cross Industry Standard Process for Data Mining
DTR	Decision Tree Regressor
EDA	Exploratory Data Analysis
ERP	Enterprise Resource Planning
KNN	K-Nearest Neighbors
PTL	Put-to-Light
RFR	Random Forest Regressor
SKU	Stock Keeping Unit
SVM	Support Vector Machine
WMS	Warehouse Management System

Resumen

El presente estudio tiene como objetivo analizar y clasificar la productividad de los operarios y la eficiencia operativa de los equipos *Put-to-Light* de un centro logístico de mercancía textil, utilizando un enfoque basado en análisis descriptivo y en pruebas estadísticas no paramétricas. Los sistemas *Put-to-Light* son una tecnología clave en la optimización del proceso de clasificación de productos, pero su rendimiento puede verse afectado por diversos factores como el tipo de equipo, el turno de trabajo, la experiencia de los operarios y la asignación de productos a las salidas del sistema. Se analizaron datos operativos de más de un año, aplicando técnicas de minería de datos inspirado en el modelo CRISP-DM. Los resultados revelan que existen diferencias significativas en los tiempos de clasificación según el operario, así como una relación entre la productividad y variables como el turno de trabajo y la ubicación de los productos en las salidas del sistema *Put-to-Light*. Estas conclusiones ofrecen una base para mejorar la gestión de los recursos humanos y optimizar la asignación de tareas, permitiendo un uso más eficiente de la tecnología *Put-to-Light* y, en última instancia, una mayor productividad operativa.

Palabras clave: equipos *Put-To-Light*, productividad, clasificación, pruebas no paramétricas, análisis de datos.

Abstract

The present study aims to analyze and classify the productivity of operators and the operational efficiency of *Put-to-Light* systems in distribution center of the apparel industry, using a descriptive analysis approach and non-parametric statistical tests. *Put-to-Light* systems are a key technology in optimizing the product classification process, but their performance can be affected by several factors such as the type of equipment, work shift, operator experience, and the assignment of products to the system's slots. Operational data from over a year was analyzed, applying data mining techniques inspired by the CRISP-DM model. The results reveal significant differences in classification times depending on the operator, as well as a relationship between productivity and variables such as work shift and the placement of products in the *Put-to-Light* system's slots. These conclusions provide a foundation for improving human resource management and optimizing task allocation, allowing for more efficient use of *Put-to-Light* technology and greater operational productivity.

Keywords: *Put-to-Light* systems, productivity, classification, non-parametric tests, data analysis.

Introducción

Los Centros de Distribución (CEDI) juegan un rol crucial en la cadena de suministro, actuando como el núcleo que conecta a los productores con los consumidores. Estos centros son responsables de recibir, almacenar, preparar y despachar mercancías; tareas que requieren una alta eficiencia operativa, especialmente en un contexto de demanda creciente impulsada por el auge del comercio electrónico. Con la evolución del comercio y las necesidades logísticas modernas, la implementación de tecnologías que optimicen estos procesos se ha convertido en una prioridad para las empresas que buscan reducir tiempos de entrega y maximizar la satisfacción del cliente.

Una de las tecnologías más utilizadas para mejorar la eficiencia en los procesos de clasificación dentro de los CEDI es el sistema *Put-to-Light* (PTL), que guía a los operarios mediante luces que indican dónde deben ubicar o recoger productos. A pesar de sus ventajas evidentes, la productividad en el uso de estos sistemas puede variar considerablemente según múltiples factores, como el tipo de equipo, el turno de trabajo, la experiencia de los operarios, y la distribución de las cargas de trabajo. Por tanto, resulta fundamental analizar con profundidad los factores que afectan la eficiencia operativa en los equipos PTL, lo que permitirá a los tomadores de decisiones optimizar las operaciones y mejorar el rendimiento general.

Este estudio tiene como objetivo principal identificar y segmentar las productividades de los operarios en equipos PTL mediante el uso de técnicas estadísticas no paramétricas. La metodología utilizada sigue el modelo CRISP-DM, ampliamente reconocido en el análisis de datos, y permite un enfoque iterativo que abarca desde la adquisición de datos hasta la evaluación de los resultados. Se utilizarán bases de datos de las transacciones de los equipos PTL de un centro logístico con mercancía del sector textil, con registros que abarcan un periodo de un año, lo que proporcionará una visión sobre los tiempos de clasificación y la eficiencia de los operarios.

A lo largo de este documento, se expondrán los resultados de un análisis detallado de los datos, en el que se examinan diversas dimensiones operativas clave para comprender mejor el comportamiento y la eficiencia de los sistemas *Put-to-Light* (PTL). El análisis incluye tanto un enfoque exploratorio de los datos recolectados como la aplicación de pruebas estadísticas no paramétricas que permitirán inferir diferencias significativas en la productividad entre diversos grupos de operarios, tipos de equipos PTL, turnos de trabajo, días laborados, y hasta en el nivel de detalle de las posiciones específicas del sistema PTL (salidas) donde los productos son ubicados.

Con estos hallazgos, se espera proporcionar herramientas y recomendaciones prácticas para mejorar el uso de los equipos PTL y optimizar la gestión de los recursos humanos y tecnológicos en los CEDI. Finalmente, se concluye con una discusión sobre las implicaciones operativas de los resultados y las oportunidades para futuras investigaciones en torno a la eficiencia de sistemas PTL.

Planteamiento del problema

Los Centros de Distribución (CEDI) son infraestructuras logísticas clave para realizar operaciones de recepción, almacenamiento, preparación y distribución de mercancías hacia otros actores de la cadena de suministro (Shiau & Lee, 2010). Debido a su rol central, la eficiencia operativa de los CEDI influye directamente en la eficiencia de toda la cadena logística (Hamdy et al., 2022). Por esta razón, cuando un CEDI maneja un amplio portafolio de clientes y grandes volúmenes de pedidos con tiempos de entrega ajustados, es necesario implementar sistemas que optimicen operativamente el proceso de preparación de pedidos (Shiau & Lee, 2010). Los sistemas *pick-to-light* y *put-to-light* son esenciales en estos casos, ya que, mediante el encendido de una luz al leer el código de un producto, se guía a los operarios del CEDI para identificar dónde recoger el producto de las estanterías o clasificarlo en las posiciones correspondientes (Ardjmand et al., 2019; Boysen et al., 2020).

Los equipos *put-to-light* (PTL) tienen la capacidad de generar métricas de productividad, permitiendo a los coordinadores de los CEDI conocer el estado de clasificación de las olas de pedidos actuales y tener acceso a información histórica. Sin embargo, la operación inadecuada de los PTL puede generar métricas poco fiables que impactan en la programación de tareas de clasificación, afectando la eficiencia operativa de los PTL y distorsionando las métricas necesarias para una gestión precisa. Investigaciones como las de Ardjmand et al. (2019) y Boysen et al. (2020), han explorado el impacto de los sistemas PTL en la eficiencia de los CEDI. Estos estudios destacan la importancia de la configuración del *layout*, la integración con sistemas de gestión y las estrategias de operación en la optimización del rendimiento de los equipos PTL.

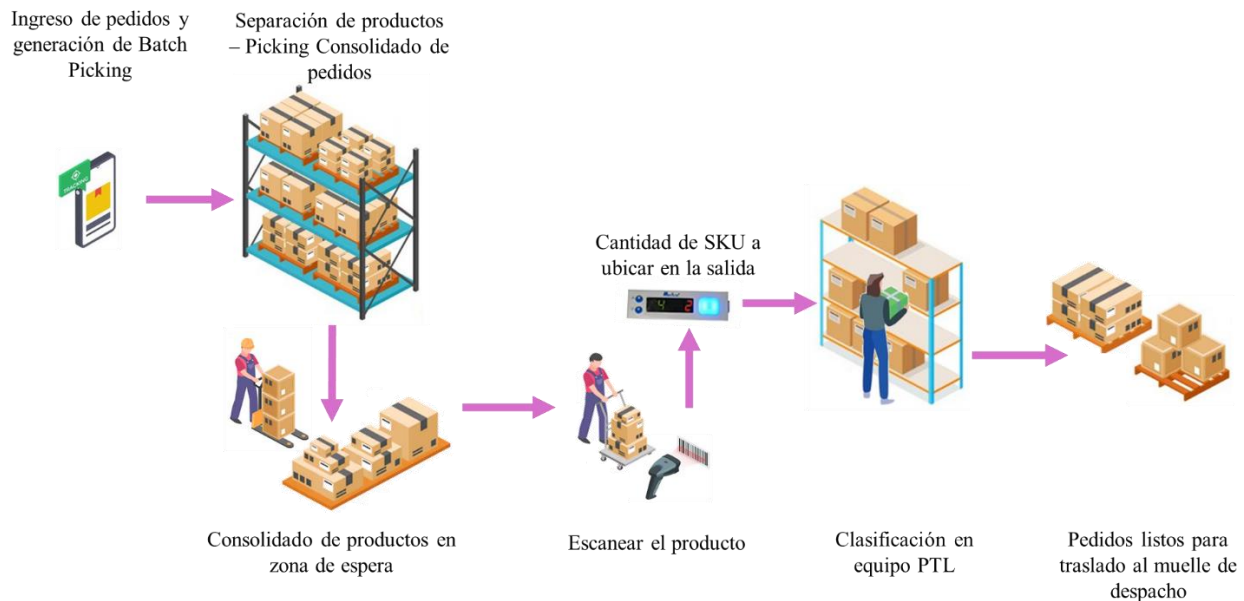
Comprender el comportamiento en las productividades en las tareas de *sorting* dentro de los CEDIs es esencial para mejorar la eficiencia en la operación de equipos PTL. Por lo tanto, el siguiente estudio busca identificar los factores clave que influyen en la productividad de los equipos PTL, proporcionando una comprensión detallada de cómo estos factores inciden en la eficiencia de las operaciones de *sorting*. Esta comprensión permitirá a los tomadores de decisiones y líderes de operaciones mejorar sus procesos y optimizar el uso de los sistemas PTL.

1 Marco Teórico

A continuación, se describe cada uno de los pasos en el proceso de operación del equipo PTL, ilustrado en la Figura 1. El proceso de operación en un Centro de Distribución (CEDI) comienza con el **Ingreso de Pedidos y Generación de *Batch Picking u Ola***. En esta fase, los pedidos de los clientes son recibidos y registrados en el sistema de gestión ERP (*Enterprise Resource Planning*) o WMS (*Warehouse Management System*). Posteriormente, se agrupan múltiples pedidos en olas (*batch*) para optimizar el proceso de recolección de los productos (Muvum, s.f.; Lightning Pick, s.f.). La siguiente etapa es la **Separación de Productos mediante el *Picking Consolidado***. En esta fase, los operarios del CEDI seleccionan los productos de las estanterías según las olas generadas previamente. Este proceso consiste en recoger los productos necesarios para cumplir con los pedidos agrupados, asegurando que todos los artículos requeridos estén listos para la siguiente etapa del proceso (Muvum, s.f.; AutoStore, s.f.).

Figura 1

Proceso de operación del equipo PTL



Nota. Basado en la documentación de Muvum Corp.

Una vez que los productos han sido recogidos, se procede al **Consolidado de Productos en la Zona de Espera**. Aquí, los productos recogidos se agrupan en una zona designada, conocida

como pulmón de espera. En esta área, los productos se organizan y preparan para ser clasificados en las ubicaciones finales. La **Clasificación en el Equipo Put-to-Light (PTL)** es una etapa crítica del proceso. En este punto, los productos se clasifican y colocan en las salidas correspondientes utilizando un sistema que emplea señales luminosas para guiar a los operarios. El proceso comienza con el escaneo de uno de los productos que se deben clasificar. Al escanear el producto, las luces en las salidas de destino se encienden, indicando a los operadores dónde deben depositar cada artículo. Los operadores luego colocan las unidades indicadas del producto en las salidas con las luces encendidas y confirman la acción, presionando un botón asociado con la luz, asegurando así la correcta clasificación de los productos (Muvum, s.f.; Lightning Pick, s.f.).

Finalmente, en la etapa de Pedidos Listos para **Llevar a Muelle de Despacho**, cuando todos los productos se clasifican y colocan en las salidas correspondientes, los pedidos completos se retiran y empaquetan en su unidad de embalaje para trasladarlos al muelle de despacho. En este punto, los pedidos ya organizados y completos se trasladan al área de despacho para ser enviados a los clientes, concluyendo así el proceso de preparación de pedidos en el CEDI (Muvum, s.f.).

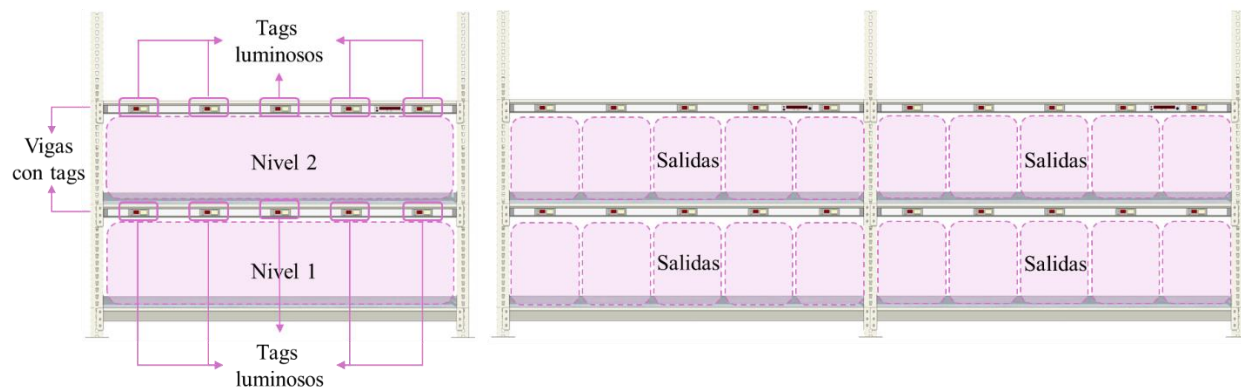
Definiciones

Es fundamental comprender varios conceptos clave que hacen parte de la composición física de los equipos PTL y su operación. La **SKU** (*Stock Keeping Unit*) es el identificador único que permite rastrear cada variación de producto desde su entrada hasta su salida del almacén (Association for Supply Chain Management, 2022). Las SKU se organizan dentro de los **Niveles**, los cuales son los espacios entre las vigas que componen un **Módulo** de estantería. Los **Módulos** son columnas de estanterías divididas por niveles, mientras que una **Zona** se refiere a un módulo de estantería o a la combinación de dos o más módulos. El *layout* del equipo PTL dependerá de los requerimientos operativos y del espacio disponible en el CEDI. Las **Salidas**, cada una identificada por un tag luminoso, son las posiciones específicas dentro del equipo PTL donde se colocan las SKU que forman el pedido de cada cliente. De este modo, se puede considerar que cada salida representa el pedido de un cliente dentro de una **Ola** de pedidos. Por último, las **Olas** (*Batch*) son grupos de pedidos que se procesan juntos (Association for Supply Chain Management, 2022). En la Figura 2(izquierda) se muestra un equipo con un solo cuerpo de estantería, mientras que la Figura

2(derecha) muestra una zona con dos cuerpos de estantería. En esa Figura se resaltan los componentes de un equipo PTL.

Figura 2

Partes de un equipo PTL. (a) Zona con un solo módulo de estantería. (b) Zona con dos módulos de estantería



Nota. Basado en la documentación de Muvum Corp.

2 Estado del arte

El estudio de la productividad en centros logísticos ha sido ampliamente investigado debido a la necesidad de optimizar la eficiencia operativa. El crecimiento acelerado del comercio electrónico ha presentado nuevos desafíos, ya que los CEDI deben gestionar un alto flujo de órdenes de pedidos (La República, 2023). Esta situación exige soluciones que no solo mejoren la productividad, sino que también se adapten rápidamente a la creciente demanda y la urgencia asociada con los pedidos en línea. Diferentes enfoques han sido explorados para modelar y predecir tiempos de ciclo y productividad, utilizando herramientas avanzadas como redes neuronales y modelos de *machine learning*.

Janikova y Bezak (2016), utilizaron redes neuronales para estimar la productividad en líneas de producción, obteniendo resultados que demostraron un rendimiento predictivo satisfactorio. Esta investigación confirmó la efectividad de las redes neuronales en la predicción de la productividad, ampliando el uso de estas técnicas en diferentes contextos industriales. Más recientemente, Krishna et al. (2022), exploraron la estimación del tiempo necesario para ejecutar tareas en líneas de ensamble como una medida para evaluar el rendimiento de los trabajadores. Este estudio incorporó variables como la experticia, experiencia, habilidad, edad y género, ofreciendo un enfoque más holístico para diferenciar el rendimiento entre operarios. Esta investigación destacó la importancia de considerar factores humanos en la evaluación de la productividad, lo que puede ser crucial en la implementación de sistemas PTL.

En el ámbito de la planeación de la producción, Atik et al. (2021), implementaron modelos de *machine learning* como *Support Vector Machine* (SVM), *Decision Tree Regressor* (DTR), *Bootstrap Aggregation* (Bagging) y *Random Forest Regressor* (RFR) para estimar tiempos de producción en la industria textil. Considerando factores como el tamaño, color y tipo de tela, los modelos Bagging y RFR demostraron un desempeño superior, con coeficientes de determinación (R^2) superiores a 0.84 y errores absolutos medios (MAE) menores que los obtenidos con otros modelos. Además, se desarrolló una interfaz gráfica para facilitar la entrada de datos y la generación de predicciones, apoyando así la toma de decisiones en la planeación de la producción.

Al-Radaideh & Al Nagi (2012), construyeron un modelo de clasificación para predecir el rendimiento de los empleados con árboles de decisión y el clasificador *Naïve Bayes*. Los autores encontraron que el título del cargo era el atributo más influyente en la predicción del rendimiento

de los empleados, seguido por el tipo de universidad de graduación. Estos hallazgos subrayan la importancia de las responsabilidades laborales y la formación académica en el desempeño de los empleados, ofreciendo una visión que podría extrapolarse al análisis de la productividad en sistemas PTL, donde la correcta asignación de tareas y la capacitación pueden jugar roles determinantes (Al-Radaideh & Al Nagi, 2012).

Complementariamente, Sarker et al. (2018), exploraron el uso de algoritmos de agrupamiento como *K-means* y árboles de decisión para analizar y predecir el rendimiento de los empleados. Este enfoque híbrido permitió segmentar a los empleados en cuatro grupos según su rendimiento: *Excelente*, *Bueno*, *Promedio* y *Pobre*; lo que facilitó la identificación de grupos que requerían intervenciones específicas, como capacitación adicional o posibilidades de promoción. Este tipo de análisis es particularmente relevante para la gestión de sistemas PTL, donde la capacidad de identificar rápidamente a los empleados con mejor rendimiento y asignarles tareas críticas puede optimizar significativamente el flujo de trabajo y la eficiencia operativa. Sabuj et al. (2023), analizaron la productividad de los trabajadores en la industria textil en Bangladesh. Utilizando modelos de regresión lineal y árboles de decisión, los autores lograron identificar factores críticos como la experiencia laboral y la carga de trabajo, que resultaron ser determinantes en la productividad de los empleados. La capacidad de interpretar fácilmente los resultados de los modelos fue clave para la implementación efectiva de mejoras operativas.

Loske et al. (2023) investigaron el impacto de los sistemas de almacenamiento en el tiempo de preparación de pedidos en almacenes logísticos. Este estudio empírico comparó los sistemas de almacenamiento de pallets completos y de alto flujo, descubriendo que los sistemas de alto flujo aceleran el proceso de picking en un 4.60%. Sin embargo, también identificaron un aumento en los tiempos de reabastecimiento del 38.65%, lo que destaca la necesidad de equilibrar ambos factores al elegir un sistema de almacenamiento (Loske et al., 2023). Este hallazgo es relevante para la optimización de la productividad en sistemas PTL, ya que subraya la importancia de evaluar tanto el proceso de picking como el reabastecimiento.

Finalmente, se ha investigado cómo un enfoque imparcial basado en algoritmos de aprendizaje automático puede predecir de manera justa el rendimiento de los empleados al considerar factores físicos, sociales y económicos que afectan su desempeño. Este estudio utilizó varias técnicas de aprendizaje automático, incluyendo la regresión logística, *Naïve Bayes*, árboles de decisión, *K-Nearest Neighbors* (KNN), SVM, y RFR. El modelo RFR alcanzó la mayor

precisión, lo que lo convierte en una herramienta poderosa para evaluar el rendimiento de los empleados de manera objetiva. La implementación de un marco como este en los CEDIs podría ayudar a evitar sesgos en la evaluación del rendimiento, asegurando que las decisiones de promoción y capacitación se basen en datos objetivos y multifactoriales (Nayem & Uddin, 2024).

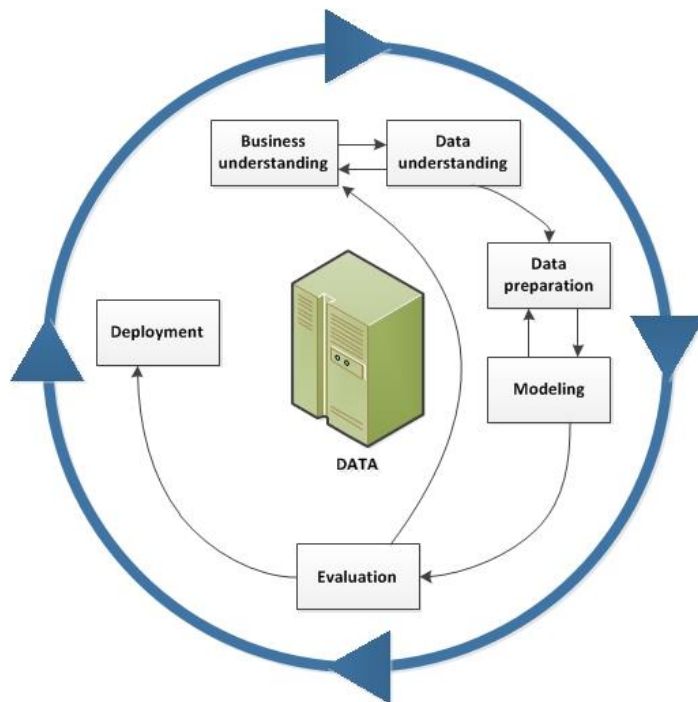
Estos estudios son altamente relevantes para la investigación sobre el comportamiento en las productividades de operación de *picking* y *sorting* en equipos PTL, ya que proporcionan una base sólida sobre el impacto de la tecnología y los sistemas de almacenamiento en la eficiencia operativa. El uso de modelos de machine learning y técnicas de predicción aplicados en diversas industrias demuestra cómo estas herramientas pueden optimizar el rendimiento en el contexto logístico. Además, la consideración de factores humanos y las diferencias entre operarios también será esencial en el análisis de la productividad de los sistemas PTL, tal como se sugiere en las investigaciones revisadas. Por lo tanto, estos enfoques y hallazgos complementan y amplían el análisis que se pretende realizar en el estudio, contribuyendo a una visión más integral sobre los factores que afectan la eficiencia operativa en centros de distribución.

3 Metodología

Para el desarrollo de este trabajo de grado, se adoptó un enfoque metodológico inspirado en la estructura de la *Cross Industry Standard Process for Data Mining (CRISP-DM)*, un modelo ampliamente reconocido y utilizado en proyectos de minería de datos y análisis *end-to-end* (Tirkel, 2011; Janikova & Bezak, 2016). Este enfoque iterativo y flexible, ver Figura 3, permitió guiar de manera estructurada todo el proceso de investigación, desde la recolección de datos hasta el hallazgo de información que permita comprender la operación de los equipos PTL a profundidad.

Figura 3

Fases del ciclo de vida de proyectos entorno a minería de datos aplicando la metodología CRISP-DM



Nota: Conceptos básicos de ayuda de CRISP-DM. IBM.

La metodología implementada sigue las principales etapas de CRISP-DM, las cuales fueron adaptadas al contexto de esta investigación: 1) comprensión del negocio, 2) entendimiento de los datos, 3) preparación de los datos, 4) modelación y 5) evaluación.

Fase 1. Identificación del problema y objetivos

Los sistemas *put-to-light* son herramientas críticas que optimizan el proceso de clasificación de pedidos, proporcionando métricas de productividad que guían las operaciones diarias; no obstante, una operación deficiente de los PTL puede generar datos inexactos, afectando la gestión operativa y la programación de las tareas de clasificación. Por ello, el objetivo de esta investigación es identificar los factores clave que impactan en la productividad de estos equipos y su relación con la eficiencia operativa en las tareas de clasificación.

Fase 2. Adquisición y preparación de los datos

Entendimiento del conjunto de datos

Una empresa dedicada al desarrollo e implementación de sistemas PTL suministró los datos de clasificación en el área de preparación y despacho de uno de sus clientes perteneciente al sector textil. En la Tabla 1, se realiza la descripción de los dos archivos *.txt recibidos, los cuales contienen registros de operación entre octubre de 2022 a octubre de 2023.

Tabla 1

Datos suministrados por Muvum Corp.

Dataset	Descripción	Campos	Registros	Tamaño archivo
Transacciones.txt	Contiene la información de operación en equipos PTL, registrando la fecha y hora cuando el trabajador realiza la primera lectura de un SKU de una Ola en una Zona.	NoOla: integer Zona: integer IdTrabajador: integer SKU: integer FechaLectura: string	4.265.663	207,64 MB

Lotes.txt	Contiene la información de operación en equipos PTL, registrando la fecha y hora cuando el trabajador realiza la confirmación de la clasificación de determinada cantidad de SKU de una Ola en las Salidas de una Zona.	NoOla: integer NoPedido: integer Zona: integer SalidaPedido: string SKU: integer DescripciónSKU: string PesoSKU: float Cantidad: integer FechaInicioOla: string FechaConfirmaciónSKU: string FechaFinalizaciónOla: string IdTrabajador: integer	6.959.870	1.225,33 MB
------------------	---	--	-----------	-------------

Con los *datasets* Transacciones y Lotes se tienen atributos con los cuáles se pueden realizar análisis de operación y extraer conocimientos valiosos que den respuesta a la problemática presentada y alcanzar los objetivos propuestos. Además, para cada *dataset* se poseen más de cuatro millones de registros, lo cual permite tener una muestra generosa de datos para realizar análisis y construir modelos de *machine learning* más confiables.

Esta etapa también comprendió la ejecución de actividades como:

- Revisión del tamaño del conjunto de datos: filas y columnas.
- Identificar el tipo de dato de cada atributo.
- Extraer estadísticas de posición y dispersión, para entender el comportamiento de los datos.
- Identificar valores duplicados, nulos o anomalías; y discutir estrategias de imputación o depuración.

Preparación de los datos

Después de realizar una exploración básica de los datos; se evidencia que existe un formato inadecuado en el tipo de variable para algunos atributos. Además, debido a la gran cantidad de

registros, la cantidad de atributos y el tipo *.txt de los archivos; influye en el tamaño del archivo para su almacenamiento en disco, haciendo a su vez menos eficiente su carga, lectura, procesamiento y guardado. En esta etapa se llevó a cabo la conversión de los *datasets* a archivo tipo *.parquet. Este es uno de los formatos de archivo más eficientes y usados en *big data* para cargar, leer y guardar información. También se procedió a revisar, imputar y/o depurar valores nulos y duplicados, según sea el caso de la variable. Además de convertir el tipo de dato de cada atributo según sea su naturaleza.

A partir de los *datasets* recibidos se crearon dos tablas adicionales aplicando reglas de normalización de bases de datos, las cuales buscan reducir la cantidad de atributos por tabla, evitar redundancias que inciden en el tamaño del archivo, reducir la complejidad de las tablas y establecer un modelo relacional entre cada una de las tablas. Se crearon las tablas Maestro de SKU: contiene información con el código de SKU, descripción, peso y familia del SKU; y Maestro de Zonas: contiene información sobre cómo se componen las zonas, las salidas por zona, la numeración de cada salida y el tipo de zona según la cantidad de salidas. De esta forma, se puede depurar los atributos PesoSKU y DescripciónSKU de la tabla Lotes, ya que estos pueden ser llamados a la tabla Lotes mediante la clave SKU de la tabla Maestro de SKU.

Por último, es fundamental determinar el tiempo requerido para que cada operario clasifique los SKUs de un pedido en el sistema *put-to-light*. Este valor es esencial para medir el rendimiento del operario en la clasificación de productos y puede calcularse a partir de dos fechas clave:

- Fecha de Lectura, presente en la tabla de Transacciones.
- Fecha de Confirmación del SKU, presente en la tabla de Lotes.

Para comprender la diferencia entre ambas fechas, es fundamental comprender que el proceso de clasificación en el sistema *put-to-light* puede descomponerse en dos etapas temporales:

Tiempo entre Lectura-Confirmación: Es el tiempo que transcurre desde que se realiza la lectura del SKU hasta la primera confirmación de las cantidades solicitadas del SKU en una de las salidas que se han iluminado en el sistema *put-to-light*. En la Figura 4 se muestra la secuencia de leer el SKU, ubicarlo en una de las salidas que se han iluminado y confirmar su colocación en la salida; el tiempo que se está mencionando considera desde la lectura del SKU hasta su confirmación en la salida.

Figura 4

Acción de lectura y confirmación de SKU

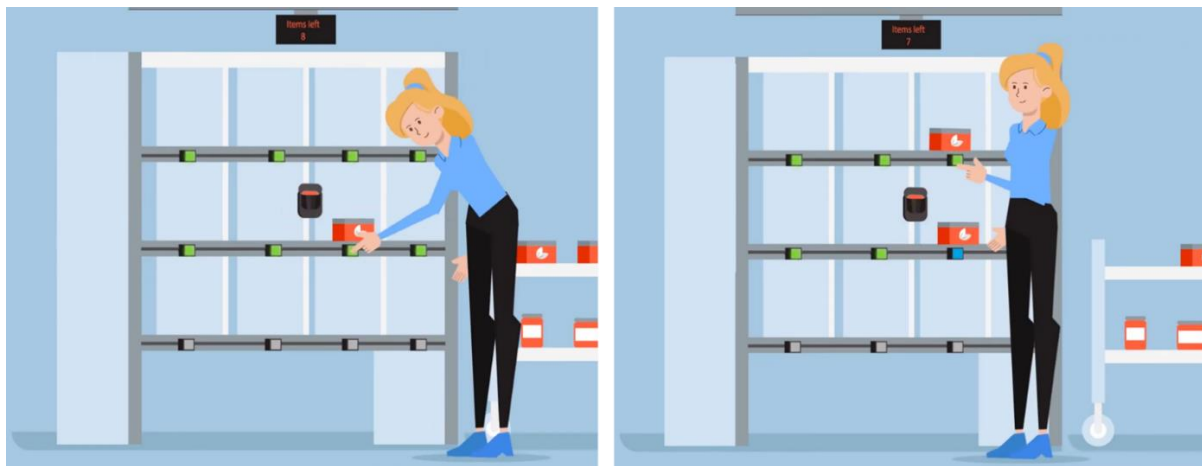


Nota: Putwall solution animation. Pcdeta Logistics Automation

Tiempo entre Confirmación-Confirmación: Después de la primera confirmación, es el tiempo que transcurre entre confirmaciones sucesivas de distintos SKUs dada la cantidad de salidas iluminadas en el equipo como se puede observar en la Figura 5, donde el operario va confirmando en cada salida cuando ubica el SKU. Por lo tanto, es el tiempo que hay entre las acciones de confirmar el SKU en la respectiva salida.

Figura 5

Acción de confirmación y confirmación de SKU



Nota: Putwall solution animation. Pcdeta Logistics Automation

Considerando estas dos etapas se puede determinar la métrica `time_sort`, la cual permite conocer el tiempo transcurrido en el proceso de clasificación. A continuación, en la Tabla 2 se describe el algoritmo utilizado para el cálculo de `time_sort`.

Tabla 2

Algoritmo para cálculo de time_sort

1	Inicio
2	Cargar tablas <i>Transacciones, Lotes</i>
3	Calcular Intersección en el <i>NoOlas</i> entre ambas tablas
4	Generar clave única $\{NoOla, Zona, IdTrabajador, SKU\}$
5	Ordenar <i>Lotes</i> por <i>FechaConfirmaciónSKU</i> Ascendente
6	Ordenar <i>Transacciones</i> por <i>FechaLectura</i> Ascendente
7	Unir <i>FechaLectura</i> de <i>Lotes</i> con <i>Transacciones</i> Como MergeDate #basado en la clave
8	<i>única y emparejamiento de las fechas</i>
9	Si <i>MergeDate</i> <> Vacío:
10	Calcular $time_sort = FechaConfirmaciónSKU - FechaLectura$
11	Sino
12	Calcular $time_sort = FechaConfirmaciónSKU_i - FechaConfirmaciónSKU_{i-1}$
13	Fin Si
14	Validar No haya registros duplicados
15	Validar No haya valores negativos en <i>time_sort</i>
16	Exportar “ <i>data_joined.csv</i> ”
17	Fin

Al finalizar esta etapa se obtuvo como producto el *dataset* “*data.csv*” el cual se usará para realizar análisis exploratorio de los datos, realizar inferencias y plantear modelos que permitan hallar información valiosa.

Fase 3. Análisis Exploratorio de los datos e inferencias

En el análisis exploratorio de datos (EDA), se utilizó la métrica *time_sort_per_sku*, calculada como el tiempo total de clasificación (*time_sort*) dividido entre la cantidad de SKU procesados (*qty_sku*), lo que permitió normalizar los tiempos y hacer comparables las productividades entre los trabajadores. A través de gráficos y medidas estadísticas, se identificaron

patrones y posibles valores atípicos en la eficiencia de los equipos PTL. Por lo que, el cálculo de `time_sort_per_sku` ayudó a identificar tiempos atípicos de manera más clara; eliminando aquellos registros donde el tiempo de clasificación por SKU superaba los dos minutos, ya que esto evidenciaba una operación incorrecta del PTL, siendo dos minutos más que suficiente para clasificar un SKU.

Se realizaron pruebas de diferencias de medias sobre `time_sort_per_sku` para evaluar la existencia de discrepancias significativas entre los tiempos de clasificación de los trabajadores. Posteriormente, se emplearon grafos para agrupar a los trabajadores en función de su productividad, agrupando a aquellos con tiempos estadísticamente similares de clasificación. Los análisis mostraron que no había diferencias significativas entre los tiempos de clasificación de los trabajadores dentro de cada clúster, confirmando una similitud en los patrones de productividad.

4 Resultados

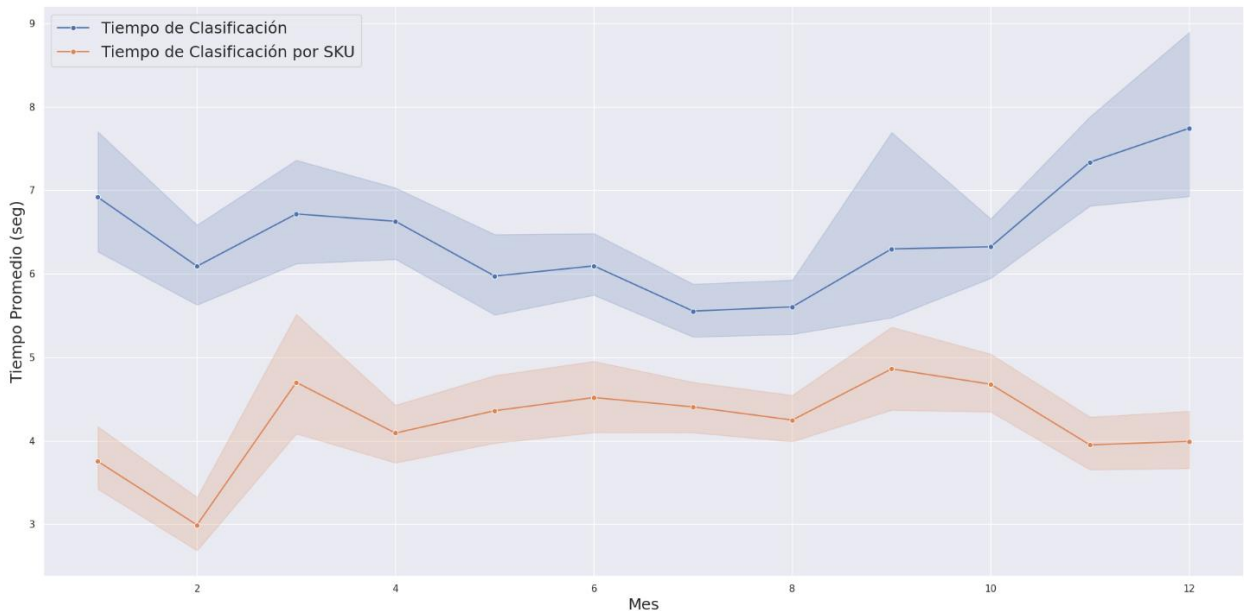
Análisis exploratorio y descriptivo de los datos

El análisis exploratorio de datos (EDA) es una fase esencial para entender el comportamiento y los factores que afectan la productividad de los equipos PTL en los CEDI. La operación de estos equipos es fundamental para la clasificación de productos en olas de pedidos, y cualquier ineficiencia o mal uso puede generar métricas poco fiables que impactan negativamente la programación de tareas. La realización del EDA busca comprender los patrones y tendencias en los tiempos de clasificación y las cantidades de SKU procesadas, con el objetivo de optimizar la operación de los PTL.

El análisis se realizó sobre datos recolectados entre octubre de 2022 y octubre de 2023 y examinando diferentes factores como las variaciones mensuales, diarias, horaria, y por turno, utilizando medidas de tendencia central y dispersión (medias y percentiles). Las métricas evaluadas incluyen el tiempo de clasificación (afectado por la cantidad de SKU clasificados) y el tiempo de clasificación por SKU (normalizado por la cantidad de SKU clasificados).

Tiempos y cantidades clasificadas por mes

En la Figura 6, los tiempos de clasificación total muestran un claro incremento durante los meses de alta demanda. Sin embargo, lo que resulta particularmente llamativo es que, a pesar del aumento en los tiempos totales de clasificación, el tiempo promedio por SKU presenta una disminución notable. Este comportamiento sugiere que, aunque los equipos PTL están procesando un mayor volumen de pedidos, se vuelven más eficientes en el procesamiento individual de SKU. Este fenómeno puede interpretarse como una mejora en los procesos operativos, impulsada posiblemente por la implementación de estrategias para manejar el incremento en la carga de trabajo.

Figura 6*Tiempos de clasificación por Mes*

La correlación de esta información con la Figura 7, que ilustra las cantidades de SKU clasificadas por mes y sus respectivas medidas de dispersión, refuerza esta hipótesis. Durante los meses de noviembre y diciembre, no solo se procesan más SKU, sino que las medidas de dispersión, como el P90, P95 y P99, indican una alta variabilidad en las cantidades procesadas. Esto sugiere que, en ciertos momentos, los equipos PTL enfrentan cargas extremas, que coinciden con los picos observados en los tiempos de clasificación total de la Figura 6. A pesar de esta variabilidad y el incremento en la cantidad de SKU procesados, la disminución en el tiempo promedio de clasificación por SKU sigue siendo un indicio de una mayor eficiencia operativa.

Este aumento en la eficiencia puede explicarse a través de la información proporcionada por la Figura 8, que presenta la disponibilidad de operarios en las tareas de clasificación durante el periodo analizado. Los datos muestran que durante los meses de noviembre y diciembre hay un incremento notable en la cantidad de operarios activos, lo que sugiere una estrategia de refuerzo en la mano de obra para responder a la mayor demanda. La mayor disponibilidad de personal durante estos meses de alta carga permite no solo manejar el aumento en los volúmenes de pedidos, sino también optimizar el tiempo de clasificación por SKU, tal como se refleja en la Figura 6 en la cual hay operarios que solo registran actividad en los meses de mayor demanda.

Figura 7
Cantidades de SKU clasificadas por mes

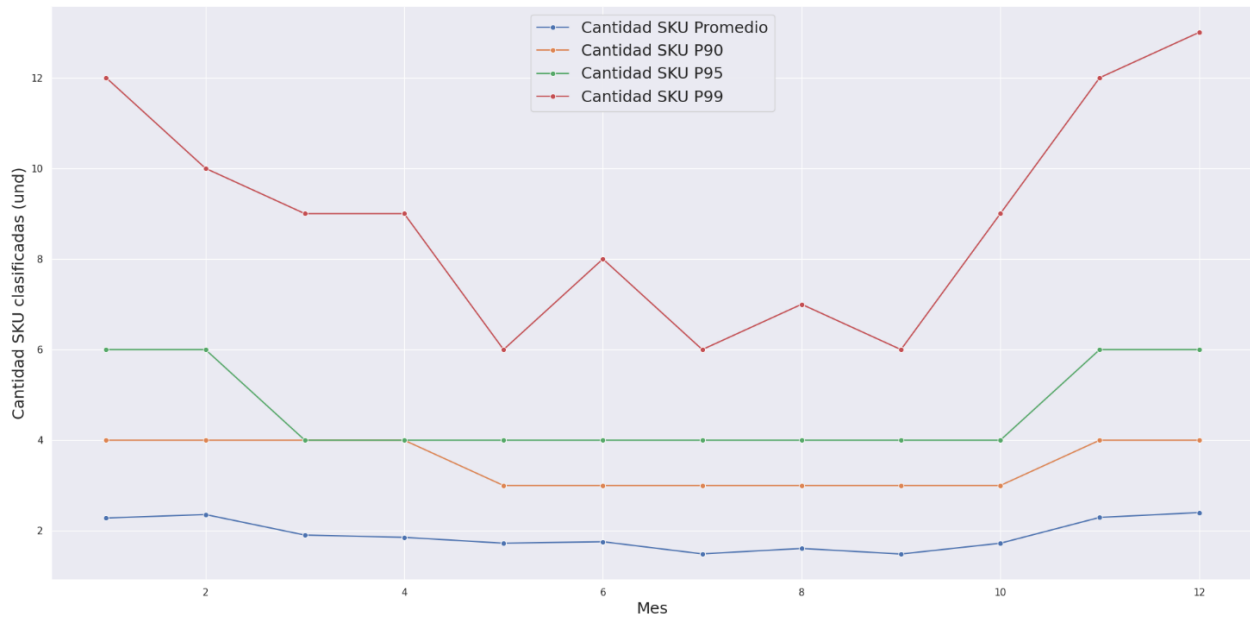
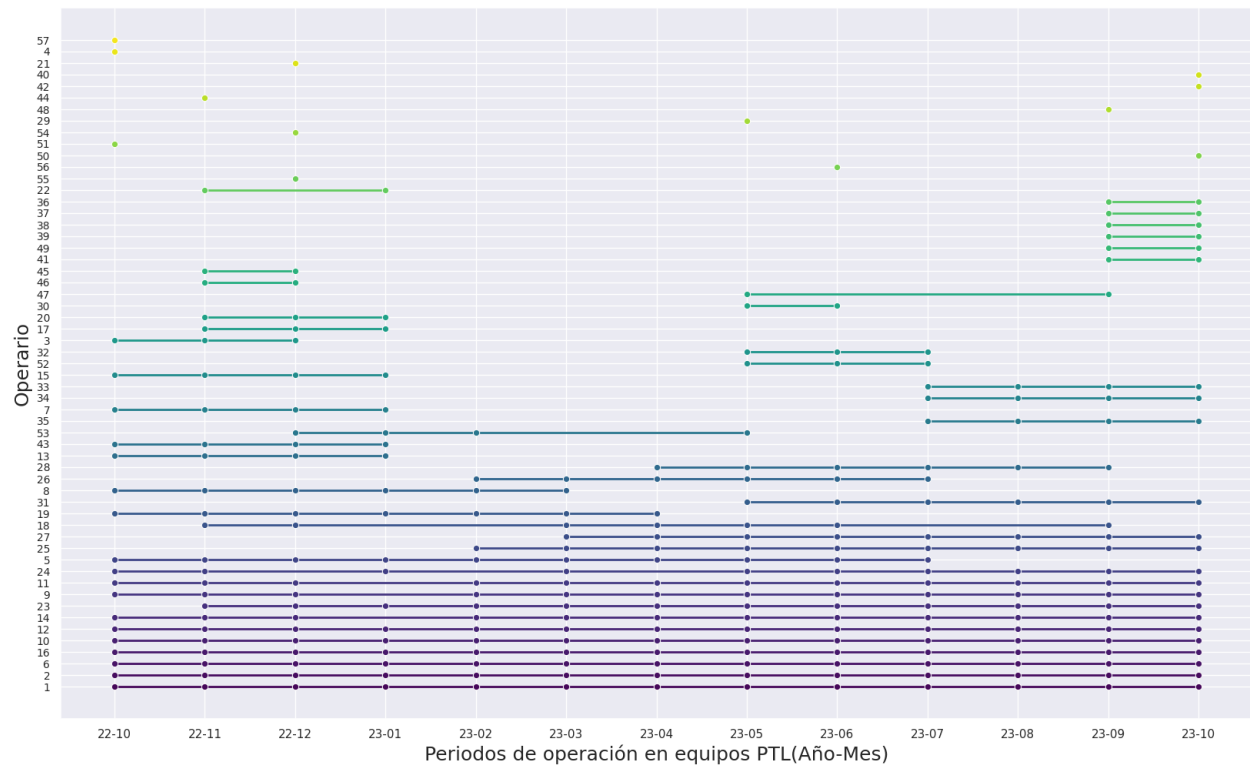


Figura 8
Línea de tiempo de los operarios asignados a tareas de clasificación en equipos PTL



Por otro lado, se observa que varios operarios han estado involucrados en tareas de clasificación de manera continua durante largos periodos, lo que indica una acumulación de experiencia en el equipo de trabajo. Este hecho probablemente ha contribuido a una mejora en la eficiencia, ya que los operarios más experimentados son capaces de procesar los SKU con mayor rapidez y precisión. La combinación de una mayor cantidad de operarios activos y una fuerza laboral más experimentada se puede reflejar en una reducción del tiempo promedio por SKU, incluso cuando las cantidades procesadas alcanzan sus niveles más altos durante los meses de noviembre y diciembre.

Es importante tener en cuenta que, según la información proporcionada por este gráfico y los datos, no se puede determinar con certeza si el hecho de que algunos operarios dejen de operar los equipos PTL se debe a que fueron reasignados a otras tareas dentro del CEDI o si su vinculación laboral terminó después de la temporada de alta demanda. Esta distinción sería relevante para comprender mejor los patrones de contratación y asignación de personal, así como para planificar los recursos humanos de manera más eficiente.

En resumen, el análisis de estos tres gráficos sugiere que los equipos PTL han sido capaces de mejorar su eficiencia operativa frente a un aumento significativo de la demanda, especialmente durante los meses de alta demanda estacional. La estrategia de aumentar la disponibilidad de operarios y el efecto positivo de una fuerza laboral experimentada han permitido reducir los tiempos promedio por SKU, lo que indica una respuesta eficiente y adaptativa a los desafíos operativos presentados por el aumento en los volúmenes de pedidos. Además, la comprensión de los patrones de reasignación o desvinculación del personal podría proporcionar una visión más integral de la gestión de recursos humanos, permitiendo optimizar la planificación de personal en respuesta a fluctuaciones estacionales en la demanda.

Tiempos y Cantidades Clasificadas por Día de la Semana

En la Figura 9, los días miércoles y jueves presentan los tiempos totales de clasificación más altos, lo que refleja una mayor carga operativa. Sin embargo, los jueves y viernes son los días más eficientes en términos de tiempo de clasificación por SKU, a pesar de que el volumen de SKU procesados aumenta hacia el final de la semana, como se muestra en la Figura 10. El ancho de las bandas en la Figura 6 es menor durante los días laborales, lo que indica una mayor consistencia

operativa, mientras que los fines de semana, y en especial los domingos, muestran bandas más anchas, lo que revela una mayor dispersión en los tiempos de clasificación debido a una operación menos predecible.

Figura 9
Tiempos de clasificación por día de la semana

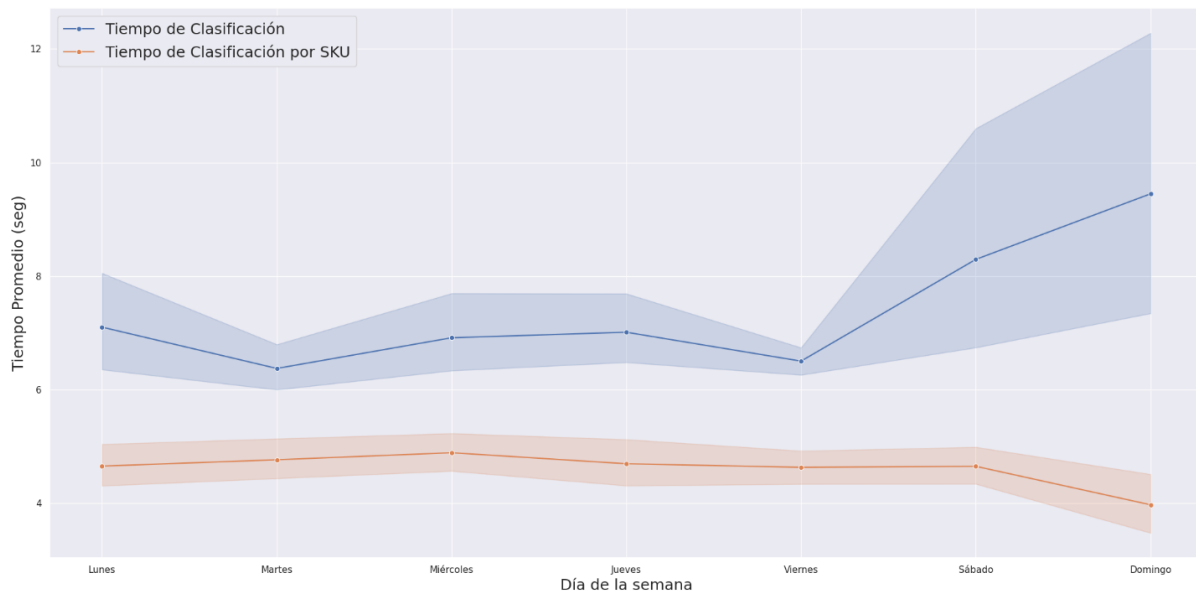
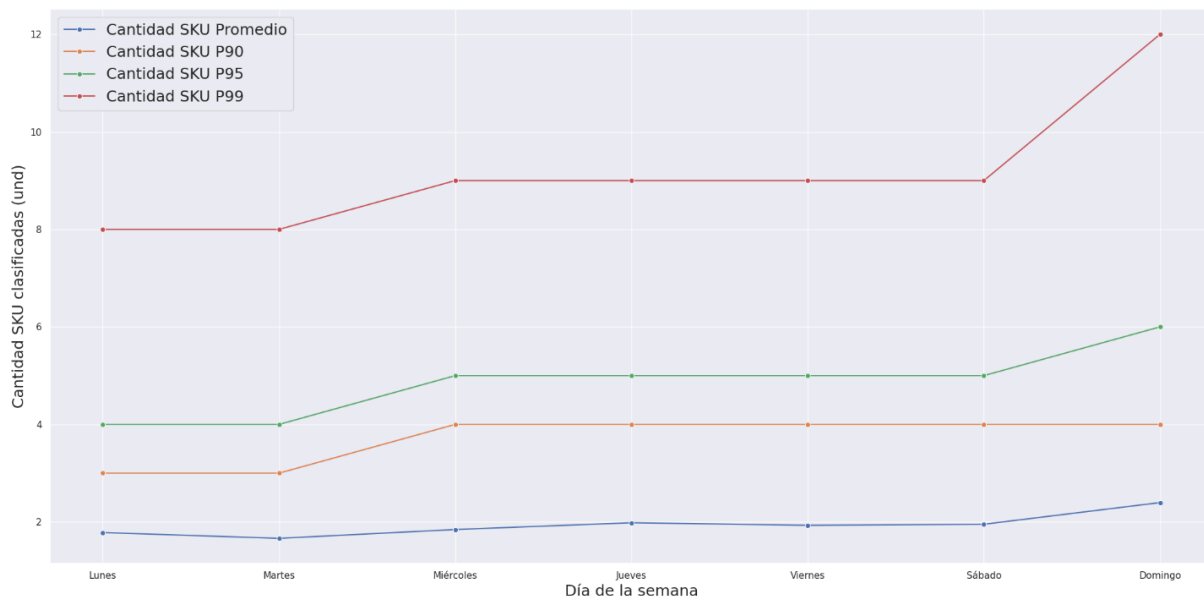


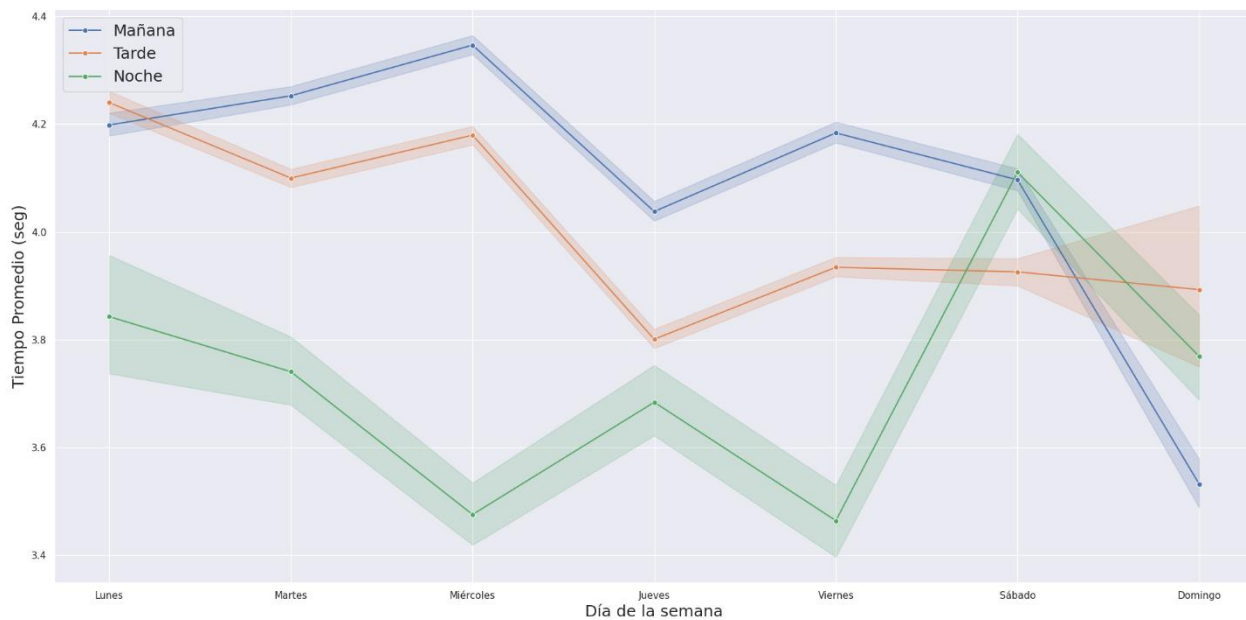
Figura 10
Cantidades de SKU clasificadas por día de la semana



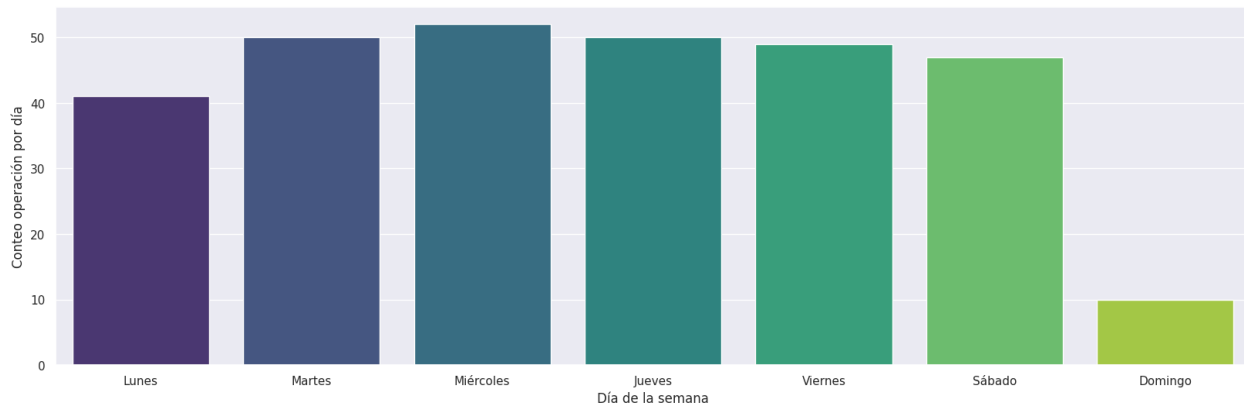
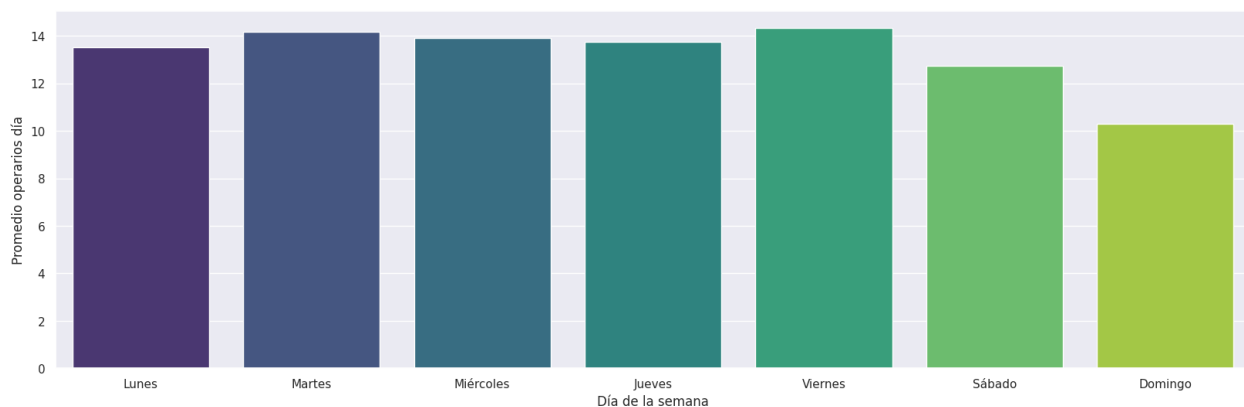
La Figura 11 revela que el turno de noche es el más eficiente en términos de tiempo de clasificación por SKU, con tiempos promedio consistentemente bajos, entre 3.4 y 3.8 segundos, especialmente a mitad de semana. Por el contrario, los turnos de mañana y tarde muestran tiempos más largos, superando los 4 segundos, y con una mayor tendencia a aumentar hacia el final de la semana, en especial los viernes. Además, los fines de semana, particularmente los domingos, presentan una mayor variabilidad en los tiempos de clasificación. En general, el turno de noche demuestra ser más eficiente y estable a lo largo de la semana.

Figura 11

Comportamiento del tiempo de clasificación por SKU según el turno y día de la semana



La frecuencia de uso de los equipos PTL, presentada en la Figura 12, es mayor los martes y miércoles, con alrededor de 50 días operativos, lo que explica los altos tiempos totales en estos días. Sin embargo, esta mayor frecuencia no se traduce en una mayor eficiencia operativa. En la Figura 13, se observa que los días con mayor número de operarios, como martes, miércoles y viernes, están preparados para gestionar la carga de trabajo, aunque la eficiencia operativa es superior en los jueves y viernes, lo que sugiere una mejor organización de los recursos al final de la semana.

Figura 12*Frecuencia de uso de los equipos PTL por día de la semana***Figura 13***Número promedio de operarios por día de la semana*

Por otro lado, los domingos tienen una baja frecuencia de uso y pocos operarios asignados, lo que genera una alta dispersión en los tiempos de clasificación, las bandas más anchas en las Figuras 9 y 11, mostrando la inconsistencia operativa durante los fines de semana. Esto pone de manifiesto la necesidad de mejorar la asignación de recursos para estabilizar la operación durante los días de menor carga.

En general las Figuras 8, 12 y 13 muestran que, aunque los días como martes y miércoles manejan grandes volúmenes de SKU y tienen tiempos de clasificación más altos, no siempre son los más eficientes en términos de tiempo por SKU. Días como jueves y viernes, con una cantidad similar de operarios, logran una mayor eficiencia por SKU, lo que sugiere una mejor organización operativa. Los fines de semana, especialmente los domingos, sufren de alta variabilidad en los

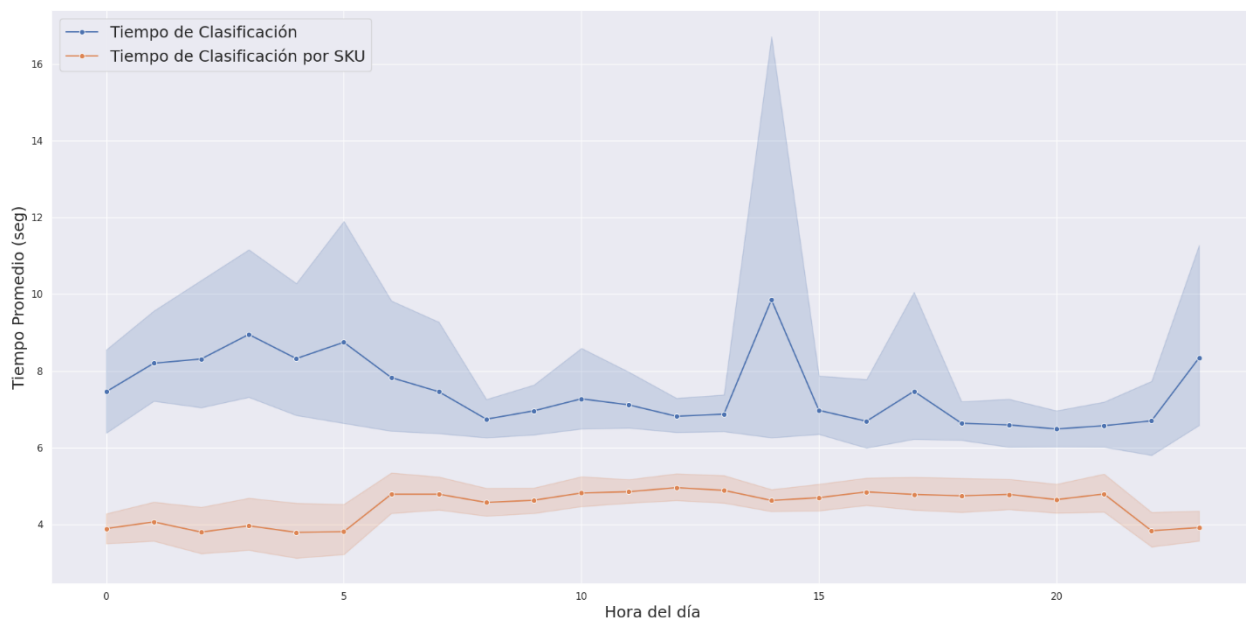
tiempos debido a la baja frecuencia de operación y la menor cantidad de personal, lo que afecta la consistencia y eficiencia en el procesamiento de SKU.

Tiempos y Cantidades Clasificadas por Hora del Día

En la Figura 14, se observa que los Tiempos Totales de Clasificación (línea azul) aumentan durante la noche y la madrugada, alcanzando sus valores más altos entre las 9 p.m. y las 5 a.m. Sin embargo, el Tiempo de Clasificación por SKU (línea naranja) disminuye en estas mismas horas, lo que indica una mayor eficiencia operativa. Este comportamiento refleja que, aunque los tiempos totales sean elevados debido al alto volumen de trabajo, el tiempo por unidad clasificada es más bajo, lo que sugiere que se están clasificando grandes cantidades de SKU en un tiempo relativamente menor. Esta relación es consistente con la Figura 15, que muestra que los mayores volúmenes de SKU se procesan precisamente en estas horas.

Figura 14

Tiempos de clasificación por hora del día

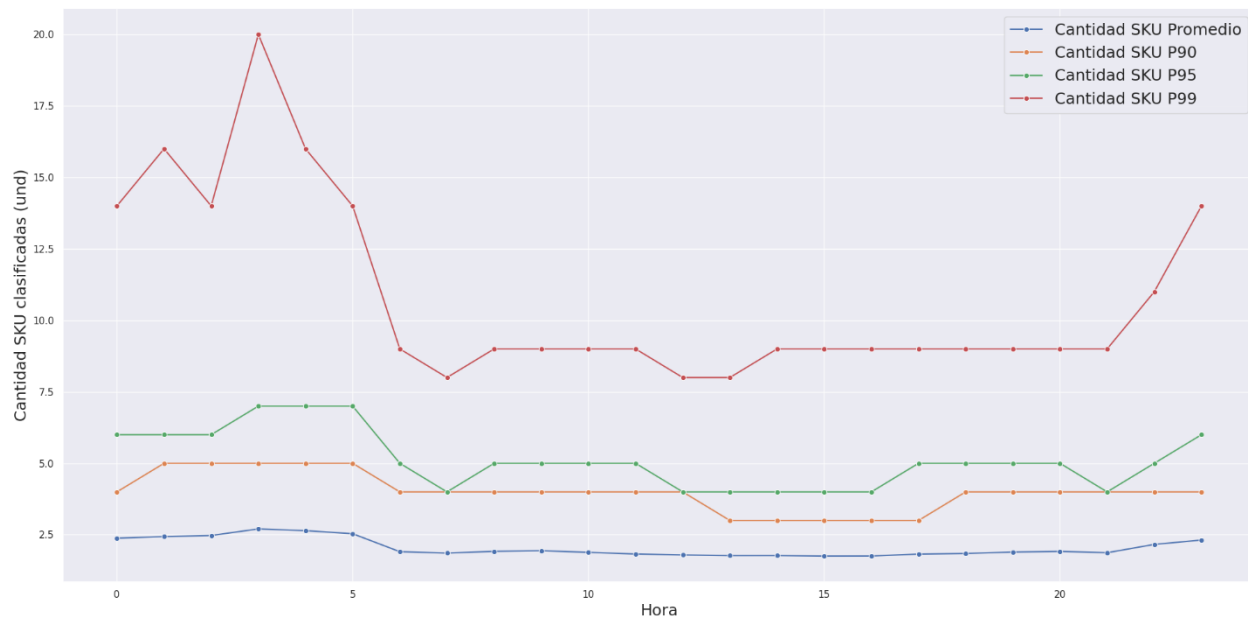


En la Figura 15, las cantidades de SKU clasificadas durante el mismo periodo también se mantienen en un volumen constante. Después de la madrugada, donde se observa el mayor volumen de procesamiento, las cantidades procesadas se estabilizan entre las 5 a.m. y las 8 p.m.,

manteniendo un nivel de trabajo controlado. Aunque las cantidades no son tan altas como en la madrugada, el flujo de SKU es suficiente para mantener una operación eficiente durante la jornada diurna, sin generar congestiones o variaciones drásticas.

Figura 15

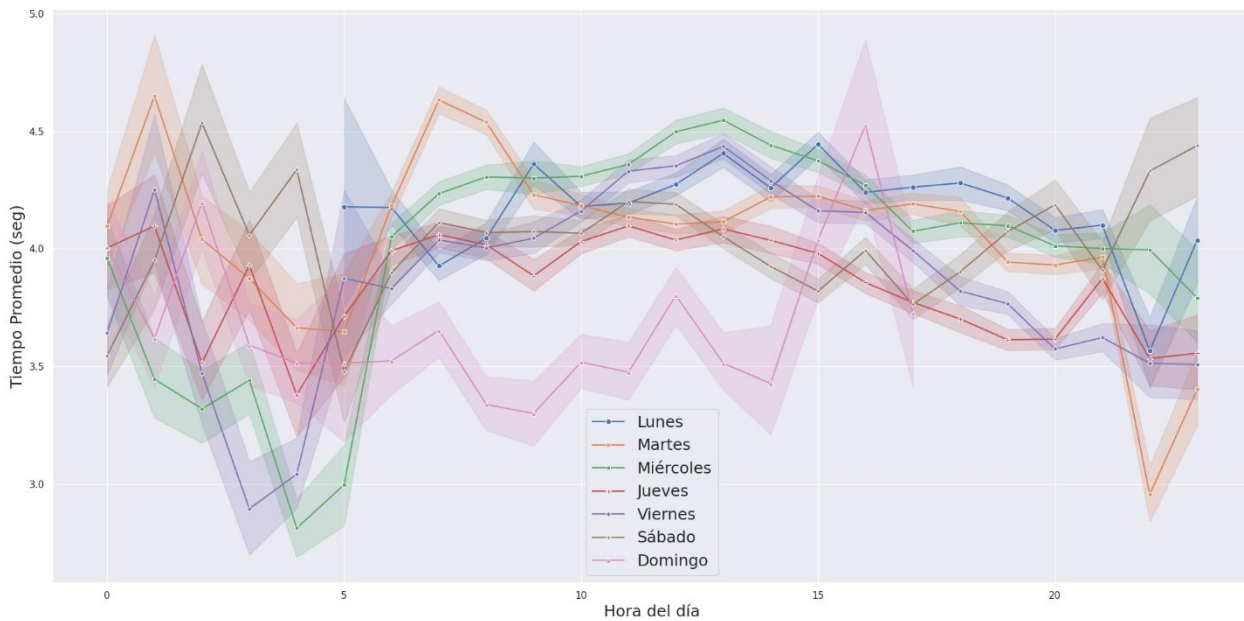
Cantidades de SKU clasificadas por hora del día



En la Figura 16, el Tiempo de Clasificación por SKU entre las 5 a.m. y las 8 p.m. durante los días laborales muestra un comportamiento estable y consistente a lo largo de estas horas. Si se trazara una línea promedio del tiempo de clasificación por SKU durante este intervalo, se observaría una tendencia estable como en la Figura 14, con pocos picos y bandas relativamente estrechas, lo que indica una mayor eficiencia operativa durante las horas centrales del turno diurno.

Figura 16

Comportamiento del tiempo de clasificación por SKU según la hora del día y día de la semana



Durante los días de semana, de lunes a sábado, el tiempo de clasificación por SKU se mantiene en un valor promedio cercano a los 4 segundos entre las 5 a.m. y las 8 p.m. Aunque estos tiempos son más altos en comparación con otras franjas horarias, su variabilidad es menor, lo que refleja una operación más estable y predecible. El tiempo de clasificación tiende a aumentar hasta el mediodía y luego desciende lentamente hasta las 8 p.m., lo que sugiere una mejora en la eficiencia hacia el final del día.

El turno de la noche presenta mayor variabilidad en los tiempos, lo que podría estar relacionado con el procesamiento de un mayor volumen de SKU, como se observa en la Figura 15. Sin embargo, a pesar de esta variabilidad, los tiempos son más bajos en promedio que durante el día, lo cual es consistente con los patrones observados en la Figura 14, que muestra franjas más amplias durante la noche y madrugada.

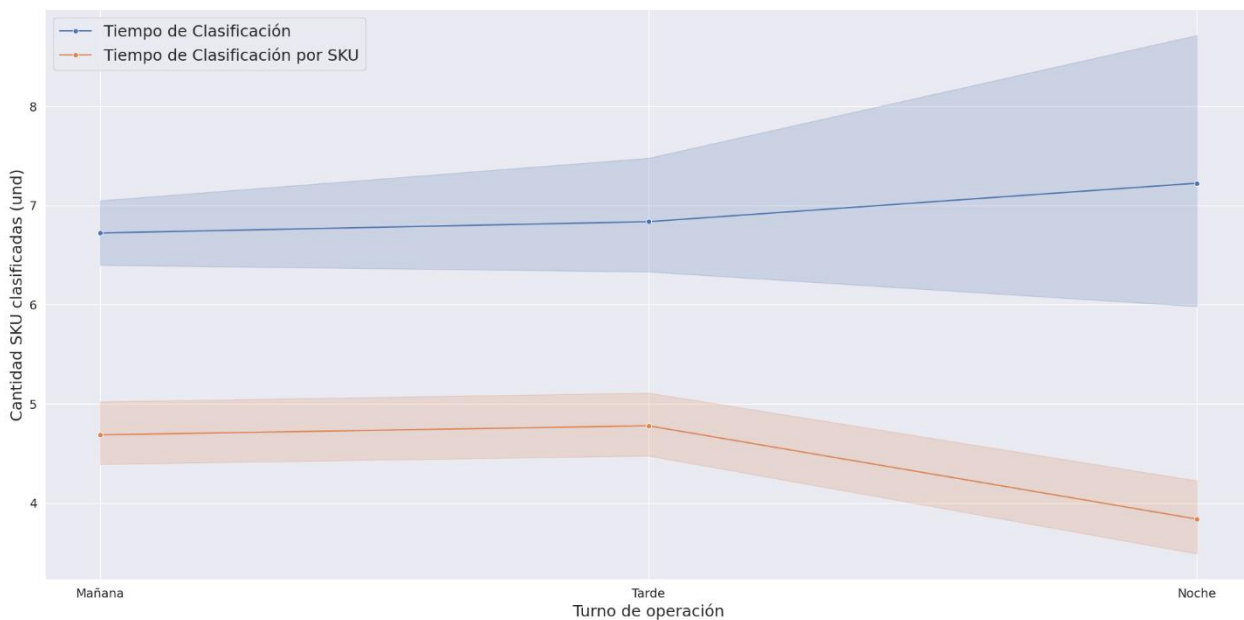
El día domingo, aunque con tiempos medios más bajos, muestra una mayor inestabilidad, con fluctuaciones al inicio y final del día. La jornada dominical finaliza a las 5 p.m., lo que podría explicar esta mayor variabilidad en comparación con los días de semana, cuando la operación se extiende hasta la noche.

Tiempos de Clasificación según el Turno de Trabajo

En la Figura 17, se observa cómo varían los Tiempos Totales de Clasificación (línea azul) y el Tiempo de Clasificación por SKU (línea naranja) según el turno de trabajo. En los turnos de mañana y tarde, se mantiene una tendencia relativamente estable, con menores tiempos por SKU y un leve incremento en los tiempos totales.

Figura 17

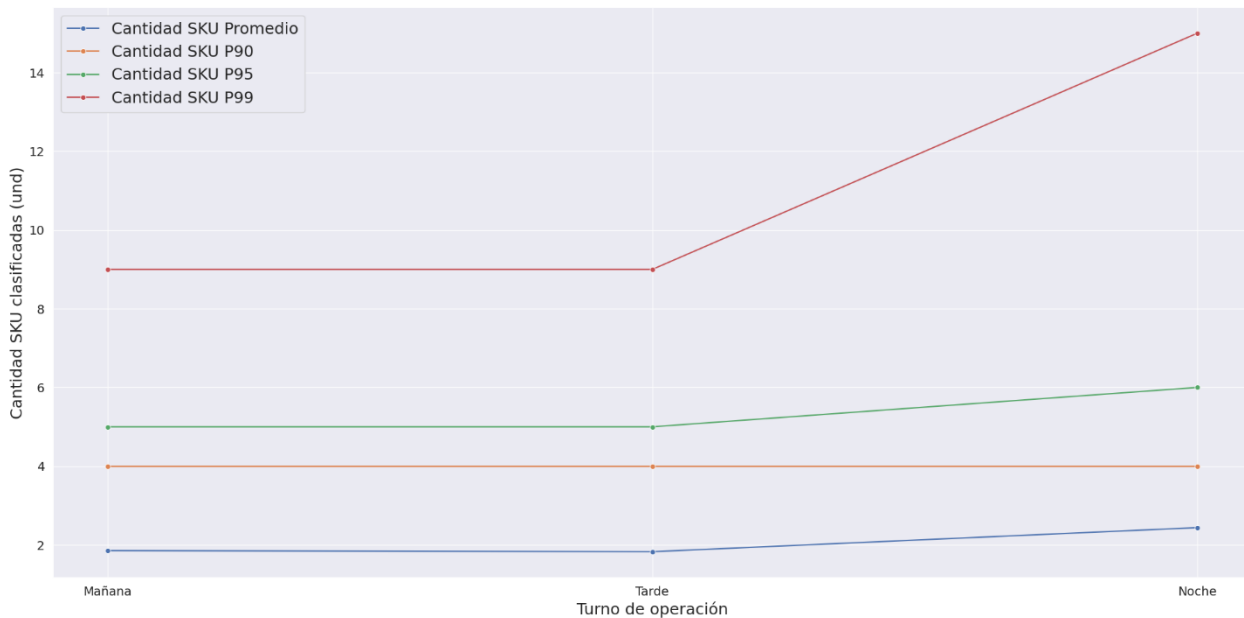
Tiempos de Clasificación según el Turno de Trabajo



En el turno de noche, el Tiempo Total de Clasificación tiende a aumentar, reflejando una mayor demanda de tiempo para completar las tareas de clasificación. Sin embargo, al mismo tiempo, el Tiempo de Clasificación por SKU muestra una disminución en este turno, lo que indica que, aunque se dedique más tiempo en total, la eficiencia operativa es mayor, ya que el tiempo promedio necesario para clasificar cada SKU es menor. Esto sugiere que, durante la noche, se procesa una mayor cantidad de SKU por unidad de tiempo, lo cual refleja una optimización de las operaciones durante este turno. El ancho de las bandas de variabilidad es mayor durante el turno nocturno para los tiempos totales, lo que indica que, aunque la operación es eficiente por SKU, hay mayor variación en los tiempos totales debido a la variabilidad en la carga de trabajo.

La Figura 18 muestra las Cantidades de SKU Clasificadas en cada turno, destacando la media y los percentiles (P90, P95, P99). Durante el turno de mañana y tarde, las cantidades clasificadas se mantienen relativamente bajas y estables, con poca variación entre los diferentes percentiles, lo cual se refleja en los tiempos de clasificación durante los turnos anteriormente mencionados, donde no se observa gran variación.

Figura 18
Cantidades de SKU Clasificadas por Turno de Trabajo



El turno de noche es el más exigente, con un notable incremento en las cantidades de SKU procesadas, especialmente en el percentil P99, lo que indica una mayor concentración de SKU procesados durante esta franja horaria. Este aumento de volumen coincide con el comportamiento observado en la Figura 15, donde las cantidades de SKU procesadas eran mayores entre 9 p.m. y 5 a.m., lo que se alinea con el turno de noche. Este incremento en la cantidad de SKU procesados es consistente con el aumento en los Tiempos Totales de Clasificación observados en la Figura 17, aunque el Tiempo de Clasificación por SKU permanece relativamente bajo, indicando una operación altamente eficiente en términos de procesamiento por unidad.

Durante el análisis del Tiempo de Clasificación por SKU, se identificaron varios factores que influyen en la eficiencia operativa de los equipos PTL. En los meses de noviembre y diciembre, caracterizados por una alta demanda, se observó que, a pesar del incremento en el volumen de trabajo, los equipos PTL y operarios mantuvieron una alta eficiencia. El tiempo por SKU se redujo a un rango de 4 a 5 segundos, gracias a una mejor planificación y la incorporación de más personal para manejar los picos de demanda.

A lo largo de la semana, los jueves y viernes destacaron como los días más productivos, con tiempos de clasificación por SKU entre 4.5 a 5 segundos, sin embargo, no se observa una gran diferencia en los tiempos de clasificación por SKU entre lunes y sábado. Por otro lado, durante la madrugada (9 p.m. a 5 a.m.), se alcanzó una eficiencia notable con tiempos de 4 a 4.5 segundos por SKU, ya que se procesaron grandes volúmenes de unidades. En el día, de 5 a.m. a 8 p.m., el tiempo por SKU se mantuvo en valores constantes cercanos 4.5 segundos.

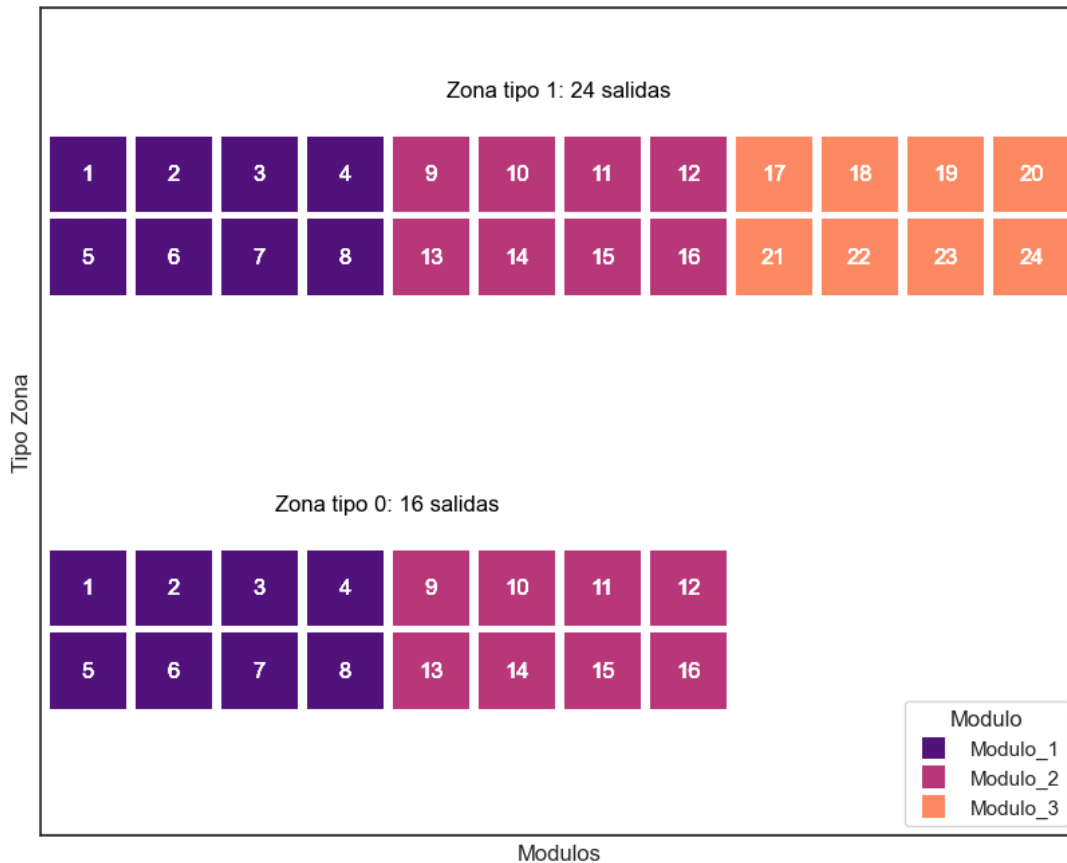
El turno de noche resultó ser el más eficiente, con tiempos promedio de 4 segundos por SKU, mientras que los turnos de mañana y tarde presentaron tiempos algo mayores, pero aún dentro del rango de 4.5 a 5 segundos. Aunque la fatiga laboral podría influir en el aumento de los tiempos totales hacia el final de cada turno, el tiempo por SKU se mantuvo estable, lo que sugiere que, a pesar de la fatiga, los equipos PTL continuaron operando de manera eficiente, aunque con mayor variabilidad dado el ancho de las bandas.

Eficiencia según el tipo de equipo PTL

La Figura 19 muestra dos tipos de equipos PTL con diferentes configuraciones de salidas y módulos de estantería. El equipo PTL tipo 1 tiene 24 salidas, organizadas en tres módulos de estantería. Por otro lado, el equipo PTL tipo 0 posee 16 salidas distribuidas en dos módulos de estantería.

El Equipo PTL tipo 1 (24 salidas), está compuesto por tres módulos de estanterías, con las salidas distribuidas secuencialmente dentro de cada módulo. El Módulo 1 contiene las salidas del 1 al 8, el Módulo 2 abarca las salidas del 9 al 16, y el Módulo 3 incluye las salidas del 17 al 24. Mientras que en el equipo PTL tipo 0 (16 salidas), las salidas están distribuidas en dos módulos de estanterías. El Módulo 1 contiene las salidas del 1 al 8, mientras que el Módulo 2 abarca las salidas del 9 al 16.

Figura 19
Asignación de salidas por tipo de equipo



La distribución de los módulos y las salidas tiene un impacto directo en la eficiencia operativa de los equipos PTL. Equipos con mayor número de salidas, como los de tipo 1, permiten manejar un volumen más alto de clasificación simultánea, mientras que los equipos de tipo 0, con menos salidas, pueden ser más apropiados para cargas de trabajo menores o espacios con restricciones de capacidad. Además, la segmentación en módulos ayuda a optimizar el acceso a las salidas y a distribuir el trabajo de manera más equilibrada entre los operarios, mejorando la productividad.

La Figura 20 presenta los tiempos de clasificación por SKU en las distintas salidas para los dos tipos de equipos PTL: el tipo 0 (16 salidas) y el tipo 1 (24 salidas). A continuación, se analizan los patrones observados, así como las diferencias en los tiempos de clasificación en la primera salida de cada módulo en ambos tipos de equipos.

Figura 20

Tiempos medio de clasificación por SKU en cada salida según el tipo de equipo PTL

S1 4.02	S2 4.40	S3 4.27	S4 4.26	S9 4.67	S10 4.60	S11 4.47	S12 4.31
S5 4.52	S6 4.02	S7 3.88	S8 3.76	S13 4.18	S14 3.98	S15 3.90	S16 3.85

Zona tipo 0: 16 salidas

S1 4.17	S2 3.94	S3 3.85	S4 3.89	S9 3.87	S10 3.77	S11 3.73	S12 3.80	S17 4.26	S18 4.19	S19 4.25	S20 4.54
S5 4.22	S6 3.70	S7 3.72	S8 3.63	S13 4.01	S14 3.93	S15 3.90	S16 4.04	S21 3.90	S22 3.85	S23 3.94	S24 4.24

Zona tipo 1: 24 salidas

En el equipo tipo 0, Figura 20 (superior), los tiempos de clasificación son más variables, con las primeras salidas del Módulo 2 (especialmente la salida S9 con 4.67 segundos) mostrando tiempos más altos. Las salidas finales del módulo son más eficientes, con tiempos cercanos a 3.76 segundos. Este equipo presenta mayores ineficiencias en las transiciones entre módulos. El equipo tipo 1, Figura 20 (inferior), es más consistente, con tiempos más bajos y estables a lo largo de las salidas. Las salidas centrales de los módulos, como la S8 y S9 (3.63 y 3.87 segundos), son las más rápidas, mientras que las primeras salidas de cada módulo, aunque más lentas, no muestran la misma variabilidad que en el equipo tipo 0.

En ambos equipos, los extremos iniciales muestran tiempos relativamente consistentes con una leve mejoría hacia el final en el equipo tipo 0, mientras que en el equipo tipo 1, hay una leve disminución en la eficiencia en las últimas salidas. El recorrido hacia los extremos podría ser un factor que influye en los tiempos de clasificación en ambos equipos. El equipo tipo 0, con solo dos estanterías, muestra una mejor eficiencia en sus extremos, mientras que en el equipo tipo 1, el mayor número de estanterías hace que el acceso a las salidas finales sea más largo, lo que afecta la eficiencia en los extremos. Esta diferencia en el diseño físico podría explicar por qué los tiempos de clasificación tienden a aumentar en las últimas salidas del equipo tipo 1.

Es importante resaltar que existe una diferencia significativa en el rango de los tiempos de clasificación entre las salidas de ambos equipos. En el equipo tipo 0, la diferencia entre el valor más alto (4.67 segundos) y el más bajo (3.76 segundos) es de 0.91 segundos, lo que representa aproximadamente un 25% del tiempo promedio de clasificación, que ronda los 4 segundos. Esta misma diferencia también se observa en el equipo tipo 1, donde el rango entre el valor más alto (4.54 segundos) y el más bajo (3.63 segundos) es igualmente de 0.91 segundos. Estas diferencias

en los tiempos de clasificación entre las salidas confirman que, a pesar de la eficiencia general, la ubicación de las salidas tiene un impacto considerable en el tiempo requerido para clasificar un SKU.

El análisis de los tiempos medios de clasificación por SKU revela que el equipo tipo 0 tiene un tiempo promedio de 4.19 segundos, mientras que el equipo tipo 1 es más eficiente, con un promedio de 3.93 segundos. Esta diferencia, aunque parezca pequeña en términos absolutos, representa una reducción de aproximadamente el 6.2% en los tiempos de clasificación para el equipo tipo 1 en comparación con el equipo tipo 0. Este porcentaje puede parecer modesto, pero dado el volumen de SKUs procesados diariamente, cualquier mejora en la eficiencia, incluso de pocos segundos, puede traducirse en un ahorro significativo de tiempo y recursos a largo plazo. Este hallazgo destaca la importancia del diseño físico y la distribución de las salidas en los equipos PTL, ya que la configuración del equipo tipo 1 parece optimizar el flujo de trabajo y reducir los tiempos de clasificación de manera más efectiva.

La Figura 21 presenta el número total de confirmaciones realizadas en cada salida de los equipos PTL, lo que refleja la frecuencia de uso de cada salida durante el periodo analizado. Este análisis nos ayuda a identificar cuáles son las salidas más frecuentadas y las posibles implicaciones para la operación de los equipos PTL en términos de distribución de la carga de trabajo.

Figura 21
Confirmaciones totales por salida en cada tipo de equipo PTL

S1	S2	S3	S4	S9	S10	S11	S12				
601,229	278,833	182,298	133,155	210,559	132,991	92,000	63,578				
S5	S6	S7	S8	S13	S14	S15	S16				
204,291	128,815	93,762	67,007	86,062	65,395	49,576	37,524				
Zona tipo 0: 16 salidas											
S1	S2	S3	S4	S9	S10	S11	S12	S17	S18	S19	S20
401,636	255,220	204,075	174,773	103,526	93,508	82,886	70,308	123,192	107,629	90,450	68,639
S5	S6	S7	S8	S13	S14	S15	S16	S21	S22	S23	S24
175,759	148,933	129,021	115,843	248,331	193,080	162,241	141,202	85,170	74,310	60,302	44,305
Zona tipo 1: 24 salidas											

Las confirmaciones totales por salida en los equipos PTL tipo 0, Figura 21 (superior) y tipo, Figura 21 (inferior), muestra que las primeras salidas y las salidas centrales son las más frecuentadas. En el equipo tipo 0, la Salida S1 tuvo 601,229 confirmaciones, siendo la más utilizada, mientras que las salidas finales, como S16, recibieron muchas menos confirmaciones

(37,524). Un patrón similar se observa en el equipo tipo 1, donde S1 tuvo 401,636 confirmaciones, mientras que la Salida S24 tuvo solo 44,305. Este comportamiento se debe, en gran parte, a cómo está configurado el equipo PTL. Las primeras salidas están más cerca del lector de código de barras, lo que facilita su uso, mientras que las salidas finales requieren un mayor recorrido, especialmente en el equipo tipo 1, que tiene más estanterías. Esta configuración provoca que las últimas salidas sean menos visitadas, lo que podría generar una distribución desigual de la carga de trabajo y afectar la eficiencia operativa.

En cuanto a las confirmaciones totales, los equipos muestran una diferencia considerable. El equipo tipo 0 alcanzó un total de 2,612,076 confirmaciones, mientras que el equipo tipo 1 (procesó un total de 3,196,340 confirmaciones durante el periodo analizado. Este mayor volumen en el equipo tipo 1 es consistente con su configuración de más salidas y su capacidad para manejar una mayor cantidad de SKU, reflejando su eficiencia superior no solo en términos de tiempos de clasificación más bajos, sino también en términos de manejo de volumen de trabajo. Además, es relevante destacar que, aunque el equipo tipo 1 maneja un mayor volumen de confirmaciones, logra mantener una eficiencia de clasificación más alta, con tiempos ligeramente menores en comparación con el equipo tipo 0. Esto refleja la optimización del layout y la capacidad de procesamiento del equipo tipo 1, lo que permite manejar más SKU de manera más eficiente.

Finalmente, la Figura 22 presenta la cantidad de veces que se visitó cada salida después de la lectura del SKU, lo que nos da una idea de las salidas priorizadas o más frecuentadas en cada equipo PTL.

Figura 22

Cantidad de confirmaciones de cada salida posteriores a la lectura del SKU

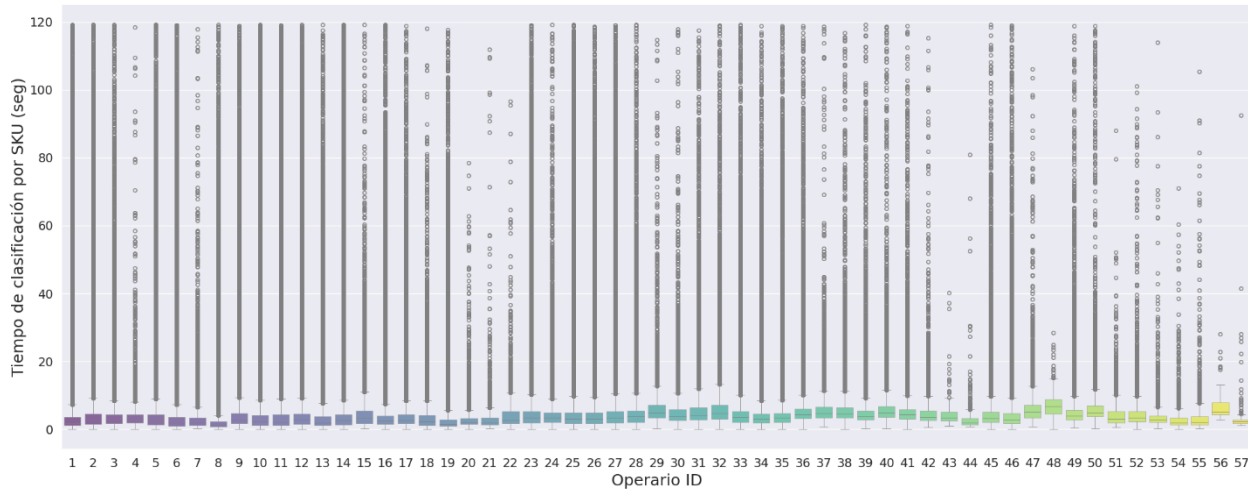
S1	S2	S3	S4	S9	S10	S11	S12				
357,780	66,337	32,577	20,658	55,445	22,422	12,195	6,776				
S5	S6	S7	S8	S13	S14	S15	S16				
55,896	21,407	12,613	8,063	11,857	6,824	4,240	3,021				
Zona tipo 0: 16 salidas											
S1	S2	S3	S4	S9	S10	S11	S12	S17	S18	S19	S20
290,372	71,071	41,076	28,725	12,004	10,004	8,245	6,976	15,694	12,605	10,118	7,755
S5	S6	S7	S8	S13	S14	S15	S16	S21	S22	S23	S24
55,275	27,851	18,882	15,142	64,514	35,711	25,010	19,856	8,912	7,141	5,812	4,413
Zona tipo 1: 24 salidas											

Las confirmaciones posteriores a la lectura del SKU muestra que en ambos equipos PTL, las salidas iniciales, especialmente la Salida S1, son las más visitadas. En el equipo tipo 0, Figura 22 (superior), la S1 registra 357,780 visitas, mientras que en el equipo tipo 1, Figura 22 (inferior) tiene 290,372 visitas, siendo claramente las posiciones más priorizadas. Las salidas centrales, como la S13 en el equipo tipo 1, también tienen una alta frecuencia de uso.

Por otro lado, las salidas finales, como S16 (3,021) en el tipo 0 y S24 (4,413) en el tipo 1, son las menos frecuentadas, lo que refleja una preferencia operativa por las salidas más cercanas. Este comportamiento está relacionado con la configuración del equipo y la distancia que los operarios deben recorrer para acceder a las salidas más lejanas. En resumen, las primeras y centrales salidas son las más utilizadas debido a su cercanía con el lector de código de barras, mientras que las últimas salidas son evitadas. Este patrón de asignación de los pedidos a las salidas del equipo muestra cómo, en el proceso de asignación, se evita sobrecargar las salidas más lejanas y de mayores tiempos por SKU para favorecer la eficiencia del proceso de preparación de pedidos.

Tiempos de clasificación por operario

La Figura 23 refleja una variabilidad notable en los tiempos de clasificación por SKU, aunque, en la mayoría de los casos, los bigotes de las cajas no superan los 15 a 20 segundos. De hecho, muchos de los bigotes inferiores comienzan cercanos a 0 segundos, lo que indica que algunos registros presentan tiempos de clasificación muy bajos, especialmente en los cuartiles inferiores. Sin embargo, La presencia de *outliers* es considerable, con puntos fuera de los bigotes que alcanzan tiempos significativamente más altos, incluso cercanos a los 120 segundos en algunos casos. Estos *outliers* sugieren que en determinadas ocasiones se presentan tareas que toman mucho más tiempo del habitual, lo que podría estar relacionado con eventos operativos atípicos o dificultades específicas en las tareas de clasificación.

Figura 23*Tiempos de clasificación por SKU de cada operario*

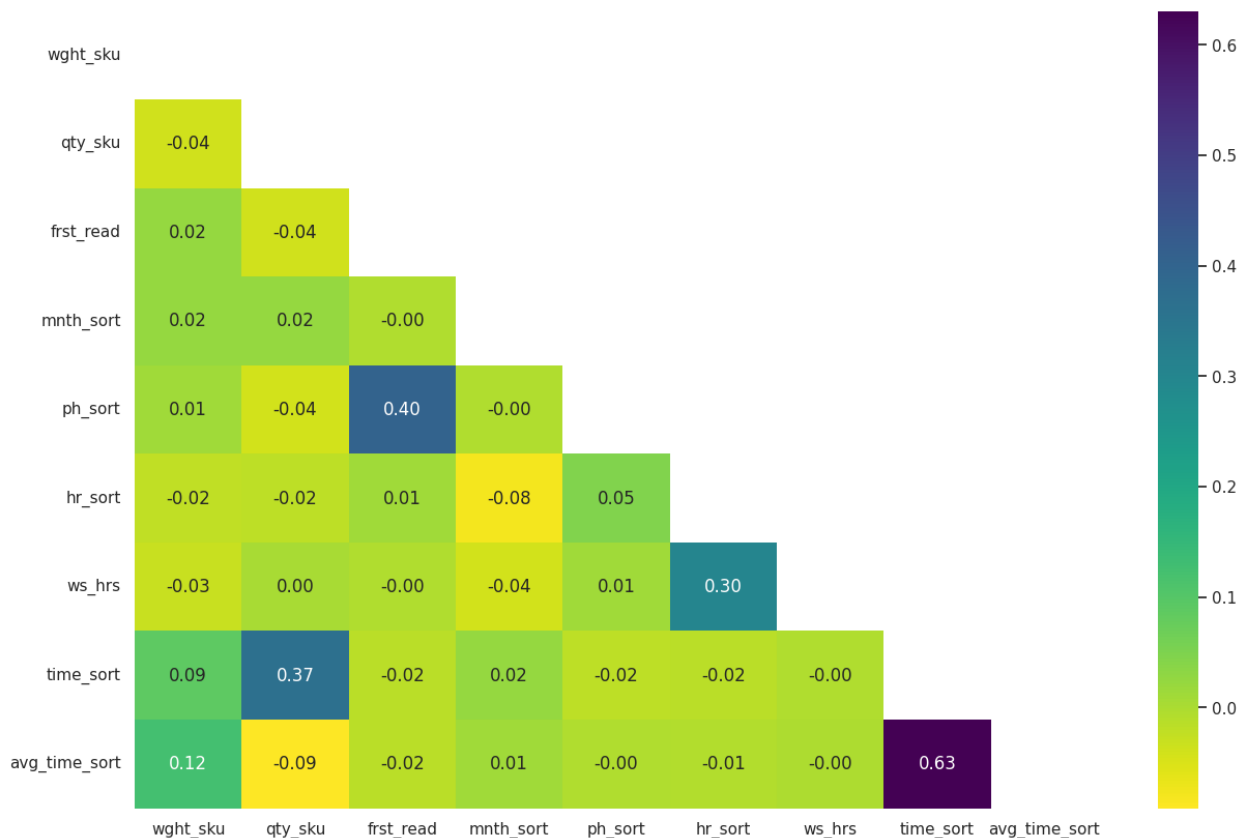
En cuanto a la cantidad de registros, algunos casos presentan poca información disponible, lo que dificulta hacer una evaluación precisa de la consistencia y el rendimiento en esas situaciones. Esto plantea un desafío al tratar de planificar o predecir tiempos de clasificación de manera eficiente.

La combinación de alta variabilidad y la presencia de *outliers*, junto con la dificultad de obtener una estimación precisa en algunos casos por la falta de datos, sugiere que los tiempos de clasificación pueden ser difíciles de prever de forma consistente. Esto indica la necesidad de optimizar los procesos para lograr una mayor homogeneidad y mejorar la eficiencia operativa. La alta variabilidad en los tiempos de clasificación por SKU y la significativa presencia de *outliers* complican el análisis de la productividad de los trabajadores. Aunque los tiempos generales de los bigotes de las cajas no superan los 15 a 20 segundos, algunos registros muestran *outliers* que superan los 100 segundos, lo que distorsiona la visión real de la eficiencia operativa. Este comportamiento en los tiempos dificulta la comparación precisa entre operarios y la evaluación de su desempeño, ya que los valores atípicos y la dispersión en los tiempos ocultan patrones consistentes. Además, la falta de suficientes registros para algunos operarios introduce incertidumbre en la evaluación de su rendimiento, complicando aún más la planificación de tareas y la mejora de los procesos.

Correlación entre variables

En la Figura 24, la relación entre `time_sort` (tiempo total de clasificación) y `qty_sku` (cantidad de SKU clasificados) muestra una correlación positiva ligera de 0.37, indicando que a medida que aumenta el número de SKUs, también lo hace el tiempo de clasificación, aunque de manera no lineal ni consistente. Esta correlación no alcanza un nivel moderado debido a la variabilidad operativa, como se refleja también en la Figura 25, donde los tiempos de clasificación presentan alta dispersión a medida que se incrementa la cantidad de SKU.

Figura 24
Matriz de correlaciones de variables

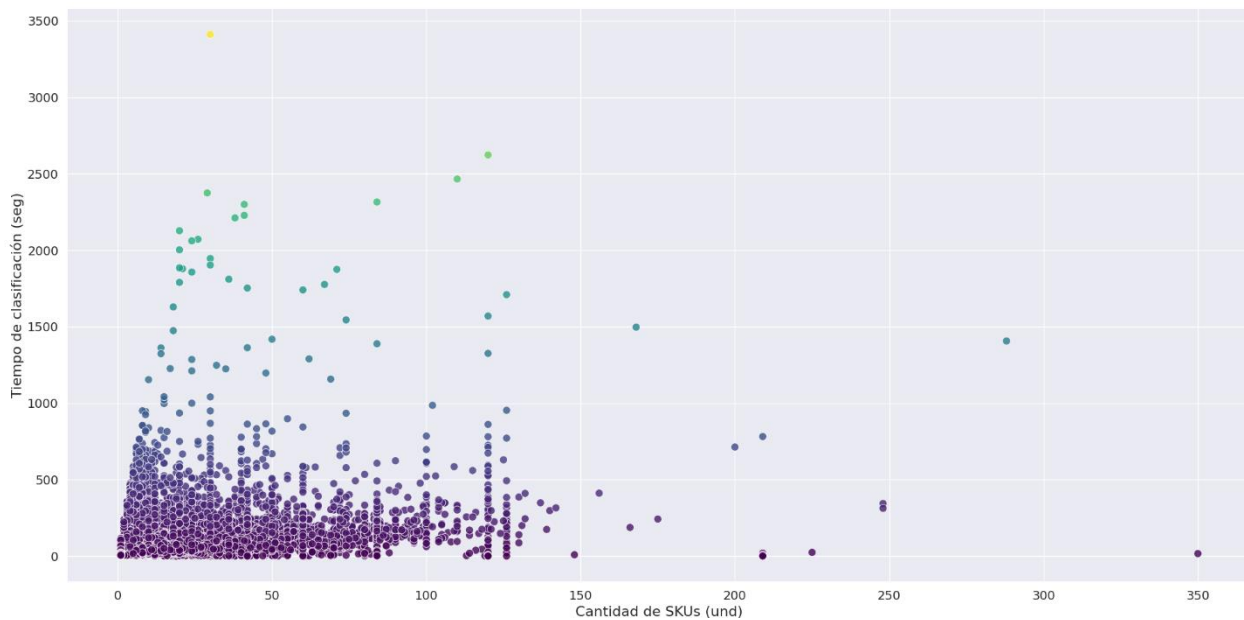


Por otro lado, la correlación débil entre `time_sort` y `wght_sku` (0.09) sugiere que el peso de los SKUs no impacta considerablemente en el tiempo total de clasificación, lo que indica que la cantidad de SKUs clasificados es un factor más relevante para determinar el tiempo necesario. Es importante señalar que la correlación entre `avg_time_sort` (tiempo de clasificación por SKU) y

time_sort es de 0.63. Sin embargo, esta correlación no debe considerarse para análisis independientes, ya que el avg_time_sort se calcula directamente a partir de time_sort dividiéndolo entre la cantidad de SKUs clasificados, lo que explica su moderada-alta correlación.

Figura 25

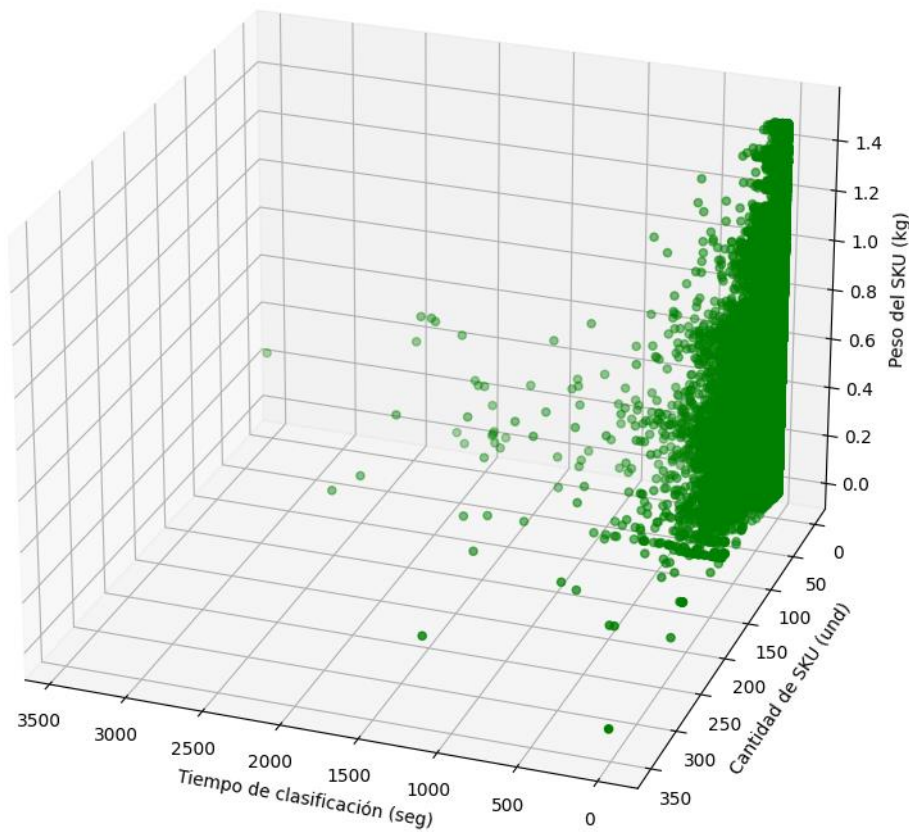
Diagrama de dispersión entre las cantidades de SKU clasificadas y el tiempo de clasificación



La Figura 26 complementa el análisis previo de la Figura 24 y la Figura 25, reafirmando que el tiempo total de clasificación (time_sort) tiene una relación más directa con la cantidad de SKUs clasificados que con el peso de estos. El peso de los SKUs tiene un impacto mínimo en el tiempo total, como lo refleja la correlación débil observada en la Figura 24 (0.09). Aunque el número de SKUs aumenta el tiempo de clasificación, la Figura 26 también muestra la variabilidad en los tiempos, sugiriendo que otros factores operativos influyen en el proceso. Esta complejidad operativa se refleja en la correlación ligera (0.37) entre time_sort y qty_sku, lo que indica que no todas las clasificaciones de grandes volúmenes de SKUs necesariamente requieren mucho más tiempo.

Figura 26

Diagrama de dispersión 3D entre el tiempo total de clasificación, la cantidad de SKU clasificados y el peso del SKU



El análisis confirma que no existe una correlación moderada-fuerte entre el tiempo total de clasificación y las variables estudiadas, como la cantidad de SKU, su peso o las otras variables listadas. Esto sugiere que el proceso de clasificación está influenciado por una variedad de factores operativos que no están reflejados directamente en estas métricas, lo que hace que predecir los tiempos de clasificación a partir de estas variables sea una tarea compleja y poco confiable.

Pruebas de estadística inferencial y no paramétricas***Pruebas de normalidad del tiempo de clasificación***

Para validar si las diferencias observadas en los Tiempos de Clasificación por SKU entre distintos factores (como mes del año, día de la semana, hora del día y turno de trabajo) son

estadísticamente significativas, se realizaron pruebas inferenciales que permitan comparar las medias o medianas de los tiempos en función de estos factores. Dado que, los tiempos de clasificación no siguen una distribución normal, como se observa en la Figura 27, y se rechaza la hipótesis de que los tiempos siguen una distribución normal ya que los valores p de la Tabla 3 indican que no se cuenta con siguiente evidencia; se sugiere emplear pruebas no paramétricas junto con el análisis de diferencias de medias.

Figura 27
Comportamiento en la distribución de los tiempos de clasificación (1) y tiempos de clasificación por SKU (2)

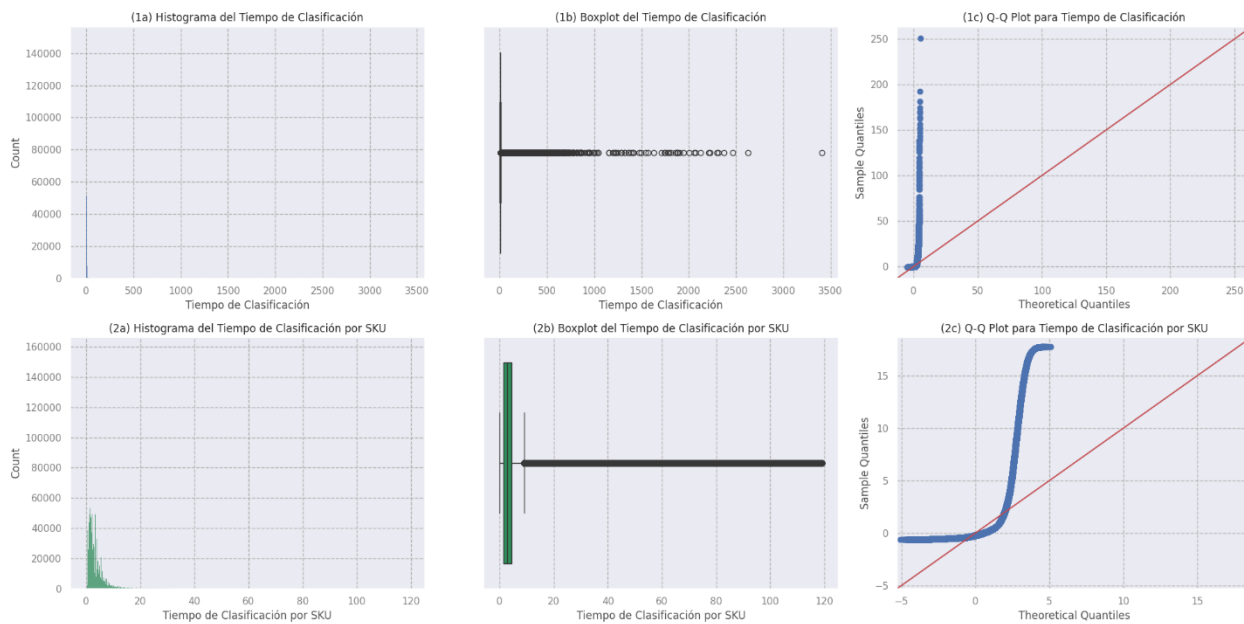


Tabla 3
Planteamiento de hipótesis y pruebas estadísticas para clasificación de tiempos ajustadas a una distribución normal

	Tiempo de clasificación	Tiempo de clasificación por SKU
Planteamiento de hipótesis	H_0 : Los tiempos de clasificación siguen una distribución normal	
	H_a : Los tiempos de clasificación no siguen una distribución normal	
Nivel de significancia α	0.05	0.05

Criterio de rechazo H_0	p-value < α	p-value < α
Prueba Shapiro-Wilk	p-value= 0.0	p-valor= 2.8646e-230
Prueba de D'Agostino's K-squared	p-value= 0.0	p-value= 0.0
Prueba Kolmogorov-Smirnov	p-value= 0.0	p-value= 0.0

Diferencia de medias en el tiempo de clasificación por SKU por tipo de equipo PTL

Se realizó la prueba de diferencia de medias utilizando el test no paramétrico Kruskal-Wallis con el objetivo de constatar las diferencias observadas en la Figura 20, donde se evidenció que el tiempo medio de clasificación por SKU en el equipo tipo 1 era inferior al del equipo tipo 0 en un 6.2%. Los resultados de la prueba arrojaron un estadístico de 35162.23 con un valor-p de 0.0, lo que indica que las diferencias entre los tiempos medios de clasificación entre ambos equipos son estadísticamente significativas. Este hallazgo confirma que la configuración física y el diseño de los equipos, con el equipo tipo 1 presentando un mejor rendimiento en términos de tiempos de clasificación, podrían estar impactando directamente en la eficiencia operativa de los trabajadores.

Diferencia de medias en el tiempo de clasificación por SKU por salidas en equipo tipo 1

Se llevó a cabo la prueba de Kruskal-Wallis para evaluar si existían diferencias significativas en los tiempos medios de clasificación por SKU entre las distintas salidas del equipo tipo 1. La hipótesis nula planteada en este contexto es que no existen diferencias significativas entre los tiempos de clasificación de las diferentes salidas, es decir, los tiempos medios de clasificación de las salidas son iguales. La hipótesis alternativa, por otro lado, establece que al menos una de las salidas tiene un tiempo de clasificación significativamente diferente.

Los resultados de la prueba de Kruskal-Wallis arrojó un valor-p de 0.0, dado que es menor que el valor de significancia de 0.05, se rechaza la hipótesis nula, concluyendo que existen diferencias estadísticamente significativas en los tiempos de clasificación de las distintas salidas del equipo tipo 1.

En la Figura 28, se presenta un grafo que agrupa aquellas salidas que no presentan diferencias estadísticamente significativas en sus tiempos de clasificación. Las aristas del grafo conectan las salidas cuyos tiempos no difieren significativamente según la prueba de Kruskal-Wallis, lo que permite identificar los grupos de salidas que muestran un comportamiento similar. Este hallazgo es crucial para optimizar las operaciones, ya que se pueden enfocar esfuerzos en mejorar las salidas con mayores diferencias en el rendimiento o aprovechar las mejores prácticas de las salidas más eficientes.

En la Tabla 4, se aprecia que el grupo de Alta Productividad tiene un tiempo de clasificación promedio de 3.627 segundos, mientras que el grupo de Baja Productividad tiene un tiempo promedio de 4.535 segundos, lo que evidencia una diferencia notable en la eficiencia operativa entre las distintas salidas del equipo PTL tipo 1. La Tabla 4 también muestra que el grupo con más salidas es el de Media-Alta Productividad, con 14 salidas, lo que indica que una parte significativa de las salidas del equipo tipo 1 presenta un rendimiento cercano al grupo 3, con tiempos de clasificación en torno a los 3.8 segundos.

Figura 28

Agrupación de las salidas del equipo PTL tipo 1 basado en las comparaciones de medias

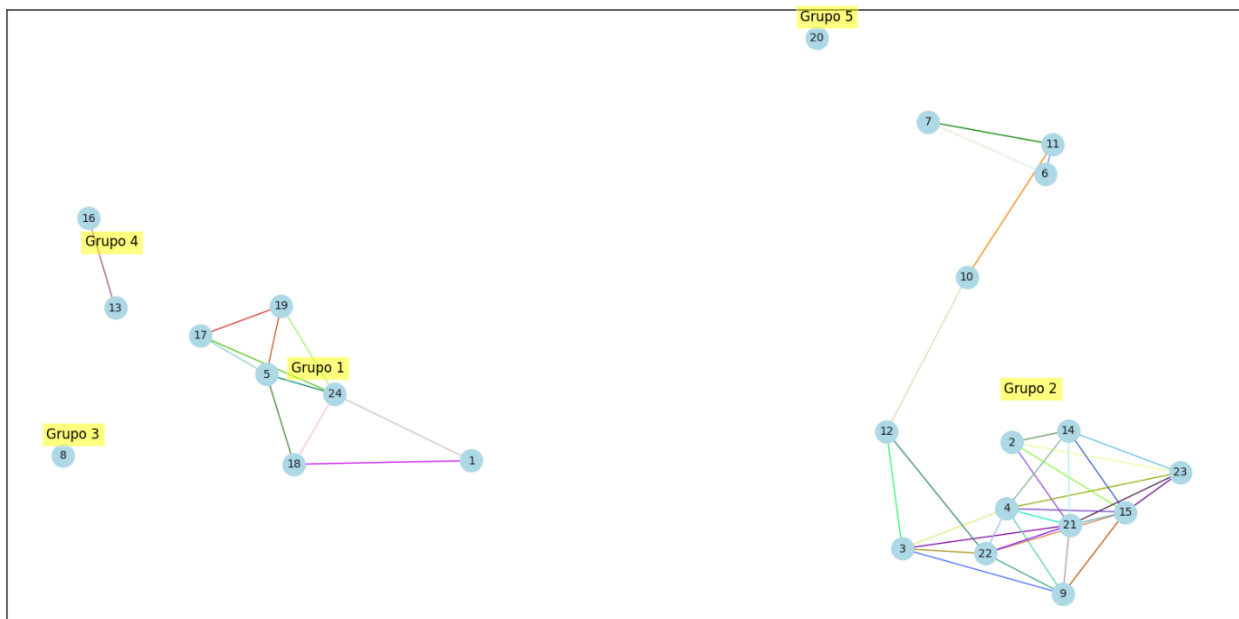


Tabla 4*Métricas de desempeño de los grupos de salidas del equipo PTL tipo 1*

Grupo	Productividad	Tiempo min	Tiempo Prom	Tiempo max	Cant. Salidas
3	Alta Productividad	3.627	3.627	3.627	1
2	Media-Alta Productividad	3.697	3.841	3.939	14
4	Media Productividad	4.008	4.024	4.039	2
1	Media-Baja Productividad	4.174	4.221	4.257	6
5	Baja Productividad	4.535	4.535	4.535	1

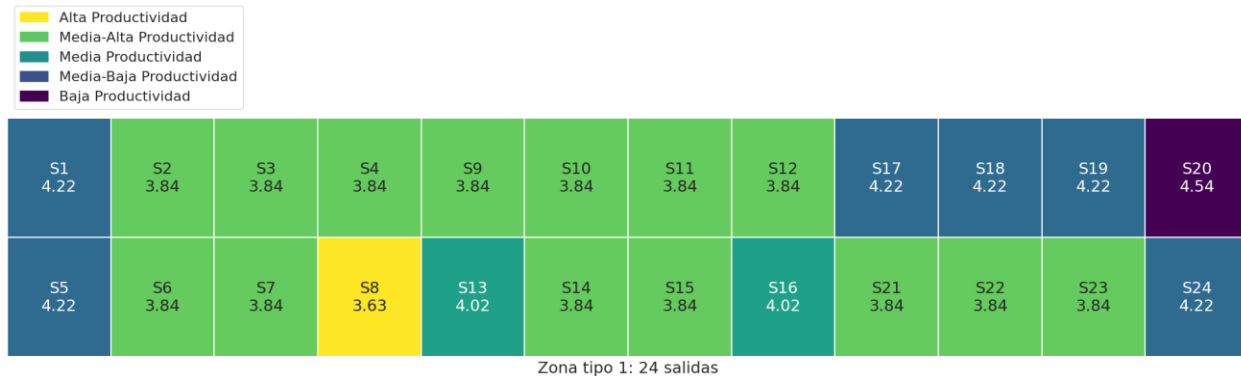
Luego de agrupar las salidas, se obtuvo la Figura 29, en la cual se analizan los tiempos de clasificación promedio y la productividad en las diferentes salidas del equipo por grupos. El lector de códigos de barras por configuración del equipo PTL, se ubica al inicio de las primeras salidas, por lo que se esperaría que las salidas más cercanas al lector, como S1, S2, S3, S5, y S6, mostraran una mayor productividad. La razón detrás de esta expectativa es que la menor distancia al lector debería permitir una velocidad de clasificación más alta.

Sin embargo, al observar los tiempos entre 3.84 y 4.22 segundos en estas salidas, se encuentra que pertenecen a las categorías de Media-Baja Productividad (en el caso de S1, S5, S17, S18, S19, y S24) o Media-Alta Productividad (S2, S3, S6, S7). Esto sugiere que, aunque estas salidas están más cerca del lector, el tiempo promedio de clasificación no es necesariamente mejor.

Puede ser que haya otros factores operativos o de manejo de SKUs que estén afectando la productividad en esta zona, como cuellos de botella en el manejo de ciertos productos o en la asignación de tareas, que ralentizan el proceso a pesar de la proximidad al lector.

Figura 29

Salidas del equipo PTL tipo 1 categorizadas por productividad en el tiempo de clasificación



Se aprecia que el rendimiento de las salidas no sigue un patrón directo de proximidad al lector de códigos de barras. Las salidas intermedias, como S8 (3.63 segundos, Alta Productividad), y algunas otras dentro de la zona central, presentan tiempos de clasificación más bajos, mientras que las salidas cercanas al lector (como S1, S5) y las más alejadas (como S20) tienen tiempos considerablemente más altos.

Esto indica que otros factores operativos, como la naturaleza de los SKUs asignados a ciertas salidas, la carga de trabajo de los operarios o la disposición física del espacio, podrían estar influyendo en la productividad más allá de la simple cercanía al lector.

Diferencia de medias en el tiempo de clasificación por SKU por salidas en equipo tipo 0

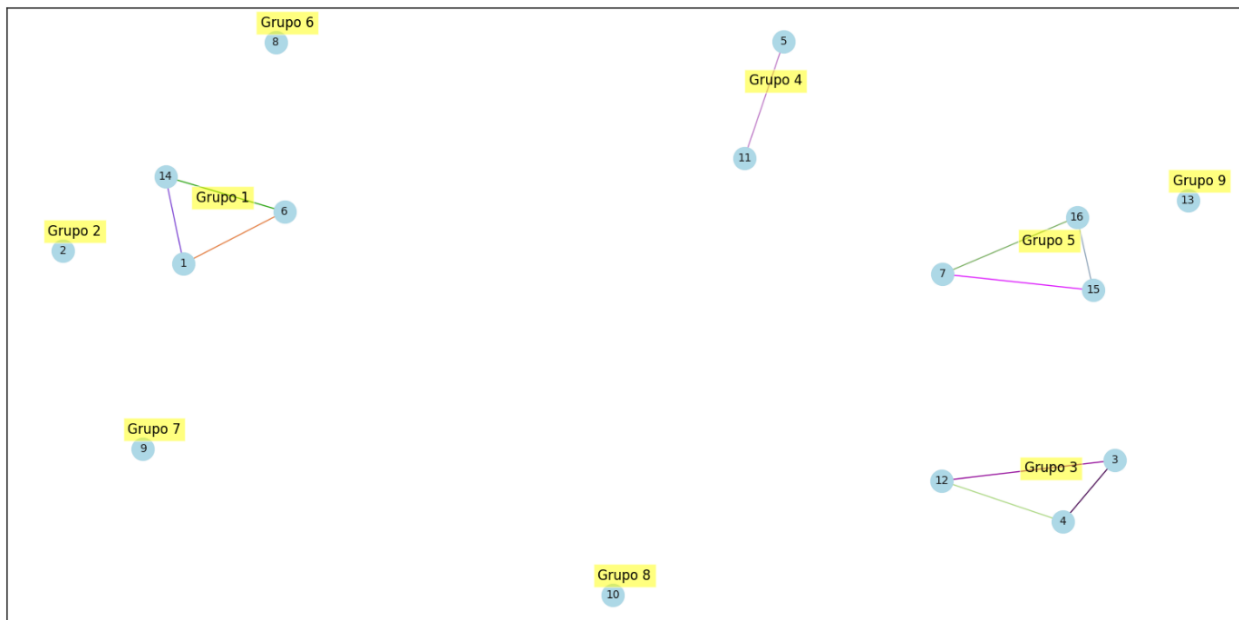
Se llevó a cabo una prueba de Kruskal-Wallis para determinar si existían diferencias significativas en los tiempos medios de clasificación por SKU entre las diversas salidas del equipo tipo 0. La hipótesis nula establecía que no había diferencias significativas entre los tiempos de clasificación de las distintas salidas, sugiriendo que los tiempos promedio de clasificación deberían ser similares. En cambio, la hipótesis alternativa planteaba que al menos una de las salidas presentaba un tiempo de clasificación significativamente diferente.

Los resultados de la prueba mostraron que el valor-p es menor que el nivel de significancia de 0.05, se rechazó la hipótesis nula, lo que significa que existen diferencias estadísticamente significativas entre los tiempos medios de clasificación de las salidas del equipo tipo 0. Este resultado indica que no todas las salidas están operando al mismo nivel de eficiencia, lo cual es un

indicador de que factores externos o específicos de cada salida podrían estar influyendo en su rendimiento. En la Figura 30, se visualiza el grafo que agrupa las salidas del equipo PTL tipo 0 que no presentan diferencias estadísticamente significativas en sus tiempos medios de clasificación. Las aristas conectan las salidas cuyos tiempos no difieren significativamente, permitiendo identificar grupos que presentan un rendimiento similar.

Figura 30

Agrupación de las salidas del equipo PTL tipo 0 basado en las comparaciones de medias



La Tabla 5 muestra los tiempos de clasificación promedio para cada grupo de productividad. El grupo con Alta Productividad tiene un tiempo promedio de clasificación de 3.760 segundos, mientras que el grupo de Muy Baja Productividad presenta un tiempo significativamente mayor, con un promedio de 4.673 segundos. Esta diferencia de tiempos refleja disparidades en el rendimiento de las salidas, que deben ser abordadas para mejorar el flujo de trabajo del equipo PTL tipo 0.

Se identifiquen 9 grupos diferentes, a pesar de que solo existen 16 salidas en el equipo PTL tipo 0, esto sugiere un alto nivel de sensibilidad en la prueba de diferencias estadísticas aplicada (prueba de Kruskal-Wallis). Este resultado puede estar relacionado tanto con el nivel de significancia seleccionado (0.05) como con la variabilidad inherente en los tiempos de clasificación.

Este nivel de sensibilidad puede resultar en una "sobre-fragmentación" de las salidas en más grupos de lo necesario, lo que puede dificultar la toma de decisiones operativas claras. Si el nivel de significancia es demasiado bajo, se corre el riesgo de que el análisis identifique diferencias menores que no necesariamente reflejan problemas operativos significativos. Esto podría llevar a que el equipo de operaciones enfoque esfuerzos de mejora en áreas que no requieren intervención, en lugar de centrarse en aquellas salidas donde las diferencias en la productividad son realmente críticas.

Tabla 5

Métricas de desempeño de los grupos de salidas del equipo PTL tipo 0

Grupo	Productividad	Tiempo min	Tiempo Prom	Tiempo max	Cant. Salidas
6	Alta Productividad	3.760	3.760	3.760	1
5	Media-Alta Productividad	3.851	3.879	3.903	3
1	Media Productividad	3.981	4.005	4.019	3
9	Media-Moderada Productividad	4.177	4.177	4.177	1
3	Moderada Productividad	4.263	4.281	4.310	3
2	Moderada-Regular Productividad	4.403	4.403	4.403	1
4	Regular Productividad	4.468	4.492	4.517	2
8	Baja Productividad	4.602	4.602	4.602	1
7	Muy Baja Productividad	4.673	4.673	4.673	1

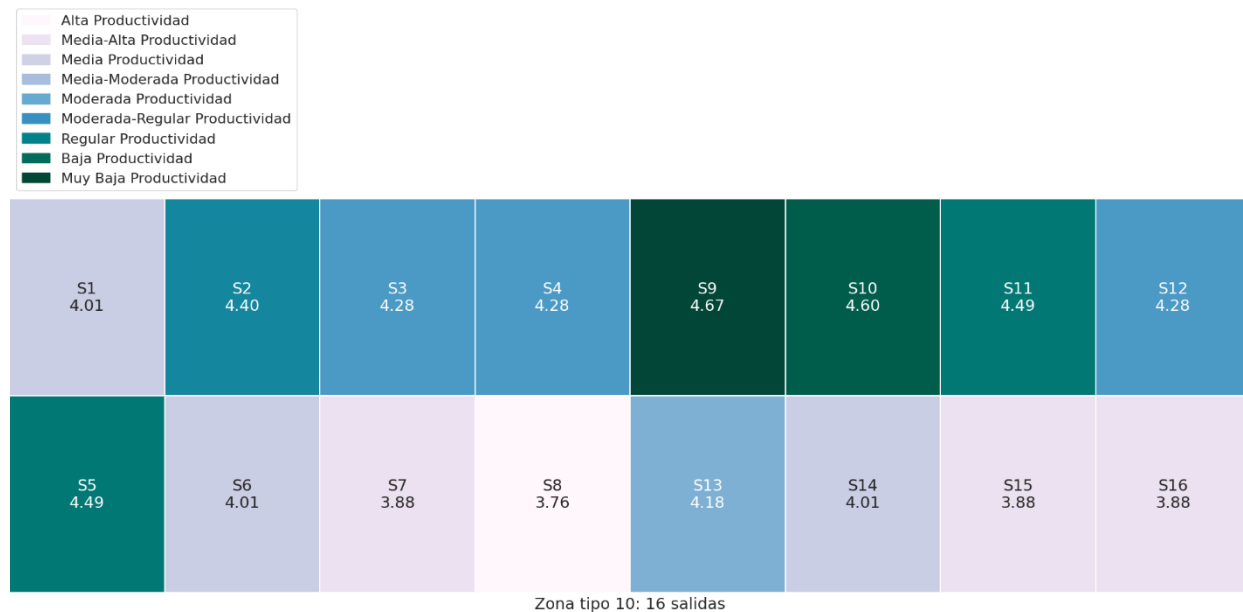
Además, se observa que la diferencia en los tiempos de clasificación entre algunos grupos es mínima. Al comparar los tiempos de clasificación entre el Grupo 5 (Media-Alta Productividad), el Grupo 1 (Media Productividad), y el Grupo 9 (Media-Moderada Productividad), se observa que

la diferencia entre el Grupo 5 y el Grupo 1 es de solo 0.126 segundos, mientras que entre el Grupo 1 y el Grupo 9 la diferencia es de 0.172 segundos. Aunque estos grupos se encuentran separados según la prueba estadística, las diferencias en los tiempos de clasificación son mínimas, lo que indica que el proceso es muy sensible a pequeñas variaciones en el rendimiento de las salidas. Este tipo de fragmentación, basado en variaciones tan pequeñas, evidencia que si bien existen diferencias estadísticamente significativas, estas no siempre se traducen en diferencias sustanciales en la operación diaria. Esto refuerza la idea de que el nivel de significancia o la sensibilidad de la prueba podría estar siendo demasiado estricta para los objetivos operativos.

En la Figura 31, se presenta la agrupación de las salidas del equipo PTL tipo 0 según su productividad, basada en el tiempo medio de clasificación. Sin embargo, se observa que no es fácil identificar un patrón claro o consistente en la ubicación de los grupos de productividad. Por ejemplo, salidas cercanas entre sí como la S2 y la S3 presentan una diferencia significativa en los tiempos de clasificación, con la S2 clasificada como Regular Productividad y la S3 en Media Productividad, a pesar de su proximidad.

Figura 31

Salidas del equipo PTL tipo 0 categorizadas por productividad en el tiempo de clasificación



Además, el comportamiento de las salidas del equipo no parece estar vinculado de manera evidente con su ubicación dentro de la zona. Las salidas que se encuentran más al centro del gráfico, como S9 y S10, tienen tiempos de clasificación más altos y están en categorías de Baja y Muy Baja Productividad, mientras que las más alejadas, como S5 y S6, tienen tiempos relativamente menores, ubicándose en Media y Media-Alta Productividad. Esto sugiere que, aunque la productividad está claramente segmentada, no hay un comportamiento espacial claro que explique las diferencias de rendimiento en función de la ubicación física de las salidas.

La creación de muchos grupos dificulta la identificación de patrones claros en la productividad de las salidas. La alta sensibilidad del análisis fragmenta las salidas en grupos, aunque las diferencias en los tiempos medios sean pequeñas. Esto puede dificultar la detección de tendencias operativas relevantes. Ajustar el nivel de significancia o la sensibilidad podría ayudar a agrupar mejor las salidas y revelar patrones más útiles.

Diferencia de medias en el tiempo de clasificación por SKU por días

Los resultados de la prueba de Kruskal-Wallis arrojaron un valor-p de 0.0. Dado que el valor-p es menor que el nivel de significancia establecido de 0.05, se rechaza la hipótesis nula. Esto

implica que existen diferencias estadísticamente significativas en los tiempos medios de clasificación entre los diferentes días de la semana. Este resultado indica que el día de la semana en el que se realiza la clasificación tiene un impacto significativo en el tiempo promedio de clasificación por SKU. A partir de estos resultados, la prueba de Welch identificó aquellos días que tienen tiempos medios de clasificación similares o significativamente diferentes. En la agrupación mediante grafos, se obtuvo que solo un grupo, está conformado por más de un día: viernes y sábado, lo que indica que estos días tienen tiempos medios de clasificación similares. En cambio, el resto de los días, como lunes, martes, miércoles, jueves, y domingo, se agrupan de manera individual, ya que no se encontró evidencia estadísticamente significativa que sugiera que estos días comparten tiempos medios similares de clasificación.

Diferencia de medias por turnos en el tiempo de clasificación por SKU

La prueba de Kruskal-Wallis arrojó un valor estadístico de 3529.76 y un valor p de 0.0. Dado que el valor p es menor que el nivel de significancia estándar de 0.05, se rechaza la hipótesis nula, concluyendo que existen diferencias estadísticamente significativas en los tiempos medios de clasificación entre los diferentes turnos de trabajo. Esto sugiere que el turno en el que se realiza la clasificación tiene un impacto en el rendimiento, y es posible que algunos turnos sean más eficientes que otros en términos de tiempo de clasificación.

Luego de realizar la prueba de Welch para hacer comparaciones entre los turnos y los tiempos medios de clasificación, se obtuvo que los tres turnos de trabajo, Mañana, Tarde, y Noche, forman grupos individuales. Esto indica que, según los resultados, no hay evidencia estadísticamente significativa que sugiera que alguno de estos turnos comparta tiempos medios de clasificación similares. Cada turno tiene un comportamiento diferenciado respecto a los tiempos de clasificación, lo que sugiere que el momento del día influye de manera importante en la eficiencia operativa. Es notable también que en este análisis no se han formado grupos con más de un turno, lo cual subraya aún más la importancia de la diferencia en los tiempos de clasificación entre cada turno de trabajo.

Diferencia de medias en el tiempo de clasificación por SKU de los operarios

El objetivo del proceso es identificar y cuantificar las diferencias en el desempeño entre los operarios en términos de su eficiencia para clasificar SKU (medido por `time_sort_per_sku`). La combinación de la prueba de Kruskal-Wallis y las comparaciones por pares con Welch's t-test permite: detectar diferencias significativas en los tiempos promedio de clasificación por SKU entre operarios; y determinar exactamente cuáles operarios tienen desempeños estadísticamente diferentes. Este análisis es crucial para evaluar la productividad y la eficiencia de los operarios de manera objetiva, proporcionando información valiosa para la toma de decisiones en cuanto a mejoras operativas, capacitación, o reconocimiento de desempeño.

Cabe resaltar que la métrica `time_sort` (tiempo total de clasificación) está influenciado por la cantidad de SKU en cada operación, lo que dificulta una evaluación directa de la eficiencia operativa. Un operario que clasifica muchos SKU en una tarea tendrá un tiempo total más alto que otro que clasifique menos, sin que esto necesariamente refleje diferencias en su eficiencia o desempeño. Por lo tanto, al trabajar con `time_sort_per_sku` se tiene una métrica normalizada que permite hacer comparaciones más justas entre los operarios, comparando de manera directa el tiempo que cada operario tarda en clasificar una unidad de SKU, independientemente del volumen total de trabajo, lo que facilita un análisis más preciso de la productividad entre operarios.

Se aplicó la prueba de Kruskal-Wallis para determinar si existían diferencias significativas en la eficiencia de clasificación. La hipótesis nula (H_0) establecía que los tiempos promedio de clasificación por SKU de todos los operarios son iguales, es decir, que no había diferencias significativas en su desempeño. Por otro lado, la hipótesis alternativa (H_a) postulaba que al menos un operario tiene un tiempo promedio de clasificación significativamente diferente en comparación con los demás.

Los resultados de la prueba arrojaron un estadístico de 486869.40 y un p-valor de 0.0, lo que implica que se rechaza la hipótesis nula. En términos prácticos, esto significa que existen diferencias estadísticamente significativas entre los operarios, al menos en uno de ellos, en cuanto a sus tiempos promedio de clasificación por SKU. Este resultado confirma que los operarios no están trabajando con la misma eficiencia, y algunos de ellos son considerablemente más rápidos o lentos al clasificar los SKU.

Clasificación de operarios por productividad

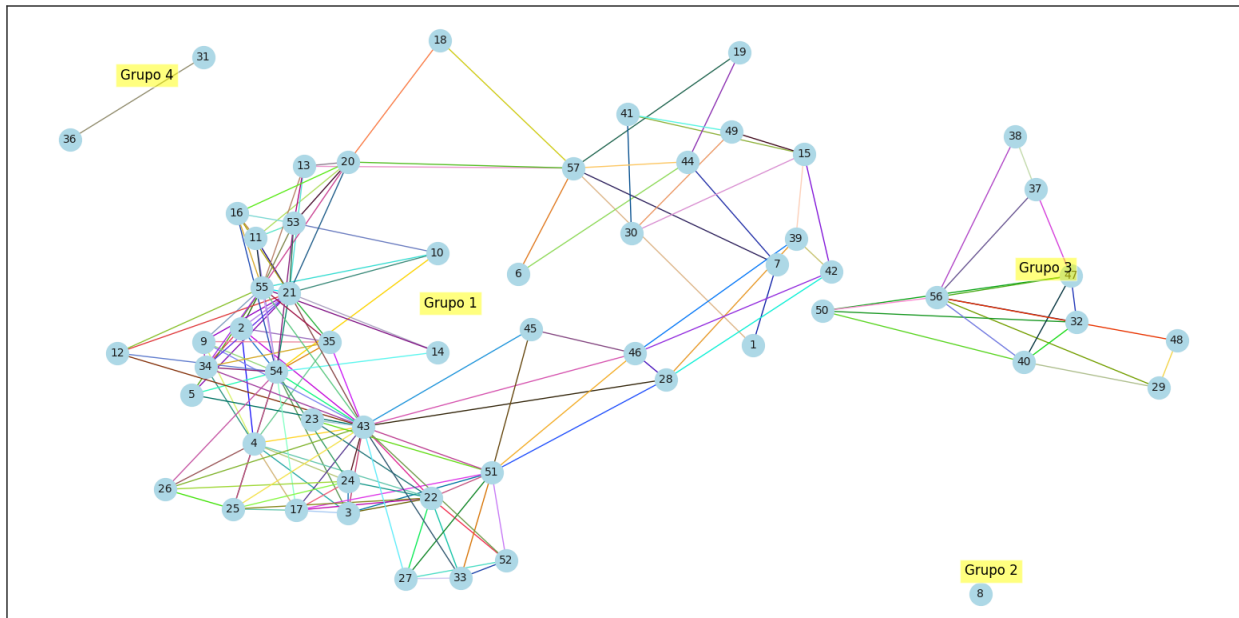
A partir de que a priori, al menos un operario presenta una diferencia significativa en sus tiempos promedios de clasificación, es necesario proceder con comparaciones por pares entre los operarios para identificar específicamente quiénes tienen tiempos de clasificación superiores o inferiores a los demás, utilizando herramientas estadísticas adicionales, como el Welch's t-test y diagramas de red para su agrupamiento. Estas comparaciones permitirán obtener una visión más detallada de las diferencias individuales en el desempeño operativo. La Tabla 32, detalla los resultados de los valores p de la comparación de medias de cada par de operarios.

Tabla 6

Fragmento del resultado de comparaciones pareadas de los operarios por sus medias de tiempo de clasificación por SKU

Trabajador 1	Trabajador 2	Valor-p	Resultado
1	2	0.000000e+00	Menor que 2
1	3	0.000000e+00	Menor que 3
1	4	4.938628e-24	Menor que 4
1	5	0.000000e+00	Menor que 5
1	6	1.630645e-63	Mayor que 6
1	7	8.838086e-01	No significativamente diferente
1	8	0.000000e+00	Mayor que 8
1	9	0.000000e+00	Menor que 9
1	10	0.000000e+00	Menor que 10
1	11	1.552799e-191	Menor que 11

El análisis de los operarios clasificados por su tiempo promedio de clasificación por SKU (time_sort_per_sku) ofrece una visión detallada de cómo se agrupan en función de su eficiencia. En la Figura 32, se observa un gráfico de red que visualiza las conexiones entre los operarios, destacando las relaciones en función de sus comparaciones estadísticas.

Figura 32*Agrupación de los operarios basado en las comparaciones de medias*

En la Figura 32, cada conexión entre pares de operarios (indicados como nodos con un número consecutivo), indica que la prueba de hipótesis Welch's t-test no permite determinar una diferencia entre ellos. Estas conexiones revelan subgrupos de operarios que comparten rendimientos similares, ya que los nodos que están más cerca entre sí representan a operarios con tiempos de clasificación comparables. Por ejemplo, los nodos más alejados o aislados, como los trabajadores identificados por los nodos 31 y 36, sugieren desempeños que se desvían significativamente de los demás, lo que podría indicar tiempos de clasificación inusualmente rápidos o lentos. Este gráfico proporciona una imagen clara de las similitudes y diferencias en el desempeño entre los operarios y destaca aquellos que requieren un análisis más profundo para entender las razones detrás de sus rendimientos atípicos.

Por otra parte, la Tabla 7 revela cómo los operarios se distribuyen en diferentes categorías de productividades según su tiempo promedio de clasificación por SKU. La mayoría de los operarios, 45 en total, se agrupan en la categoría *standard*, lo que indica que su rendimiento es el esperado y no presenta anomalías significativas. Estos trabajadores se consideran dentro de la condición normal de operación en cuanto a eficiencia operativa. También se observa un grupo de 9 trabajadores en la categoría *beginner*, lo que indica que sus tiempos promedio de clasificación son significativamente más largos, lo que sugiere que pueden necesitar apoyo adicional o

entrenamiento para mejorar su eficiencia. Finalmente, 2 trabajadores se encuentran en la categoría *learning*, lo que refleja que están en un proceso de mejora, pero aún no alcanzan el rendimiento esperado.

Tabla 7

Clasificación y cantidad de operarios por categoría de productividad

Grupo	Productividad	Cant. Operarios
2	advanced	1
3	beginner	9
4	learning	2
1	standard	45

Destaca también un único trabajador en la categoría *advance*, el operario 8, un caso particular en el análisis presentado. En el grafo de la Figura 32, el operario 8 se encuentra aislado del resto de los nodos, lo que indica que su rendimiento es significativamente diferente al de otros operarios. Este aislamiento sugiere que sus tiempos de clasificación por SKU pueden ser atípicos, ya sea por ser mucho más rápidos o mucho más lentos que la media de los demás trabajadores. Dado que este operario no muestra muchas conexiones con otros, esto también puede implicar que su rendimiento no es comparable directamente con los de otros trabajadores, lo que lo convierte en un caso importante para revisar a fondo. Puede ser útil observar su historial de rendimiento y analizar las posibles razones de su comportamiento atípico. Este tipo de situaciones podría deberse a factores como una experiencia superior, la naturaleza de las tareas asignadas, o incluso variabilidad en las condiciones operativas.

Análisis de productividades por grupo

La Tabla 8 proporciona un análisis de los tiempos mínimos, medios y máximos de clasificación por SKU en cada categoría de productividad. El operario clasificado como *advanced* presentan un tiempo medio de clasificación significativamente bajo, de aproximadamente 2.7 segundos por SKU, lo que resalta su eficiencia superior. En contraste, los operarios *beginner* muestran un tiempo medio mucho más elevado, alrededor de 6.7 segundos por SKU, lo que indica

que su eficiencia es considerablemente menor. Los operarios en la categoría *learning* tienen un tiempo medio de 5.6 segundos por SKU, lo que sugiere que, si bien están mejorando, todavía no alcanzan los niveles de eficiencia de los operarios en la categoría estándar. Por otro lado, la mayoría de los operarios, clasificados como *standard*, tienen un tiempo promedio de 4.3 segundos por SKU, lo que refleja un desempeño eficiente y esperado, creando un operario tipo a partir del cual se pueden estimar tiempos para la programación de operaciones los equipos PTL.

Tabla 8

Métricas de desempeño de cada una de las categorías de operarios por productividad

Grupo	Tiempo min	Tiempo prom	Tiempo máx	Confirmaciones	Cant. Operarios	Confirmaciones/Cant. Operarios
advanced	2,7037	2,7037	2,7037	178.841	1	178.841
beginner	6,1824	6,6953	7,3884	107.042	9	11.894
learning	5,6435	5,6515	5,6595	111.803	2	55.902
standard	3,0483	4,3291	5,5242	5.383.728	45	119.638

Al observar los tiempos de clasificación por SKU en cada una de las categorías de productividad, se destaca que los extremos de cada categoría están claramente definidos y no se cruzan con los de la categoría siguiente. Esto significa que no hay solapamiento entre las categorías, lo que refuerza la idea de que los límites que separan a los trabajadores de diferentes niveles de habilidad están bien delineados. Esta clara separación demuestra que el método utilizado para segmentar a los trabajadores fue efectivo, ya que permite una clasificación precisa y consistente de acuerdo con su desempeño. Los operarios dentro de cada categoría comparten características comunes en términos de productividad, y las diferencias en el rendimiento entre categorías son objetivas y sustanciales. Esto facilita la identificación de los más productivos y de aquellos que necesitan mejorar, permitiendo tomar decisiones de mejora específicas para cada grupo sin riesgo de confusión o superposición entre los niveles de habilidad.

Análisis de productividades por grupo en cada tipo de equipo PTL

En la zona tipo 1, Tabla 9, el operario *advanced* destaca por su alta eficiencia, con un tiempo de clasificación promedio de 2.6871 segundos, manejando una carga de trabajo considerable. Los *beginner* y *learning* tienen tiempos mucho más altos, entre 5.8642 y 7.4571 segundos, lo que indica que su rendimiento es menor, mientras que los *standard* presentan un desempeño sólido con un promedio de 4.5724 segundos.

Tabla 9

Métricas de desempeño de cada una de las categorías de operarios en el equipo PTL tipo 1

Grupo	Tiempo	Tiempo	Tiempo	Confirmaciones	Cant.	Confirmaciones/Cant.
	min	prom	máx		Operarios	Operarios
advanced	2,6871	2,6871	2,6871	166.893	1	166.893
beginner	5,8642	7,4571	9,5158	31.813	9	3.535
learning	5,7158	6,0358	6,3557	56.293	2	28.147
standard	2,5731	4,5724	6,9350	3.099.340	45	68.874

Por otro lado, en la zona tipo 0, Tabla 10, el operario *advanced* también es altamente eficiente, con un tiempo promedio de 2.9360 segundos, pero con una carga de trabajo menor. Los *beginner* y *learning* en esta zona muestran tiempos más bajos que en la zona tipo 1, con un promedio entre 3.9964 y 6.2988 segundos, lo que sugiere que esta zona, al tener menos salidas, facilita un mejor rendimiento incluso para los operarios con menos experiencia. Los *standard* también son más eficientes, con un promedio de 4.1997 segundos.

Tabla 10

Métricas de desempeño de cada una de las categorías de operarios en el equipo PTL tipo 0

Grupo	Tiempo	Tiempo	Tiempo	Confirmaciones	Cant.	Confirmaciones/Cant.
	min	prom	máx		Operarios	Operarios
advanced	2,9360	2,9360	2,9360	11.948	1	11.948
beginner	3,9964	6,2988	6,8566	75.229	9	8.359
learning	5,2393	5,3849	5,5304	55.510	2	27.755
standard	2,7024	4,1997	5,4974	2.284.388	45	50.764

En ambas zonas, el operario *advanced* mantiene tiempos de clasificación excepcionalmente bajos. Los *beginner* y *learning* de la zona tipo 0 tienen un rendimiento más eficiente que en la zona tipo 1, lo que podría estar relacionado con la menor cantidad de salidas y la simplicidad operativa de la zona tipo 0. Los *standard* también muestran una ligera ventaja en la zona tipo 0 en términos de eficiencia, lo que refuerza la idea de que la configuración del layout y la cantidad de salidas afectan el rendimiento global de los equipos PTL.

5 Conclusiones

El presente estudio analizó la eficiencia operativa de los equipos PTL en los CEDI, identificando los factores que influyen en los tiempos de clasificación y la productividad de los trabajadores. A través de un análisis exhaustivo de los tiempos de clasificación por SKU y la segmentación de los trabajadores por categorías de productividad, se buscó obtener información clave para mejorar la gestión operativa y optimizar los recursos humanos y tecnológicos en los procesos de clasificación.

El análisis global de los equipos PTL, que incluyó la variación de los tiempos de clasificación por SKU según las temporadas del año, los días de la semana, las horas del día y los turnos de trabajo, reveló patrones consistentes de eficiencia, a pesar de las fluctuaciones en el volumen de SKU procesados. Durante las temporadas de alta demanda, como los meses de noviembre y diciembre, el volumen de SKU clasificados aumentó de manera considerable, pero los tiempos de clasificación por SKU se mantuvieron estables en un rango de 4 a 5 segundos, lo que refleja una optimización en la planificación de recursos y la capacidad de los operarios para mantener una alta productividad. Del mismo modo, los turnos de noche, aunque procesan mayores volúmenes de SKU, mostraron tiempos de clasificación por SKU más eficientes, cercanos a 4 segundos, destacándose como el turno de mejor desempeño.

Al analizar los días de la semana, los jueves y viernes presentaron tiempos eficientes similares, mientras que los fines de semana, particularmente los domingos, mostraron una mayor variabilidad operativa. Esto sugiere que, aunque la operación es eficiente en condiciones de alta carga, algunos periodos específicos de menor actividad o personal afectan la estabilidad de los tiempos de clasificación. A pesar de esto, los tiempos de clasificación por SKU se mantuvieron en rangos permisibles, incluso en momentos de alto volumen operativo.

Además, los gráficos de dispersión y la matriz de correlación no mostraron una relación moderada o fuerte entre las variables analizadas (peso del SKU, cantidad de SKU, horas trabajadas, entre otras). La única excepción fue una ligera correlación positiva (0.37) entre el tiempo total de clasificación (time_sort) y la cantidad de SKU clasificados (qty_sku), lo que indica que, aunque el tiempo total de clasificación aumenta ligeramente con la cantidad de SKU, esto no afecta significativamente el tiempo por unidad procesada.

Las pruebas inferenciales y de estadística no paramétrica, como Kruskal-Wallis, confirmaron diferencias estadísticamente significativas en los tiempos promedio de clasificación por SKU entre los operarios, lo que respalda la clasificación en grupos de productividad. La metodología utilizada para realizar estas pruebas, junto con el agrupamiento basado en comparaciones por pares, mediante la validación y segmentación de los grupos usando Welch's t-test y grafos, ofrece una forma efectiva de identificar patrones de eficiencia y establecer planes de acción para mejorar la operación. El método utilizado fue efectivo, segmentando a los operarios en cuatro categorías claramente diferenciadas (*advanced, standard, learning y beginner*), sin solapamientos entre los tiempos de clasificación de cada grupo. Esto permite una identificación precisa de los trabajadores con mayor productividad y de aquellos que requieren mayor apoyo o capacitación, brindando una herramienta valiosa para la planificación de recursos humanos y la optimización de procesos en los CEDI. La segmentación mostró que los operarios en la zona tipo 1 (con 24 salidas) tuvieron un rendimiento promedio ligeramente inferior al de la zona tipo 0 (con 16 salidas), lo que sugiere que un diseño físico más compacto facilita un mejor desempeño de los operarios.

Las pruebas de Kruskal-Wallis para los equipos PTL tipo 1 y tipo 0 confirmaron que existen diferencias estadísticamente significativas en los tiempos medios de clasificación entre las distintas salidas de cada equipo. La diferencia en la productividad entre ambos equipos evidencia la existencia de factores operativos y logísticos que influyen en el rendimiento. En el caso del equipo PTL tipo 1, los resultados revelaron que las salidas cercanas al lector de códigos de barras no necesariamente tienen los mejores tiempos medios de clasificación. A pesar de la proximidad, estas salidas (como S1, S2, S3, S5, S6) pertenecen a categorías de media o media-baja productividad, lo que sugiere que otros factores, como la asignación de productos o la frecuencia de uso, podrían estar influyendo en la eficiencia de estas salidas. En el extremo derecho de la Figura 29, las salidas más alejadas del lector presentan peores tiempos, siendo la salida S20 la de peor rendimiento, clasificada como Baja Productividad.

Las salidas del equipo tipo 0 reveló una gran cantidad de grupos, 9 grupos para 16 salidas. Esta alta fragmentación puede deberse a la sensibilidad del proceso o al nivel de significancia establecido en las pruebas. Aunque los tiempos medios de clasificación entre grupos pueden diferir, las diferencias en el tiempo entre grupos contiguos son mínimas, como en los grupos 5, 1, y 9, con diferencias de menos de 0.2 segundos entre ellos. Esto pone de manifiesto que, aunque se

encuentren diferencias estadísticamente significativas, su relevancia práctica en algunos casos puede ser limitada.

Los resultados de la diferencia de medias por día de la semana mostraron que solo un grupo, conformado por Viernes y Sábado, compartía tiempos medios similares. El resto de los días de la semana no mostró evidencia de tiempos medios significativamente iguales, lo que sugiere que el día de la semana influye de manera importante en la variabilidad de los tiempos de clasificación, probablemente debido a variaciones en la carga de trabajo o la organización operativa.

En cuanto a los turnos de trabajo, no existe evidencia de tiempos medios de clasificación similares entre los diferentes turnos. Cada turno (Mañana, Tarde, Noche) formó su propio grupo o estuvo individualizado, lo que refleja diferencias significativas en la eficiencia de clasificación según el turno.

Los equipos PTL y la organización de los operarios han demostrado una alta eficiencia, incluso en condiciones variables de demanda. La segmentación de los trabajadores por categorías de productividad, la identificación de *outliers* y las diferencias de desempeño según la configuración de las zonas brindan hallazgos clave para la mejora continua de los procesos de clasificación en los centros de distribución. Estos resultados son esenciales para planificar mejor los recursos, mejorar la gestión operativa y asegurar una operación más eficiente y flexible ante cambios en el volumen de SKU y las demandas del mercado.

6 Limitaciones y Trabajo Futuro

En este estudio se buscó en un principio predecir el tiempo de clasificación por SKU en equipos PTL, considerando variables como el operario o su experticia, la familia del SKU, el peso del SKU, el tipo de zona, y la salida en la que se iba a ubicar el SKU. Sin embargo, la alta variabilidad de los datos afectó la precisión de los modelos de predicción, ya que el R^2 no superó 0.20, y los errores como MAE, MAPE y MSE fueron elevados. A pesar del *feature engineering* y la aplicación de varios modelos de regresión y *machine learning* (*Ridge*, *Lasso*, *Random Forest*, *ANN* y *XGBoost*), los resultados fueron insatisfactorios debido a la presencia de *outliers* y problemas operativos, como sesiones no cerradas en los equipos PTL.

Además, los gráficos de dispersión y la matriz de correlaciones no revelaron una correlación moderada o fuerte entre las variables y los tiempos de clasificación. La única excepción fue una correlación ligera (0.37) con la cantidad de SKU (*qty_sku*). Sin embargo, la comprensión del tiempo de clasificación según la cantidad de SKU clasificados llevaba a inconsistencias, ya que clasificar un gran volumen de SKU (decenas) a veces tomaba el mismo tiempo que clasificar unos pocos SKU. Este ruido en los datos dificulta entender el comportamiento real del tiempo y sugiere que los operarios podrían estar incurriendo en otras actividades no capturadas en los registros, como buscar el SKU o realizar pausas, lo que incrementa el tiempo total de clasificación.

Para abordar estos inconvenientes, se optó por una estrategia de clasificación de trabajadores por productividad basada en pruebas no paramétricas como Kruskal-Wallis y Welch's t-test, complementada con grafos para agrupar a los operarios cuyos tiempos no mostraban diferencias significativas. Esta metodología fue más efectiva que los modelos tradicionales de clasificación, como *K-means* o *K-NN*, ya que evitó clasificaciones erróneas debidas a la alta dispersión de los datos.

Una limitación clave fue la falta de información detallada sobre los operarios, como su experiencia, contrato o asignación de turnos, lo que impidió una correlación más precisa entre las características de los trabajadores y su desempeño.

En resumen, aunque no se lograron predicciones satisfactorias con los modelos tradicionales, la clasificación por pruebas no paramétricas resultó ser una solución efectiva para segmentar la productividad de los operarios, lo que ofrece un enfoque más preciso para la planificación y optimización de los procesos en los equipos PTL.

En cuanto a trabajos futuros, se recomienda realizar una limpieza minuciosa de los datos, enfocándose en la eliminación de *outliers* sin comprometer la integridad y la información valiosa del conjunto de datos. Dado que los resultados del presente estudio indicaron que los tiempos de clasificación en su mayoría oscilan entre 2 y 10 segundos, sería adecuado definir límites basados en este rango para descartar los valores extremos que no reflejan el comportamiento real del proceso. Esta estrategia permitiría reducir el ruido y la variabilidad, facilitando la obtención de predicciones más precisas y mejorando la comprensión de los factores que afectan los tiempos de clasificación.

Adicionalmente, sería útil incorporar más datos históricos, cubriendo al menos dos o tres años de operaciones, para tener una visión más robusta del impacto de la demanda en temporadas altas, días festivos y otras variaciones en la carga de trabajo. Asimismo, la recopilación de datos adicionales sobre los operarios—como su experiencia, edad, turnos asignados y formación—sería clave para correlacionar estos factores con la productividad y hacer análisis más precisos.

Otra recomendación está enfocada en mejorar el control operativo, asegurando que los tiempos registrados reflejen el trabajo real, y que los operarios cierren correctamente las sesiones de los equipos PTL al finalizar sus turnos. Esto evitaría errores en los registros y permitiría obtener datos más consistentes para futuras predicciones y análisis de eficiencia. Estas mejoras ayudarán a optimizar la operación de los equipos PTL y a generar estrategias más efectivas para la planificación de recursos y la mejora de procesos.

Finalmente, tener conciencia de las diferencias entre los tiempos promedios (por SKU por operario, por salida o equipo asignado) permitirá a las compañías usuarias de equipos PTL y a sus fabricantes, asignar eficientemente las cargas de trabajo en dichos equipos de manera que se consigan mejores eficiencias operativas en estos procesos.

7 Referencias

- Al-Radaideh, Q. A., & Al Nagi, E. (2012). Using data mining techniques to build a classification model for predicting employees performance. *International Journal of Advanced Computer Science and Applications*, 3(2), 144-151. <http://dx.doi.org/10.14569/IJACSA.2012.030225>
- Ardjmand, E., Sanei, O., & Youssef, E. (2019). Using list-based simulated annealing and genetic algorithm for order batching and picker routing in put wall based picking systems. *Applied Soft Computing Journal*, 75, 106–119. <https://doi.org/10.1016/j.asoc.2018.11.019>
- Atik, C., Kut, A., & Birol, S. (2021). Estimating lead time using machine learning algorithms: A case study by a textile company. *2021 Innovations in Intelligent Systems and Applications Conference (ASYU)*. <https://doi.org/10.1109/asyu52992.2021.9599012>
- AutoStore. (s.f.). Optimizing efficiency: Understanding put-to-light picking systems. Recuperado en julio de 2024, de <https://www.autostoresystem.com/insights/optimizing-efficiency-understanding-put-to-light-picking-systems>
- Boysen, N., Füßler, D., & Stephan, K. (2020). See the light: Optimization of put-to-light order picking systems. *Naval Research Logistics*, 67(1), 3–20. <https://doi.org/10.1002/nav.21883>
- Hamdy, W., Al-Awamry, A., & Mostafa, N. (2022). Warehousing 4.0: A proposed system of using node-red for applying internet of things in warehousing. *Sustainable Futures*, 4, 100069. <https://doi.org/10.1016/j.sftr.2022.100069>
- Janikova, D., & Bezak, P. (2016). Prediction of production line performance using neural networks. *IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, 59-63. <https://doi.org/10.1109/ICAIPR.2016.7585212>
- Krishna, P., Kaur, P., & Singh, S. (2022). Modelling and prediction of worker task performance. *Journal of Industrial Engineering and Management*, 15(3), 103-115. <https://doi.org/10.1016/j.ijpe.2022.108657>
- La República. (2023). El sector del comercio electrónico en Colombia proyecta crecimiento de 17% en 2024. *La República*. Recuperado en julio de 2024, de <https://www.larepublica.co/internet-economy/el-sector-del-comercio-electronico-en-colombia-proyecta-crecimiento-de-17-en-2024-3781454>
- Lightning Pick. (s.f.). Put-to-Light systems. Recuperado en junio de 2024, de <https://lightningpick.com/products/put-to-light/>
- Loske, D., Klumpp, M., Grosse, E. H., Modica, T., & Glock, C. H. (2023). Storage systems' impact on order picking time: An empirical economic analysis of flow-rack storage systems. *International Journal of Production Economics*, 261, 108887. <https://doi.org/10.1016/j.ijpe.2023.108887>
- Muvum. (s.f.). MUV Put-to-Light (PTL). Recuperado en mayo de 2024, de <https://muvumcorp.com/services/muv-put-to-light-ptl/>

- Nayem, Z., & Uddin, M. A. (2024). Unbiased employee performance evaluation using machine learning. *Journal of Open Innovation: Technology, Market, and Complexity*, 10, 100243. <https://doi.org/10.1016/j.joitmc.2024.100243>
- Pittman, P. H., & Atwater, J. B. (Eds.). (2022). *ASCM Supply Chain Dictionary* (17th ed.). APICS, Inc.
- Sabuj, M. S., Biswas, S., & Sultana, S. (2023). Interpretable garment workers' productivity prediction in Bangladesh. *International Journal of Productivity and Performance Management*. <https://doi.org/10.1109/ICCIT57492.2022.10054863>
- Sarker, A., Shamim, S. M., Zama, M. S., & Rahman, M. M. (2018). Employee's performance analysis and prediction using K-means clustering & decision tree algorithm. *Global Journal of Computer Science and Technology*, 18(1), 1-6.
- Shiau, J.-Y., & Lee, M.-C. (2010). A warehouse management system with sequential picking for multi-container deliveries. *Computers & Industrial Engineering*, 58(2), 382-392. <https://doi.org/10.1016/j.cie.2009.04.017>