

# Rhombus

UNIVERSIDAD LATINOAMERICANA DE CIENCIA Y TECNOLOGÍA

VOLUMEN 5, NÚMERO ESPECIAL, OCTUBRE DEL 2025. ISSN: 1659-162355



Rhombus es una revista digital multidisciplinaria de carácter científico, de la Universidad Latinoamericana de Ciencia y Tecnología. Es una publicación semestral, gratuita y de acceso abierto, cuya intención es difundir investigaciones inéditas en el campo de las ciencias empresariales, ingenierías, ciencias sociales y áreas afines. Está dirigida a investigadores, docentes, estudiantes, profesionales y expertos nacionales e internacionales..

## CONSEJO EDITORIAL

**M.Sc. Gina Montero,**  
Universidad Latinoamericana de  
Ciencia y Tecnología, Costa Rica.

**M.Sc. Danilo Montero,**  
Universidad Latinoamericana de  
Ciencia y Tecnología, Costa Rica.

**Dr. Fulvio Lizano,**  
Universidad Latinoamericana de  
Ciencia y Tecnología, Costa Rica.

## CONSEJO CIENTÍFICO

**M.Sc. Antonio Ezequiel Di Genova,**  
Universidad de Belgrano, Argentina

**Dr. Albert Miró,**  
ESERP Digital Business and Law School, España.

**M.Sc. Marco Rojas Lizano,**  
Universidad de Costa Rica, Costa Rica.

**Dra. Silvia Irene Verdugo Guzmán,**  
Universidad Internacional de La Rioja, España.

## EQUIPO EDITORIAL

### EDITOR EN JEFE

M. Sc. Erick Lobo  
elobo@ulacit.ac.cr

### DIAGRAMACIÓN

Carlos Fonseca

### DOCUMENTALISTA

Fabían Mata Monge



Universidad Latinoamericana de Ciencia y Tecnología  
Barrio Tournón, San José, Costa Rica  
Teléfono: (506) 2523-4000  
www.ulacit.ac.cr

Para cualquier notificación o consulta escriba a [rhombus@ulacit.ac.cr](mailto:rhombus@ulacit.ac.cr)

La Universidad Latinoamericana de Ciencia y Tecnología, Costa Rica, no se hace responsable de la opinión vertida por las personas autoras en los distintos artículos.

Hecho el depósito legal.



Licencia Creative Commons Reconocimiento-NoComercial-CompartirIgual 4.0 Internacional (CC BY NC-SA 4)

# RHOMBUS

Revista semestral.

Volumen 5, Número especial, octubre 2025. ISSN: 1659-162355

## CONTENIDOS

<b>PRESENTACIÓN</b>	V
Carolina Guzmán, Decana de la Facultad de Ingeniería	
<b>LA IA Y EL MACHINE LEARNING APLICADA A LA INGENIERÍA INDUSTRIAL A FAVOR DE LA MEJORA EN LA GESTIÓN DE OPERACIONES</b>	1
Rodolfo Fernando Álvarez Calvo	
<b>APLICACIÓN DE MODELOS DE COLAS Y SIMULACIÓN PARA EVALUAR LA EFICIENCIA DEL SERVICIO DE HELADOS EN MCDONALD'S CHAPINERO</b>	21
Laura Sofía Guerrero Guavita, Gabriela Zipaquirá Antolínez & Francly Rocío Castellanos Oviedo	
<b>ANÁLISIS DEL COMPORTAMIENTO DE COMPRA EN TIENDAS DE ALIMENTACIÓN MEDIANTE INTELIGENCIA ARTIFICIAL: UNA MIRADA A LA ATENCIÓN, TIEMPOS Y PATRONES DEL CONSUMIDOR</b>	49
Katherine Yahaira Montañez Chara, Danna Alejandra Rodríguez López & Francly Rocío Castellanos Oviedo	
<b>CACAO FOR PEACE AND REGION PROGRAMME – PROCAPAZ: UNA ESTRATEGIA DE DESARROLLO COMUNITARIO</b>	69
Lady Johanna Peñaloza Farnán	

**INTEGRACIÓN DE REDES Y PROTOCOLOS DE COMUNICACIÓN 91**  
**PARA UNA CELDA DE MANUFACTURA AVANZADA**

Fernando Ricardez Rueda, Hugo Enrique Noriega Pérez, Carlos Adalberto Valenzuela Melchor,  
Beatriz Adriana Esparza Ramírez

**ANÁLISIS Y PREDICCIÓN DE LA SINIESTRALIDAD VIAL EN IBAGUÉ 103**  
**MEDIANTE INTELIGENCIA ARTIFICIAL: UN ENFOQUE DESDE**  
**LA INGENIERÍA INDUSTRIAL**

David Trujillo Durán

## PRESENTACIÓN

### Congreso de Ingeniería Industrial Sin Fronteras 2025

El Segundo Congreso Sur y Centroamericano de Ingeniería Industrial 2025, celebrado en el mes de mayo bajo el lema "Ingeniería Industrial Sin Fronteras, Integrando Tecnologías y Personas", reunió a académicos, investigadores y profesionales de Costa Rica, México y Colombia en un espacio de diálogo e intercambio científico sobre los desafíos más apremiantes de nuestra disciplina. Organizado conjuntamente por ULACIT (Costa Rica), BUAP y Universidad Panamericana (México), y las Universidades de Ibagué y El Bosque (Colombia), este evento consolidó la colaboración regional como estrategia fundamental para enfrentar la transformación digital industrial.

Durante dos jornadas intensas de trabajo, los participantes exploraron seis ejes temáticos centrales: economía circular en el marco de la digitalización industrial, inteligencia artificial en la ingeniería industrial, colaboración humano-máquina, redes inteligentes de innovación abierta, integración de tecnologías 4.0 para servicios especializados, y su aplicación en la producción agrícola. Las presentaciones evidenciaron cómo estas temáticas no solo representan tendencias tecnológicas, sino transformaciones profundas en la manera de concebir y ejercer la ingeniería industrial en nuestros contextos latinoamericanos.

El congreso inició con un conversatorio inaugural que estableció el marco conceptual y contextual del evento, destacando la urgencia de adaptar tecnologías avanzadas a las realidades socioeconómicas de la región. Las conferencias magistrales de Omar Mora Navarro sobre la evolución del análisis de causa raíz mediante *machine learning* y Diseño de Experimentos (DOE en inglés), y del Dr. Lars Maaseidvaag sobre el uso del modelado de procesos para fortalecer el

Análisis de Modo de Fallos y Efectos (en inglés FMEA), proporcionaron visiones integradoras que conectaron metodologías tradicionales con herramientas emergentes de la industria 4.0. La jornada vespertina presentó investigaciones aplicadas en inteligencia artificial. La Universidad Panamericana compartió avances sobre el papel de la IA en la ingeniería industrial.

La Dra. Helga Patricia Bermeo abrió la segunda jornada con una reflexión provocadora sobre cómo el cambio tecnológico está transformando lo conocido, invitando a la audiencia a "prepararse activamente para surfear las olas de innovación en lugar de ser arrastrados por ellas". Su llamado a mantener el rol humano en la toma de decisiones, configuración de valor y liderazgo estratégico resonó a lo largo de todo el evento como un recordatorio esencial: la tecnología potencia, no reemplaza, la capacidad humana de crear soluciones innovadoras. Tal como lo reforzó la conferencia de Juan Pablo Rojas, cofundador de IMAGIA, donde presentó "Ingenieros del Futuro: Potenciando la Industria con Inteligencia Artificial" con ejemplos y casos de aplicaciones en industria.

Esta edición especial de la Revista Rhombus presenta una selección de trabajos que capturan la diversidad y profundidad de las investigaciones presentadas en el congreso, organizados temáticamente para reflejar los principales ejes del evento.

Álvarez presenta un marco conceptual y metodológico sobre la aplicación de la inteligencia artificial y el *machine learning* en la gestión de operaciones industriales, ofreciendo una guía práctica para los profesionales en ingeniería industrial interesados en incorporar estas herramientas a sus procesos de mejora continua.

Dos investigaciones complementarias abordan la aplicación de IA en el sector retail. Guerrero, Zipaquirá y Castellanos analizan la eficiencia del servicio de helados en McDonald's Chapinero mediante modelos de colas y simulación, mientras que Montañez, Rodríguez y Castellanos utilizan inteligencia artificial para estudiar el comportamiento de compra en tiendas D1, revelando patrones en atención, tiempos y conducta del consumidor que permiten optimizar la experiencia de compra.

Ricardez, Noriega y Valenzuela documentan la integración de redes y protocolos de comunicación en celdas de manufactura avanzada, demostrando cómo la interoperabilidad de sistemas es fundamental para la implementación exitosa de soluciones de industria 4.0, particularmente en entornos educativos donde se prepara a la próxima generación de ingenieros industriales.

Trujillo desarrolla un modelo de análisis y predicción de siniestralidad vial en Ibagué mediante inteligencia artificial, ilustrando cómo las herramientas de la industria 4.0 pueden aplicarse a problemáticas de alto impacto social, demostrando el potencial de la ingeniería industrial para contribuir a la seguridad ciudadana y la política pública.

Peñaloza documenta el programa "Cacao For Peace And Region Programme – Procapaz", una estrategia integral que contribuye al fortalecimiento comunitario en zonas rurales de Colombia. Este trabajo evidencia cómo la ingeniería industrial puede articular procesos productivos con innovación social, transferencia de conocimiento y construcción de paz en territorios afectados por el conflicto armado.

Cada uno de estos artículos representa no solo un aporte académico, sino una respuesta contextualizada a los desafíos específicos que enfrentan nuestras sociedades latinoamericanas. Los autores han demostrado que la integración de

tecnologías avanzadas con conocimiento local y enfoque humano no es solo posible, sino necesaria para un desarrollo industrial sostenible e inclusivo.

Esta edición especial es, en definitiva, un testimonio del compromiso de las universidades organizadoras con la generación de conocimiento pertinente y la formación de profesionales capaces de liderar la transformación industrial de la región sin perder de vista la dimensión humana y social de la ingeniería.

Ing. Carolina Guzmán Arroyo, MAP.

Decana de la Facultad de Ingeniería

---

**Nota sobre el proceso de elaboración:** El presente texto fue elaborado con el apoyo de Claude (inteligencia artificial de Anthropic), utilizada para la síntesis y organización de ideas, estructuración del documento y refinamiento del estilo. La interpretación crítica, el análisis de contenido y las conclusiones reflejan el trabajo analítico del autor, complementado por las capacidades de procesamiento y organización textual de la herramienta de inteligencia artificial.

# La IA y el Machine Learning aplicada a la ingeniería industrial a favor de la mejora en la gestión de operaciones

## AI and Machine Learning applied to industrial engineering to improve operations management

---

Rodolfo Fernando Álvarez Calvo<sup>1</sup> 

---

Fecha de recepción: 22-08-2025 | Fecha de aprobación: 25-09-2025

### Resumen

La gestión de operaciones en entornos empresariales enfrenta un proceso de transformación impulsado por tecnologías emergentes como la inteligencia artificial (IA) y el aprendizaje automático (Machine Learning, ML). Estos avances permiten optimizar recursos, automatizar procesos y apoyar la toma de decisiones en tiempo real, aspectos esenciales para incrementar la competitividad en la ingeniería industrial. La evolución de los sistemas de planificación, desde los MRP hasta los actuales ERP y CRM inteligentes, evidencia la transición hacia entornos ciberfísicos más adaptativos y predictivos, donde la integración de big data, IoT y computación en la nube redefinen la trazabilidad y la eficiencia operativa. Esta transición de la Industria 4.0 a la 5.0 refleja un cambio hacia un modelo más humano y sostenible, que combina la colaboración entre máquinas inteligentes y trabajadores con un enfoque en la personalización y la responsabilidad social. A nivel mundial, estudios de firmas como McKinsey, Deloitte y Gartner proyectan un impacto económico de billones de dólares para 2025, destacando aplicaciones en manufactura, farmacéutica, energía y servicios. Sin embargo, persisten retos significativos en la adopción del ML, como la escalabilidad, el versionado de modelos y la alineación con objetivos estratégicos. En consecuencia, la integración estratégica de IA y ML constituye un pilar clave para la innovación, productividad y sostenibilidad de las operaciones industriales.

### Palabras clave

IA, Machine Learning, IoT, gestión de operaciones, ERP, MRP

---

<sup>1</sup> Doctor en Docencia, Magíster en Proyectos, Licenciado en Ingeniería de Calidad, ULACIT, Auditor para Organizaciones Educativas ISO 21001:2018 e ISO 19011:2018, Green Belt Six Sigma, Especialista en Mediación de Entornos Virtuales, actualmente docente universitario de ingeniería industrial, administración de empresas y docente Área Técnica para el MEP en la Especialidad de Procesos Industriales. Correo electrónico: [ralvarez@cfia.or.cr](mailto:ralvarez@cfia.or.cr)

## Abstract

Operations management in business environments is facing a transformation driven by emerging technologies such as artificial intelligence (AI) and machine learning (ML). These advances enable resource optimization, process automation, and support for real-time decision-making—all essential aspects for increasing competitiveness in industrial engineering. The evolution of planning systems, from MRP to today's intelligent ERP and CRM systems, demonstrates the transition toward more adaptive and predictive cyber-physical environments, where the integration of big data, IoT, and cloud computing redefines traceability and operational efficiency. This transition from Industry 4.0 to 5.0 reflects a shift toward a more human and sustainable model, combining collaboration between intelligent machines and workers with a focus on personalization and social responsibility. Globally, studies by firms such as McKinsey, Deloitte, and Gartner project an economic impact of trillions of dollars by 2025, highlighting applications in manufacturing, pharmaceuticals, energy, and services. However, significant challenges remain in the adoption of ML, such as scalability, model versioning, and alignment with strategic objectives. Consequently, the strategic integration of AI and ML constitutes a key pillar for the innovation, productivity, and sustainability of industrial operations.

## Keywords

AI, Machine Learning, IoT, Operations Management, ERP, MRP

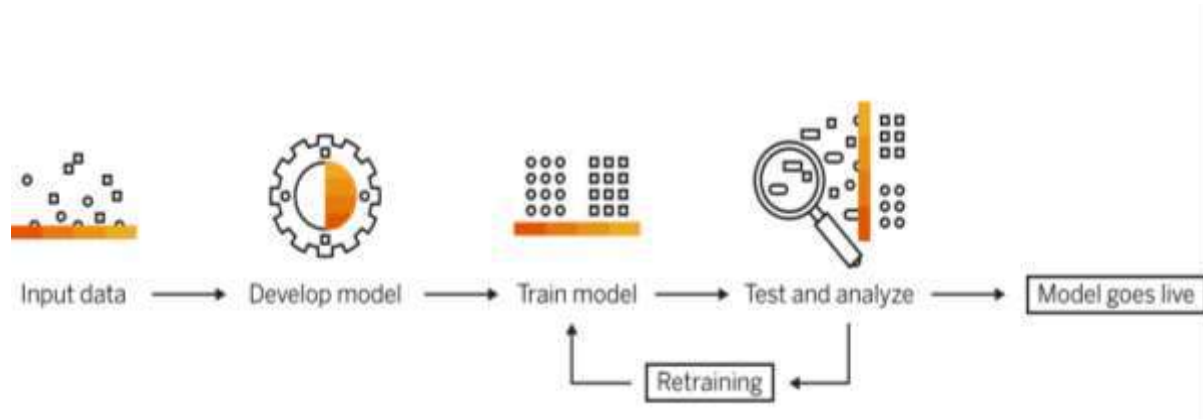
## Introducción

En los contextos empresariales contemporáneos, la competitividad se ha convertido en un factor determinante que influye directamente en la supervivencia y el éxito de las compañías. En este escenario dinámico y globalizado, la gestión de operaciones desempeña un papel crucial, ya que constituye el eje mediante el cual se articulan los recursos, procesos y estrategias para alcanzar altos niveles de eficiencia y productividad. Hoy en día, las empresas no solo deben responder con rapidez a los cambios del mercado, sino también anticiparse a ellos mediante la implementación de metodologías y tecnologías que optimicen la toma de decisiones y garanticen resultados sostenibles.

La planificación de operaciones, en particular, se ha transformado en un área estratégica donde convergen algoritmos avanzados y técnicas de modelación matemática orientadas a la optimización de recursos. Según Cabanillas (2020), el uso de tecnologías emergentes se perfila como una herramienta invaluable para los ingenieros industriales, quienes encuentran en estas soluciones la posibilidad de procesar grandes volúmenes de datos (*big data*) y abordar modelos de gran complejidad que resultan imposibles de gestionar de forma tradicional (Troncoso et al., 2023).

**Figura 1**

*Funcionamiento del Machine Learning*



Fuente: Tomado de SAP. (2025).

En este marco, la revolución tecnológica ha introducido innovaciones disruptivas que están cambiando radicalmente los paradigmas industriales. La inteligencia artificial (IA), que en sus primeras fases se enfocaba en automatizar tareas básicas o replicar procesos cognitivos humanos, está evolucionando hacia la inteligencia artificial generativa (IAG). Esta última se distingue por su capacidad adaptativa y por generar contextos cada vez más complejos y evolucionados, lo cual resulta altamente pertinente en ámbitos industriales donde la automatización, la flexibilidad y la toma de decisiones basadas en datos cuantitativos constituyen ventajas competitivas determinantes. Sin embargo, la falta de una definición unificada del concepto de IA y de criterios claros para la selección de algoritmos adecuados puede conducir a resultados ambiguos o a un alcance limitado de los beneficios esperados (Coto, 2021).

Por lo tanto, el presente artículo tiene como finalidad analizar la incidencia de la IA y del Machine Learning (ML) en los procesos de planificación y programación de operaciones industriales. A través de esta exploración se pretende describir el funcionamiento de estas tecnologías, comprender su estructura y proponer posibles integraciones en áreas específicas de la ingeniería industrial. De esta manera, se busca contribuir a la discusión sobre la optimización operacional y la transición hacia un modelo de gestión más inteligente, ágil y sostenible.

## **Contexto actual de la gestión de operaciones**

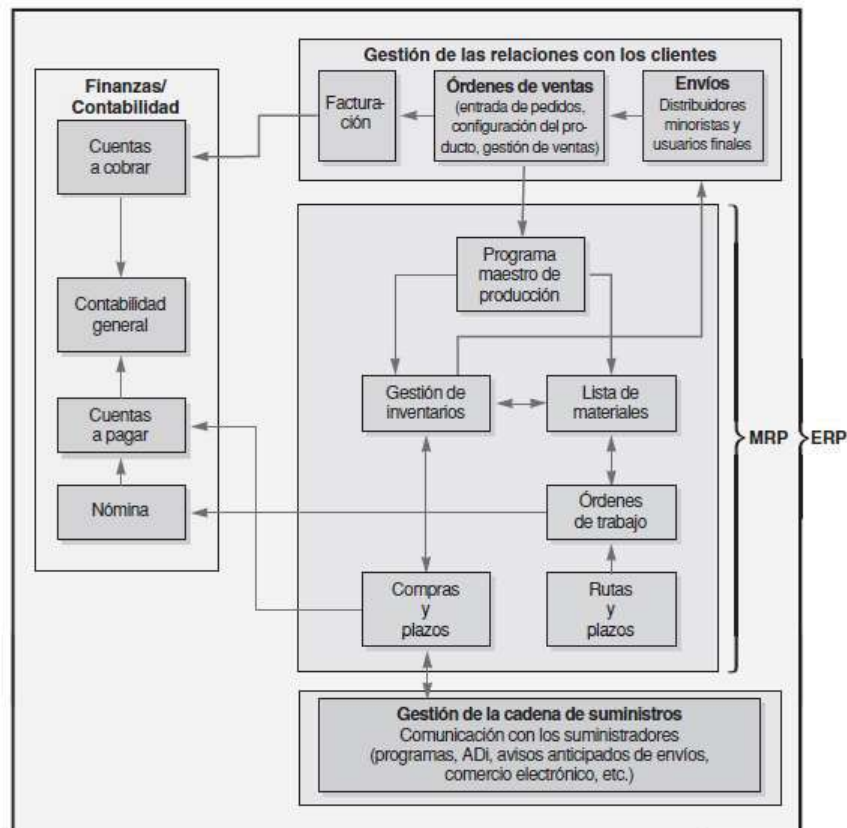
### ***Sistemas ERP***

Los sistemas de planificación de recursos empresariales, más conocidos como ERP (Enterprise Resource Planning), constituyen una de las herramientas más relevantes en la historia reciente de la gestión organizacional. Estos sistemas computacionales existen desde hace varias décadas, y su

génesis respondió a la necesidad de gestionar de manera más eficiente los recursos empresariales que, en contextos crecientes de competitividad, exigían un mayor control y coordinación. Un ejemplo claro de su aplicabilidad se encuentra en la previsión de requerimientos sobre demandas futuras, función que se convirtió en un pilar estratégico para reducir incertidumbres y anticipar necesidades (Heizer & Render, 2007).

**Figura 2**

*Flujo de información del ERP con áreas funcionales de la empresa.*



Fuente: Tomado de Heizer y Render (2007).

Según Chapman (2006)

los ERP han demostrado ser mucho más que un software administrativo, ya que logran extender un control organizacional integral al abarcar prácticamente todos los departamentos de la compañía. De esta forma, áreas como ventas, finanzas, recursos humanos, logística y producción quedan integradas en una misma plataforma, lo cual permite no solo optimizar la planificación de actividades cotidianas, sino también asegurar la coherencia entre los objetivos estratégicos y las operaciones diarias (p. 151).

La implementación de ERP no solo se traduce en una mayor eficiencia, sino también en múltiples beneficios medibles: incremento de la productividad, trazabilidad detallada de la información, reducción de costos operativos, mejora en la capacidad de respuesta frente a clientes y mercados, así como un fortalecimiento del proceso de toma de decisiones. Estas ventajas evidencian que los ERP han dejado de ser herramientas opcionales para convertirse en sistemas imprescindibles en el contexto actual de digitalización empresarial.

### ***Evolución de los Sistemas MRP***

La Planificación de Requerimientos de Materiales (MRP I) surgió en los años 60 como una solución para automatizar el cálculo de necesidades de materiales en función de una demanda independiente y dependiente, integrando listas de materiales (BOM), inventarios y el plan maestro de producción. Su implementación permitió reducir inventarios, mejorar la disponibilidad de materiales y planificar la producción de manera más eficiente (Chapman, 2006).

Posteriormente, el MRP evolucionó hacia el MRP II (Manufacturing Resource Planning), ampliando el enfoque original del MRP I al integrar no solo los materiales, sino también los recursos productivos, incluyendo capacidades de planta, gestión financiera, control de piso y planificación de la demanda. Para Jacobs y Chase (2018) este sistema se convirtió en una herramienta clave para alinear los objetivos estratégicos de la empresa con las operaciones diarias, promoviendo una visión más integral de la planificación operativa. El MRP II introdujo conceptos de retroalimentación cerrada y permitió simulaciones de escenarios futuros, mejorando la toma de decisiones con la implementación de más datos y variables (Heizer & Render, 2007, p. 174).

**Figura 3**  
*Evolución del MRP I y MRP II*



Fuente: Tomado de Lauramajon4 (2018)

En la actualidad, se discute la posibilidad de un MRP III, aún no estandarizado, como una extensión natural que incorpore IA, aprendizaje automático (ML), y análisis predictivo. Esta visión busca un sistema autónomo, capaz de anticipar interrupciones en la cadena de suministro, adaptar planes dinámicamente, y optimizar en tiempo real las decisiones productivas en entornos complejos e inciertos (Ivanov et al., 2018). MRP III se conceptualiza como una evolución ciberfísica del sistema tradicional, en donde los datos de sensores IoT, algoritmos de IA y plataformas cloud colaboran para transformar la planificación de recursos en una actividad cognitiva y proactiva, más que reactiva.

### **Los Sistemas CRM**

Hoy en día la transformación digital y las crecientes expectativas de los consumidores han impulsado a las organizaciones a buscar herramientas tecnológicas que garanticen una gestión de operaciones más eficiente y centrada en el cliente. El Customer Relationship Management (CRM) es una filosofía empresarial respaldada por una plataforma tecnológica que permite gestionar la información de los clientes, automatizar procesos y generar inteligencia de negocios orientada a la personalización y fidelización (Kotler & Keller, 2016).

**Figura 4**

*Usos de un CRM*



Fuente: Tomado de Initium Software (2024)

En este contexto, los sistemas CRM han evolucionado de ser simples plataformas de registro de contactos a convertirse en potentes motores de integración de datos, análisis predictivo y soporte a la toma de decisiones estratégicas en la gestión operativa (Buttle & Maklan, 2019). Esta

transformación ha tenido un impacto significativo en industrias como manufactura, servicios, comercio electrónico y logística, donde el enfoque en el cliente se entrelaza con la productividad.

Los CRM actuales, basados en IA, ML, big data e IoT, representan una evolución hacia un CRM inteligente, donde se predice el comportamiento del cliente, se automatizan tareas operativas y se facilita la toma de decisiones en tiempo real. En este sentido, la gestión de operaciones se vuelve más ágil, adaptable y orientada al cliente (Wamba et al., 2020).

### ***El avance de la industria 4.0 a industria 5.0***

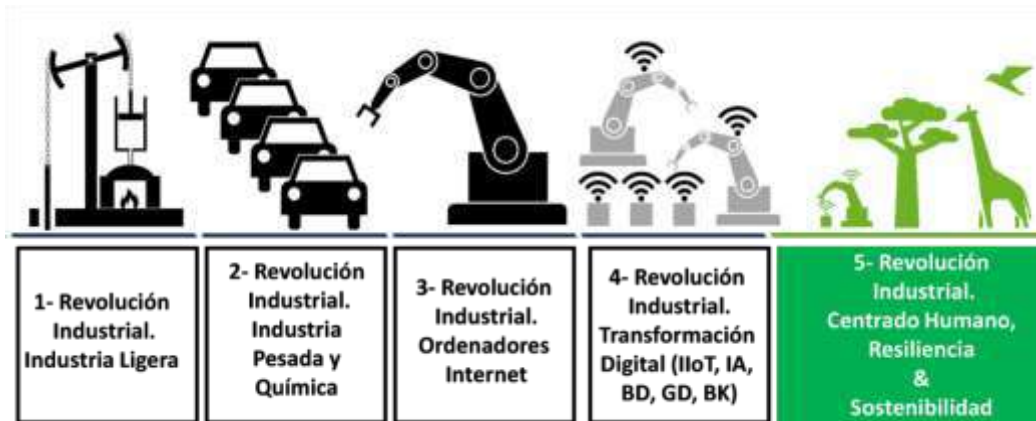
El concepto de industrias inteligentes de hace un tiempo atrás está siendo desplazado por parte de la IA. Las organizaciones están migrando a sistemas cada vez más complejos en sus razonamientos, siendo estos más adaptativos para la toma de decisiones. Los avances en los sistemas computacionales son de todos los días un vivo ejemplo es el paso de los sistemas tradicionales en las empresas las cuales cuentan con sus equipos físicos, pero que ya muchas organizaciones han dado el salto a ERP's en la nube, con estas modificaciones las corporaciones apuestan por nuevas formas de atender los requerimientos existentes (Jacobs y Chase, 2018).

La Cuarta Revolución Industrial, o Industria 4.0, ha estado marcada por la digitalización y la automatización inteligente, caracterizada por la integración de tecnologías como el Internet de las Cosas (IoT), la computación en la nube, la robótica avanzada, la inteligencia artificial y el big data en los procesos industriales. Su propósito principal ha sido alcanzar una producción más flexible, personalizada y eficiente (Schwab, 2016). Esta revolución ha permitido el desarrollo de sistemas ciberfísicos, fábricas inteligentes y cadenas de suministro digitalizadas.

Sin embargo, es relevante afirmar que la Industria 5.0 surge como una respuesta crítica a las limitaciones de la Industria 4.0, especialmente en lo que respecta al papel del ser humano en la industria de manufactura (Verma, 2014). Esta nueva fase incorpora valores éticos, inclusión y resiliencia como pilares fundamentales del desarrollo industrial, buscando un equilibrio entre eficiencia tecnológica y responsabilidad social.

**Figura 5**

*Historia de las revoluciones industriales*



Fuente: Tomado de Amendiola (2022)

En lugar de una automatización total, la Industria 5.0 promueve la colaboración entre humanos y máquinas inteligentes, humanizando las labores y haciendo énfasis de la personalización en masa, la sostenibilidad ambiental y el bienestar del trabajador (Nahavandi, 2019). La integración de las filosofías asociadas con el equilibrio y desarrollo ambiental toman fuerza hoy en día para muchas corporaciones que desde ya vislumbran la importancia de esta fusión: social-económica y tecnológica.

***Relación de la Inteligencia artificial (IA) con el Machine Learning (ML)***

La IA, de acuerdo con Forero & Negre (2024), se puede entender como la capacidad de diseñar máquinas y sistemas pensantes que simulen procesos de razonamiento humano (p. 2). En un sentido amplio, la IA constituye un campo interdisciplinario de la informática que busca desarrollar sistemas capaces de realizar tareas tradicionalmente reservadas a la inteligencia humana, como la toma de decisiones, el reconocimiento de patrones complejos, la comprensión del lenguaje natural, la resolución de problemas y el aprendizaje a partir de la experiencia acumulada.

Dentro de este vasto campo, el ML ocupa un lugar privilegiado al representar la rama de la IA que se especializa en diseñar algoritmos capaces de aprender de manera autónoma a partir de los datos. Según Luan & Tsai (2021), el ML se focaliza en el aprendizaje por medio del análisis masivo de información, lo que le permite a los sistemas mejorar su desempeño de forma continua sin necesidad de ser reprogramados explícitamente para cada tarea. Esta característica resulta especialmente útil en entornos industriales donde la variabilidad, el volumen y la velocidad de los datos hacen indispensable un enfoque flexible y adaptativo.

**Figura 6**

*Tabla comparativa de IA y ML*

Concepto	Inteligencia Artificial (IA)	Machine Learning (ML)
Definición	Ciencia de hacer máquinas inteligentes	Subcampo de la IA que aprende de datos
Alcance	Más amplio	Más específico
Ejemplos	Robótica, visión por computadora, NLP	Clasificación, regresión, clustering
Método de aprendizaje	Puede ser programado o aprender	Aprende automáticamente de datos
Relación	ML es una forma de implementar IA	Parte de la IA

Fuente: Elaboración propia a partir de Forero & Negre (2024).

En síntesis, la relación entre IA y ML puede comprenderse mediante una analogía sencilla pero ilustrativa: la IA puede considerarse como el cerebro en su totalidad, el cual busca replicar la inteligencia humana en sus diferentes dimensiones, mientras que el ML es la capacidad específica de aprender, adaptarse y evolucionar dentro de ese *cerebro*. En el contexto de la gestión de operaciones industriales, esta sinergia abre oportunidades significativas para la innovación, la optimización de procesos y la creación de soluciones más eficientes y sostenibles.

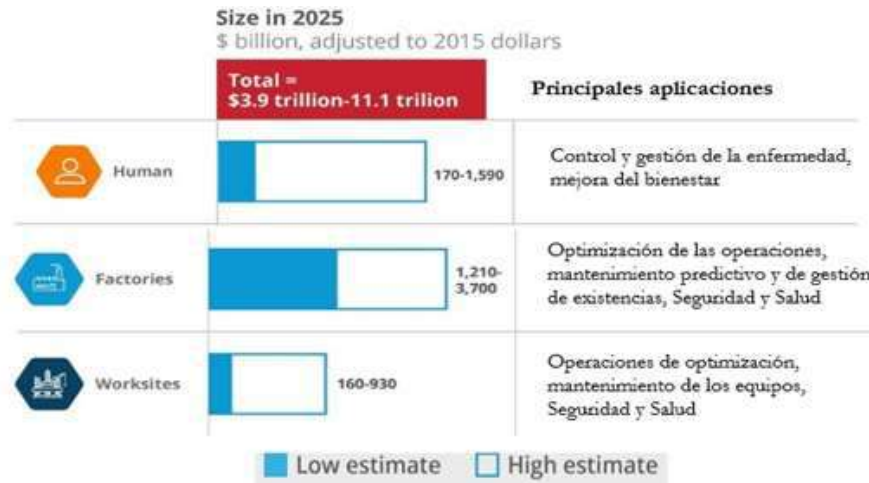
### **Análisis del impacto tecnológico - empresarial a nivel mundial**

#### ***Impacto económico potencial de parte IoT para el 2025***

Según la firma McKinsey (en Williams, 2025) habrá una estimación considerable del impacto del Internet de las Cosas, tanto a nivel de factor humano, como factor industrial, así como en las cargas de trabajo a nivel general. La imagen a continuación presenta la proyección antes mencionada:

**Figura 7**

*Impacto económico del IoT*



Fuente: Tomado de Williams (2025).

Se consideran tres áreas clave de aplicación: humana (Human), fábricas (Factories) y sitios de trabajo (Worksites), con un rango total de impacto que oscila entre \$3.9 y \$11.1 billones. Esta amplia variación en las estimaciones refleja la incertidumbre inherente a la adopción tecnológica, pero también señala un potencial significativo de transformación económica mediante el uso de tecnologías avanzadas como el IoT, inteligencia artificial y sistemas ciberfísicos.

En el ámbito humano, el impacto económico estimado es el más bajo entre las tres categorías, con un rango entre \$170 y \$1,590 mil millones. Sin embargo, su relevancia radica en los beneficios sociales, como el control y gestión de enfermedades y la mejora del bienestar general. Estas aplicaciones están alineadas con el desarrollo de tecnologías médicas, dispositivos portátiles para monitoreo de salud, y sistemas de alerta temprana que pueden mejorar la calidad de vida y reducir costos hospitalarios mediante intervenciones tempranas y personalizadas.

Por otro lado, las fábricas representan el sector con mayor potencial económico, con un impacto estimado entre \$1.21 y \$3.7 billones. Esto se explica por la aplicación de tecnologías para la optimización de operaciones, mantenimiento predictivo y gestión eficiente de inventarios. Además, este sector se beneficia significativamente de la automatización y digitalización, lo que incrementa la productividad y reduce costos operativos. Finalmente, los sitios de trabajo (Worksites), con un impacto de entre \$160 y \$930 mil millones, también muestran un potencial considerable, especialmente en industrias como la construcción y minería, donde la seguridad, la eficiencia del uso de equipos y el mantenimiento son claves. En conjunto, estos datos subrayan la importancia de

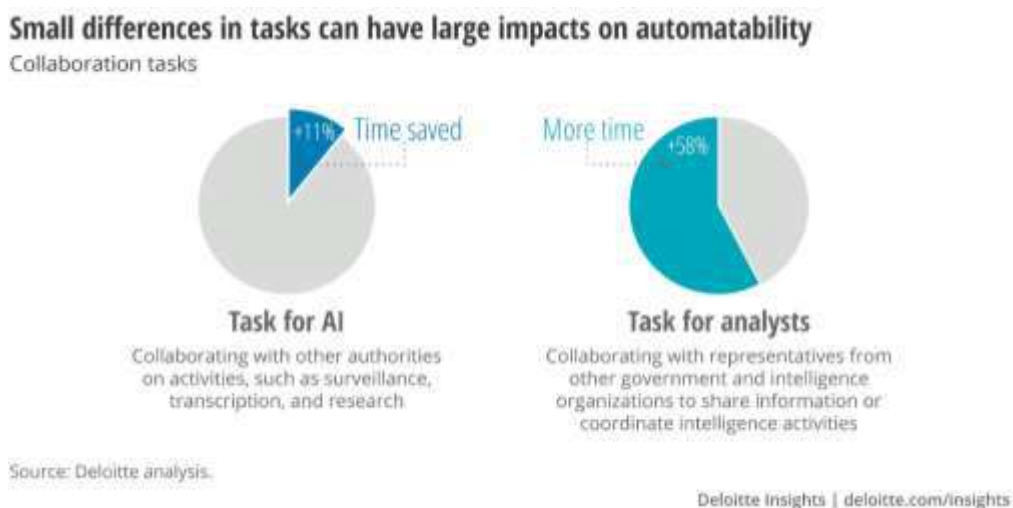
seguir invirtiendo en transformación digital y la adopción de tecnologías emergentes para asegurar competitividad, sostenibilidad y bienestar en el futuro cercano.

### **Impacto en el análisis de inteligencia de negocios**

Para la firma Deloitte, la IA ha pasado de ser casi ciencia ficción a una realidad común en diversas aplicaciones empresariales. En el análisis de inteligencia, la IA ya se está implementando para etiquetar imágenes y analizar grandes cantidades de datos, lo que ayuda a los humanos a distinguir la señal entre el ruido. Pero lo que la comunidad de inteligencia (CI) está haciendo ahora con la IA es solo un adelanto de lo que está por venir. Estas primeras aplicaciones apuntan a un futuro en el que la IA, implementada de forma inteligente, potenciará la capacidad de los analistas para extraer valor de la información. Esta empresa realizó un estudio que describe el tema a discusión:

#### **Figura 8**

*Comparativos de realización de tareas por IA vs tareas hechas por humanos*



Fuente: Tomado de Mitchell et al. (2019)

La imagen proporcionada ilustra cómo pequeñas diferencias en las tareas pueden tener un impacto significativo en la automatización, específicamente en el ámbito de la colaboración dentro del análisis de inteligencia. A través de dos gráficos circulares, se compara la automatización de tareas llevadas a cabo por IA frente a las realizadas por analistas humanos. El gráfico de la izquierda muestra que la IA puede ahorrar un 11 % de tiempo en tareas colaborativas como vigilancia, transcripción e investigación, lo que sugiere que estas actividades son parcialmente automatizables y que la tecnología ya está logrando eficiencias notables en tareas estructuradas y repetitivas.

En contraste, el gráfico de la derecha muestra que cuando la colaboración implica actividades más complejas —como coordinar con otros representantes gubernamentales y organizaciones de inteligencia, los analistas humanos todavía dedican un 58 % más de tiempo. Esto refleja la menor automatización de tareas que requieren juicio, interpretación de contexto y toma de decisiones estratégicas. A diferencia de las tareas técnicas, estas actividades colaborativas de alto nivel exigen habilidades interpersonales, pensamiento crítico y adaptabilidad, capacidades que actualmente la IA no puede replicar plenamente.

Este contraste evidencia que la integración efectiva de la IA en el análisis de inteligencia no implica sustituir al analista humano, sino redistribuir el trabajo. La IA puede encargarse de tareas repetitivas y de procesamiento masivo de datos, permitiendo que los analistas se enfoquen en el razonamiento estratégico, la interpretación contextual y la toma de decisiones críticas. A medida que las tecnologías maduren, será esencial que las organizaciones diseñen procesos colaborativos hombre-máquina, donde la IA actúe como un asistente que potencie las capacidades humanas, especialmente en ámbitos donde la automatización total no es viable ni deseable.

**Aplicación de la IA por sector industrial**

La empresa consultora Gartner elaboró un estudio que muestra el uso de la IA por sector industrial y por labor funcional dentro la compañía.

**Figura 9**

*Uso de la IA por sector industrial*

	Sectores							
	Automoción y fabricación de vehículos	Medios de comunicación	Arquitectura e ingeniería	Energía y servicios públicos	Sistema sanitario	Fabricación de productos electrónicos	Fabricación	Industria farmacéutica
Diseño de medicamentos								●
Ciencia de materiales	●			●		●		
Diseño de chips						●		
Datos sintéticos	●		●	●	●	●	●	●
Diseño generativo (de piezas)	●		●				●	

Fuente: Tomado de Wiles (2023)

Según la imagen anterior, se puede comprender que cada sector adopta la IA de manera distinta, dependiendo de sus necesidades y procesos clave. Por ejemplo, el diseño de medicamentos es

exclusivo de la industria farmacéutica, lo cual tiene sentido dado que esta aplicación requiere conocimientos altamente especializados en biotecnología y química. Por otro lado, el diseño de chips se limita a la fabricación de productos electrónicos, lo que refleja la especificidad técnica de esta tarea. La ciencia de materiales aparece en sectores como medios de comunicación y energía, lo cual puede estar relacionado con el desarrollo de nuevos materiales para dispositivos o infraestructuras.

Finalmente, el caso de uso más ampliamente distribuido es el diseño generativo (de piezas), que se aplica en al menos cinco sectores distintos. Esto sugiere que esta tecnología tiene una gran versatilidad, permitiendo optimizar el diseño de componentes físicos en industrias tan variadas como la automotriz, la ingeniería, la electrónica y la farmacéutica. Asimismo, los datos sintéticos también tienen una presencia transversal, lo cual indica su utilidad para entrenar modelos de IA en contextos donde los datos reales son escasos o sensibles. En conjunto, el gráfico demuestra cómo la IA generativa está transformando múltiples industrias al ofrecer soluciones adaptadas a sus desafíos particulares.

### ***Impacto de IA en Productividad y Toma de Decisiones***

Para Krajewski et al (2008) el uso de sistemas informáticos como la simulación es fundamental para decidir de manera acertada. El impacto de la inteligencia artificial en la productividad empresarial y en la calidad de la toma de decisiones se ha convertido en un tema de creciente interés en la literatura académica y en los estudios de consultoría. La Escuela de Negocios española IEBS ha desarrollado un estudio sobre la aplicación de la IA en diferentes sectores empresariales, cuyos resultados revelan un panorama amplio y diverso respecto al grado de adopción tecnológica. [Figura 10]

En el estudio, cada barra del gráfico de resultados se encuentra segmentada por colores que representan distintos niveles de adopción: desde aquellas empresas que aún no utilizan ni exploran aplicaciones de IA, hasta aquellas que han logrado desplegarla de manera completa y sistemática en sus procesos. Esta clasificación permite observar cómo la IA está siendo asimilada en fases diferenciadas, dependiendo tanto de la madurez digital de las organizaciones como de las características de cada industria.

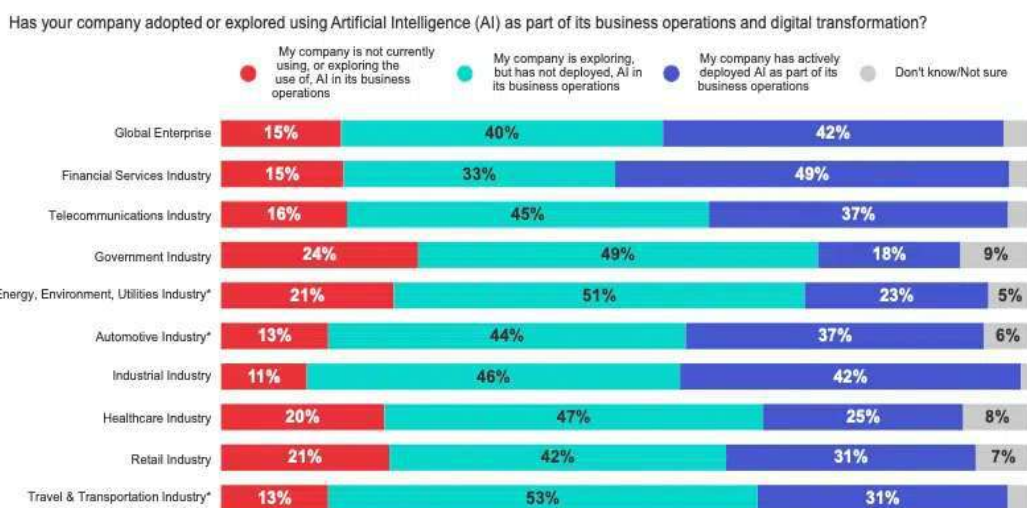
En conjunto, los resultados ofrecen una visión clara del estado actual de la transformación digital a nivel global. Se evidencia que, si bien algunos sectores como los servicios financieros y las telecomunicaciones muestran niveles avanzados de adopción, otros como el comercio minorista o el transporte aún permanecen en etapas tempranas. Esto implica que la IA no solo está redefiniendo los procesos internos de las organizaciones, sino que también está configurando nuevas brechas de competitividad entre industrias. Por tanto, el reto no se limita a la incorporación de tecnología,

sino también a la capacidad de diseñar estrategias que permitan aprovechar el potencial de la IA como una herramienta de apoyo a la productividad y como un motor de decisiones más rápidas, precisas y fundamentadas.

**Figura 10**

*Uso de la IA para la mejora de la productividad en toma de decisiones*

Enterprises within the financial services are most likely to be using AI, with nearly half of IT Professionals in that industry reporting their enterprise has actively deployed AI.



Base: IT Professionals at Enterprises (organizations = 1,035 employees). Global Enterprise = 2,242; Financial Services = 216; Telecommunications = 103; Government = 149; Energy = 156; Automotive = 88; Industrial = 302; Healthcare = 164; Retail = 139; Travel = 88. \*Sample size is between 50 and 99. Note: Media & Entertainment, Consumer/US/Car, and Aerospace & Defense industry samples sizes are too low to show.

Fuente: Tomado de López (2025).

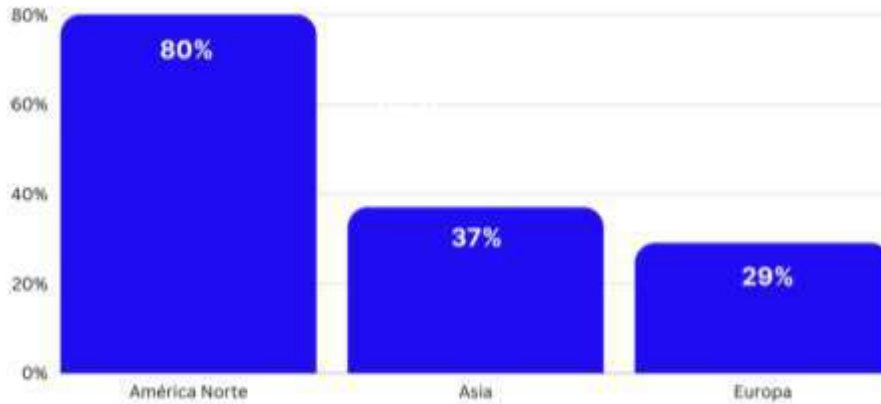
**Datos del ML a nivel mundial**

El aprendizaje automático o Machine Learning ha tomado un papel protagónico para algunas economías y sectores empresariales claves. La adopción del ML es un indicador clave de competitividad tecnológica y capacidad de innovación países que apuesten por esto se caracterizan por ser líderes en el sector comercial donde se desenvuelvan.

Por otro lado, países que estén con regulaciones más estrictas, menor inversión o una adopción más cautelosa de nuevas tecnologías pueden quedar rezagadas por esta tendencia tecnológica. Las siguientes imágenes presentan unas series de datos interesantes sobre el tema de la adopción del ML a nivel mundial:

**Figura 11**

*Adopción del Machine Learning a nivel geográfico*



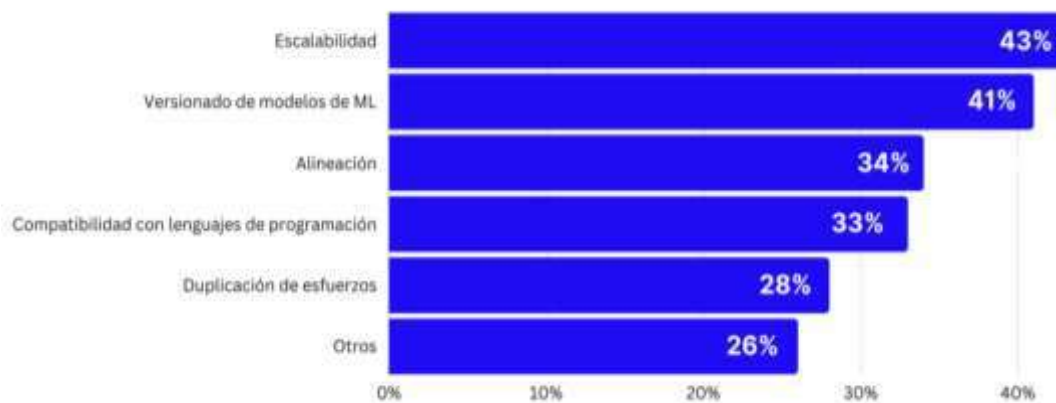
Fuente: Tomado de Gil (2023).

Lo entendido con el gráfico anterior es que existe una apertura completa en Norteamérica para la implementación del ML en sus actividades corporativas. Adicionalmente, la firma McKinsey asevera que el 51% de las organizaciones están en una fase inicial del ML y un 15% dice ser usuarios avanzados del ML (Gil, 2023).

Por otro lado, otra estadística llamativa es lo relacionado con los retos que conlleva la implementación del ML a nivel empresarial, los resultados muestran:

**Figura 12**

*Retos al implementar ML*



Fuente: Tomado de Gil (2023).

## **Conclusiones**

### ***Proyección económica***

En el sector de fábricas, el uso de IA y Machine Learning se enfoca especialmente en la optimización de operaciones, mantenimiento predictivo y gestión de inventarios, áreas que son pilares fundamentales en la ingeniería industrial. Esto permite una mejor asignación de recursos, reducción de tiempos muertos y aumento de la productividad.

Se estima que la inteligencia artificial generará entre \$3.9 y \$11.1 billones de dólares en valor económico para 2025, lo que evidencia su creciente papel estratégico en sectores clave de la ingeniería industrial, como recursos humanos, plantas industriales (factories) y sitios de trabajo (worksites).

### ***Automatización y tareas colaborativas***

Las tareas realizadas por sistemas de IA presentan un ahorro de tiempo de hasta 11%, mientras que las tareas ejecutadas por analistas humanos requieren 58% más tiempo. Esto resalta el alto potencial de la IA para automatizar procesos repetitivos o de vigilancia, especialmente en áreas administrativas o de análisis de datos dentro de la gestión operativa.

Aunque la IA puede realizar tareas más rápido, las interacciones complejas entre analistas y organismos para coordinar acciones siguen siendo menos automatizables. Esto implica que la IA debe ser vista como una herramienta de soporte estratégico, que libere tiempo a los profesionales para actividades de mayor valor agregado en la toma de decisiones y la mejora continua.

### ***Aplicabilidad en sectores***

Aplicaciones como el diseño generativo de piezas y los datos sintéticos son especialmente relevantes en sectores como fabricación, energía, automoción y defensa. Esto abre oportunidades en la ingeniería industrial para reducir tiempos de desarrollo, optimizar diseño de productos y simular escenarios sin necesidad de pruebas físicas costosas.

La presencia de IA en industria farmacéutica, automoción y servicios públicos evidencia su madurez en sectores donde la precisión, la seguridad y la trazabilidad son críticas, lo cual representa un reto y una oportunidad para la gestión de operaciones industriales avanzadas.

### ***Adopción por tipo de industria***

Aunque industrias como servicios financieros (48%) y telecomunicaciones (47%) tienen una alta adopción de IA, sectores industriales clave como manufactura (46%) y automoción (41%) también muestran una integración activa. Esto indica que la IA está trascendiendo sectores digitales para tener un impacto real en operaciones industriales físicas.

En sectores como el transporte y comercio minorista, se observa una baja implementación práctica, lo cual puede representar una oportunidad para los ingenieros industriales de liderar la transformación digital a través de la aplicación estratégica de IA y ML.

### ***Adaptación por zona geográfica***

América del Norte refleja un ecosistema tecnológico más maduro, con mayor inversión en infraestructura digital, talento especializado y cultura de innovación. Esta alta adopción sugiere que las organizaciones norteamericanas están aprovechando el ML para optimizar procesos industriales, reducir costos operativos y mejorar la toma de decisiones estratégicas. En contraste, Asia y Europa presentan niveles de adopción más bajos, esto puede deberse a factores como regulaciones más estrictas o una menor integración entre la academia, la industria y el gobierno. Esta brecha tecnológica puede traducirse en una menor competitividad industrial si no se abordan con políticas de transformación digital más agresivas.

Desde la perspectiva de la ingeniería industrial, esta diferencia en la adopción implica que el impacto del ML en la gestión de operaciones no es uniforme a nivel global. Mientras algunas regiones avanzan hacia modelos de producción inteligentes y adaptativos (Industria 5.0), otras aún enfrentan barreras estructurales que limitan su evolución más allá de la Industria 4.0.

### ***Retos del ML y la IA***

Si bien el potencial del ML en la ingeniería industrial es ampliamente reconocido, su implementación enfrenta barreras técnicas y organizativas significativas. El principal reto identificado es la escalabilidad (43%), lo cual indica que muchas organizaciones logran desarrollar modelos funcionales en entornos controlados, pero encuentran dificultades al intentar llevarlos a producción a gran escala. Esto es especialmente relevante en la ingeniería industrial, donde los sistemas deben operar con grandes volúmenes de datos en tiempo real y bajo condiciones variables.

El versionado de modelos (41%) también representa un desafío crítico. En entornos industriales, donde la trazabilidad y la reproducibilidad son esenciales para la mejora continua y la auditoría de procesos, la falta de control sobre las versiones de modelos puede comprometer la confiabilidad de las decisiones automatizadas. Asimismo, la alineación entre los modelos de ML y los objetivos del negocio (34%) sugiere una desconexión entre los equipos técnicos y estratégicos, lo que puede derivar en soluciones que, aunque técnicamente sólidas, no generan valor operativo. Otros retos como la compatibilidad entre lenguajes de programación (33%) y la duplicación de esfuerzos (28%) reflejan la necesidad de una mayor estandarización y colaboración entre equipos. Estos problemas limitan la integración fluida de modelos de IA en sistemas como ERP o CRM, fundamentales en la gestión de operaciones industriales.

## Referencias

- Amendiola, L. (2022, 2 de marzo). *¿Por qué la industria 5.0? ¿Qué opinan los profesionales?* [Blog] <https://www.linkedin.com/pulse/por-qu%C3%A9-la-industria-50-opinan-los-profesionales-amendola-ph-d/>
- Buttle, F., & Maklan, S. (2019). *Customer Relationship Management: Concepts and Technologies* (4th ed.). Taylor and Francis Group.
- Cabanillas, M., Rivadeneyra, R., Palacios, C. & Hernández, B. (2020). *Habilidades socioemocionales en las instituciones educativas*. SciComm. Universidad Autónoma de Chile.
- Chapman, S. (2006). *Planificación y Control de la producción*. Pearson Education. Pearson / Prentice Hall.
- Chopra S., & Meindl P. (2013). *Administración de la Cadena de Suministro: Estrategia, Planeación y Operación* (5ta. Ed.). Pearson Education.
- Coto M. (2021). Consideraciones para la incorporación de la Inteligencia Artificial en un programa de pregrado de Ingeniería Eléctrica. *Revista Actualidades Investigativas en Educación*, 21 (2), 1-25. <https://www.kerwa.ucr.ac.cr/server/api/core/bitstreams/8b6c33c4-9b44-499f-8a0a-3229d8dce555/content>
- Forero W. & Negre F. (2024). Técnicas y aplicaciones del Machine Learning e Inteligencia Artificial en educación: una revisión sistemática. *Revista Iberoamericana de Educación a Distancia*, 27 (1), 209-253. DOI: <https://doi.org/10.5944/ried.27.1.37491>
- Gil, C. (2023, 31 de agosto). Estadísticas de machine learning: Tendencias a conocer en 2023. [Página web Sortlist] <https://www.sortlist.es/datahub/reports/estadisticas-de-machine-learning/>
- Heizer J., & Render B. (2007). *Dirección de la Producción y Operaciones. Decisiones estratégicas* (8va. Ed.). Pearson Education.
- Initium Software. (2024, 21 de agosto). Maximizando la eficiencia empresarial: ¿Para qué sirve el CRM? [Página web] [https://www.initiumsoft.com/blog\\_initium/para-que-sirve-el-crm/](https://www.initiumsoft.com/blog_initium/para-que-sirve-el-crm/)
- Ivanov D., Dolgui A. & Sokolov B. (2018). The impact of digital technology and Industry 4.0 on the ripple effect and supply chain risk analytics. *International Journal of Production Research*, 57 (3), 829-846. DOI: <https://doi.org/10.1080/00207543.2018.1488086>
- Jacobs R., & Chase R. (2018). *Administración de Operaciones: Producción y Cadena de Suministros* (15va. Ed.). McGraw Hill.

- Kotler P., & Keller, K. (2016). *Marketing Management* (15th ed.). Pearson Education.
- Krajewski L., Ritzman L., and Malhotra M. (2008). *Administración de Operaciones: Procesos y Cadena de valor* (8va. Ed.). Pearson Education.
- Lauramajon4. (2018, 17 de marzo). *Software a medida*. [Blog *Diarium Blogdelauramajon4*] <https://diarium.usal.es/lauramajon4/2018/03/17/software-a-medida/>
- López, S. (2025, 11 de febrero). Estadísticas Esenciales de Inteligencia Artificial para 2025: ¿Quién la usa y para qué? [IEBS Business School página web] <https://www.iebschool.com/hub/estadisticas-esenciales-de-inteligencia-artificial-para-2025-quien-la-usa-y-para-que-tecnologia/>
- Luan, H. & Tsai, C. (2021). A Review of Using Machine Learning Approaches for Precision Education. *Educational Technology and Society*, 24 (1), 250-266. <https://www.jstor.org/stable/26977871>
- Mitchell, K., Routh, A., Mirkow, A., Mariani, J., & Keyal, A. (2019). The future of intelligence analysis. A task-level view of the impact of artificial intelligence on intel analysis [Deloitte página web]. <https://www.deloitte.com/us/en/insights/industry/government-public-sector-services/artificial-intelligence-impact-on-future-intelligence-analysis.html>
- Nahavandi, S. (2019). Industry 5.0: A human-centric solution. *Sustainability*, 11(16), 4371- <https://doi.org/10.3390/su11164371>
- SAP. (2025). ¿Qué es machine learning? [Página web] <https://www.sap.com/latinamerica/products/artificial-intelligence/what-is-machine-learning.html>
- Schwab, K. (2016). *The Fourth Industrial Revolution*. World Economic Forum. <https://www.weforum.org/about/the-fourth-industrial-revolution-by-klaus-schwab>
- Troncoso M., Duenas Y. & Verdecia E. Inteligencia artificial y educación: nuevas relaciones en un mundo interconectado. *Revista Estudios del Desarrollo Social: Cuba y América Latina*, 11 (2). [http://scielo.sld.cu/scielo.php?script=sci\\_arttext&pid=S2308-01322023000200014&lng=es&tlng=es](http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S2308-01322023000200014&lng=es&tlng=es)
- Wamba, S. F., Gunasekaran, A., Akter, S., Ren, S. J.-F., Dubey, R., & Childe, S. J. (2020). Big data analytics and firm performance: Effects of dynamic capabilities. *Journal of Business Research*, 70, 356–365. <https://doi.org/10.1016/j.jbusres.2016.08.009>
- Verma, D. (2014). Industry 5.0: A human-centric and sustainable approach to industrial development. *International Journal of Social Relevance & Concern*, 12 (5), 17-21. <https://doi.org/10.26821/IJSRC.12.5.2024.120507>

Wiles, J. (2023, 26 de enero). *Por muy revolucionario que sea, ChatGPT es solo el principio; los usos de la IA generativa en la empresa pueden ser mucho más sofisticados* [Gartner página web].

<https://www.gartner.es/es/articulos/mas-alla-de-chatgpt-el-futuro-de-la-ia-generativa-para-empresas>




Williams, I. (2025). *¿Estamos a las puertas de una industria 5.0?* [PRLInnovación página web]

<https://www.prlinnovacion.com/tecnologia-prevencion-estamos-a-las-puertas-de-una-industria-5-0>

# Aplicación de modelos de colas y simulación para evaluar la eficiencia del servicio de helados en McDonald's Chapinero

## Application of Queueing Models and Simulation to Evaluate the Efficiency of the Ice Cream Service at McDonald's Chapinero

---

Laura Sofía Guerrero Guavita<sup>1</sup>, Gabriela Zipaquirá Antolínez<sup>2</sup> & Francy Rocío Castellanos Oviedo<sup>3</sup>

---

Fecha de recepción: 12-08-2025 | Fecha de aprobación: 07/10/2025

### Resumen

McDonald's es una cadena de comida rápida reconocida mundialmente por la eficiencia operativa y estandarización en sus procesos. En Chapinero, Bogotá, se estableció un punto de venta especializado en helados, aprovechando su ubicación estratégica en una zona comercial y estudiantil con alto tránsito peatonal. Este punto, enfocado en productos como McFlurrys, conos y sundaes, generó inicialmente largas filas y tiempos de espera elevados, lo cual motivó la presente investigación.

El estudio tuvo como propósito analizar y optimizar el desempeño operativo del punto de venta mediante modelos de investigación de operaciones e inteligencia artificial. Se recopiló información real sobre tiempos de llegada y atención al cliente, analizando estadísticamente estos datos para identificar la distribución que mejor los representaba, encontrándose la distribución Gamma como la más apropiada. A partir de estos resultados, se realizaron simulaciones computacionales en Python y FlexSim para evaluar diferentes indicadores críticos como tiempo promedio de espera, tasa de servicio y nivel de saturación del sistema. Estas simulaciones permitieron visualizar claramente los procesos internos, identificar cuellos de botella y generar alternativas operativas viables.

- 
- 1 Estudiante Facultad de Ingeniería Industrial, Universidad Santo Tomás, Bogotá, Colombia. Grupo de Investigación en Procesos Organizacionales – GIPO, Semillero SIGEO. Contacto: [laura.guerreo@usantotomas.edu.co](mailto:laura.guerreo@usantotomas.edu.co), celular: 3124964843.
  - 2 Estudiante Facultad de Ingeniería Industrial, Universidad Santo Tomás, Bogotá, Colombia. Grupo de Investigación en Procesos Organizacionales – GIPO, Semillero SIGEO. Contacto: [zipaquiragabriela@usantotomas.edu.co](mailto:zipaquiragabriela@usantotomas.edu.co)
  - 3 Ingeniera Industrial, Magíster en Ingeniería – Ingeniería Administrativa. Docente Facultad de Ingeniería Industrial Universidad Santo Tomás, Bogotá, Colombia. Grupo de Investigación en Procesos Organizacionales – GIPO, Semillero SIGEO. Contacto: [francycastlellanos@usta.edu.co](mailto:francycastlellanos@usta.edu.co)

Finalmente, el análisis permitió proponer mejoras concretas relacionadas con la capacidad instalada, la asignación efectiva de personal y la distribución óptima de tareas en función de la demanda real, logrando así una significativa reducción en los tiempos de espera y una mejora sustancial en la eficiencia del servicio ofrecido a los clientes.

### **Palabras Clave**

Teoría de colas, Simulación, Tiempo de espera, Eficiencia Operativa, Optimización

### **Abstract**

McDonald's is a globally recognized fast-food chain known for operational efficiency and standardized processes. In Chapinero, Bogotá, a specialized ice cream outlet was established, strategically located in a busy commercial and student area with high pedestrian traffic. This outlet, focused on products such as McFlurrys, cones, and sundaes, initially experienced long queues and extensive wait times, prompting this research.

The study aimed to analyze and optimize the operational performance of the outlet using advanced industrial engineering techniques and artificial intelligence. Real data on customer arrival and service times were collected and statistically analyzed to identify the distribution that best represented these times, determining the Gamma distribution as the most suitable. Subsequently, computational simulations were performed using Python and FlexSim to evaluate critical indicators such as average waiting time, service rate, and system saturation levels. These simulations provided clear visualization of internal processes, identified bottlenecks, and facilitated the exploration of alternative operational scenarios.

Finally, the analysis led to concrete improvement proposals related to installed capacity, efficient personnel allocation, and optimal task distribution based on actual demand. These changes significantly reduced wait times and substantially enhanced service efficiency and customer experience.

### **Keywords**

Queueing theory, Simulation, Waiting time, Operational efficiency, Optimization

### **Abreviaciones:**

- SAS: *Statistical Analysis System*. Software utilizado para el análisis estadístico y ajuste de distribuciones.
- M/M/1: Modelo de cola con llegadas Poisson (M), tiempos de servicio exponenciales (M) y un solo servidor (1).

- M/M/c: Modelo de cola con llegadas Poisson (M), tiempos de servicio exponenciales (M) y múltiples servidores (c).
- M/G/1: Modelo de cola con llegadas Poisson (M), tiempos de servicio con distribución general (G) y un solo servidor (1).
- M/G/k: Modelo de cola con llegadas Poisson (M), tiempos de servicio con distribución general (G) y múltiples servidores (k).
- KPI: *Key Performance Indicator*. Indicadores clave de rendimiento usados para evaluar la eficiencia del sistema.
- AD: *Anderson-Darling*. Prueba estadística para evaluar el ajuste de distribuciones.
- K-S: *Kolmogorov-Smirnov*. Prueba de bondad de ajuste utilizada para comparar distribuciones.
- $W^2$ : *Cramér-von Mises*. Prueba estadística para evaluar la bondad de ajuste de una distribución.

## Introducción

El análisis de la capacidad operativa y la eficiencia desempeña un papel crucial en el sector de servicios, particularmente en aquellos contextos donde la demanda es variable y la atención oportuna es un factor determinante en la satisfacción del cliente. En este ámbito, se destacan herramientas fundamentales como la teoría de colas, la simulación de eventos discretos y el análisis de distribuciones de probabilidad, que permiten modelar, analizar y optimizar procesos de servicio, como es el caso del sector de comidas rápidas.

La industria de comida rápida enfrenta desafíos crecientes para satisfacer las expectativas de los clientes en términos de rapidez y calidad del servicio. Los largos tiempos de espera representan un reto clave para los gerentes, quienes buscan mejorar la percepción del cliente y garantizar un servicio eficiente. Como señala Villarreal et al. (2021),

tener una larga espera para acceder a múltiples servicios se cataloga como uno de los principales retos de los gerentes, directores y aquellos que toman decisiones para mejorar esa percepción de los usuarios y de esta manera responder a las quejas que a las empresas les permita generar un mejoramiento continuo en el servicio que brindan a sus clientes.

En el contexto de la pandemia de COVID-19, las cadenas de comida rápida han adoptado tecnologías como quioscos de autoservicio, pedidos móviles y aplicaciones de entrega a domicilio para reducir la presión sobre los puntos de venta físicos, similar a la virtualización observada en sectores como las agencias de viajes (Villarreal et al., 2021). Estas innovaciones reflejan la necesidad de optimizar recursos limitados frente a una demanda variable, un problema central en el punto de helados de McDonald's Chapinero, donde las largas filas y tiempos de espera prolongados motivaron esta investigación.

La teoría de colas ha sido ampliamente utilizada en Ingeniería Industrial para describir sistemas en los que se presentan llegadas aleatorias de entidades que requieren atención limitada por uno o varios servidores. Al comprender métricas como el tiempo promedio en cola, la tasa de servicio y la longitud media del sistema, es posible identificar cuellos de botella y proponer mejoras operativas (Gross & Harris, 2008). En este contexto, la distribución Gamma ha demostrado ser útil para modelar tiempos de servicio variables, especialmente cuando el proceso presenta asimetrías o desviaciones significativas respecto a la media (Walpole et al., 2012).

La teoría de colas se ha convertido en un método analítico clave para abordar problemas complejos y dinámicos, permitiendo analizar características específicas como el número medio de entidades en espera y el tiempo medio de permanencia en el sistema según (Camelo et al., 2010). Existen diversos modelos desde el punto de vista de la Investigación de operaciones; éstos se diferencian por características como: la distribución que siguen los tiempos entre llegadas y los tiempos de servicio, la población, la disciplina de la cola, la capacidad total del sistema y la cantidad de servidores. Particularmente los del tipo M/M/1, son efectivos para analizar tasas de servicio que se distribuyen exponencialmente relacionadas normalmente con sistemas de atención al cliente (Narvéez-Gómez et al., 2018)

Además, la teoría de colas permite cuantificar el equilibrio entre la eficacia en la prestación de servicios y la eficiencia operativa, identificando claramente los parámetros clave para evaluar y optimizar los sistemas de servicio (Singer et al., 2008). De igual manera, la gestión eficiente de las líneas de espera reduce significativamente los tiempos de espera, incrementando la satisfacción del cliente y disminuyendo pérdidas económicas asociadas a la mala calidad del servicio (Villarreal et al., 2021).

La simulación de sistemas de transporte público mediante modelos de redes de colas abiertas permite evaluar el desempeño integral del sistema, incluyendo la estimación precisa de tiempos de espera y longitud de filas (Ortiz & Serrano, 2006). Finalmente, la correcta aplicación de la teoría de colas y simulación en instituciones financieras ha demostrado ser efectiva para mejorar procesos críticos, como los tiempos de atención al cliente (Gómez, 2008).

La simulación computacional es otra herramienta clave para la evaluación de sistemas de servicio. Según Law & Kelton (2014), la simulación permite replicar el comportamiento de procesos reales mediante la construcción de modelos virtuales, facilitando la experimentación sin alterar el entorno físico. La combinación de simulación en lenguajes de programación como Python y entornos gráficos como FlexSim permite tanto el análisis cuantitativo como la visualización tridimensional de flujos operativos.

La simulación dinámica de procesos permite representar sistemas complejos a través de modelos comprensibles, que apoyan decisiones estratégicas y tácticas (Roark et al., 2019). En este mismo

sentido, la simulación de eventos discretos ayuda a determinar requerimientos cuantitativos esenciales, considerando la aleatoriedad de los procesos y sus interacciones (Pérez & Riaño, 2007).

Además, la incorporación de técnicas de análisis estadístico apoyadas en software como SAS Studio ha permitido mejorar la calidad del ajuste de datos recolectados en campo. Mediante pruebas de bondad de ajuste como Kolmogorov-Smirnov, Cramér-von Mises y Anderson-Darling, es posible validar la representatividad de una distribución teórica frente a datos empíricos, fortaleciendo la credibilidad del modelo (SAS Institute Inc., 2024).

Llorente et al. (2001), también resaltan el valor de la simulación en sistemas de atención, subrayando cómo aumentar recursos disponibles no siempre implica mejoras significativas en tiempos de espera, destacando así la importancia de modelos precisos para decisiones operativas eficientes.

Diversos estudios aplicados al sector de alimentos han demostrado que la simulación de sistemas de atención en puntos de venta permite identificar con claridad los momentos críticos del día, optimizar la asignación de personal y mejorar la experiencia del cliente (Toneguzzi, 2022; Corral González, 2024). En el caso particular de McDonald's, su modelo operativo altamente estandarizado constituye un entorno ideal para aplicar herramientas de ingeniería industrial con fines de mejora continua.

McDonald's es una de las cadenas de comida rápida más reconocidas a nivel mundial, destacada por su eficiencia operativa y estandarización de procesos (Corral, 2024). La sede ubicada en Chapinero, Bogotá, fue seleccionada como caso de estudio por su ubicación estratégica en una zona comercial y estudiantil con alto flujo peatonal. Para atender esta demanda, se implementó un punto de venta exclusivo para helados hacia la calle, con productos como McFlurrys, conos y sundaes (Toneguzzi, 2022). Sin embargo, esta estrategia generó largas filas y tiempos de espera considerables, lo cual motivó esta investigación. El objetivo fue caracterizar el rendimiento del servicio mediante herramientas de inteligencia artificial aplicadas a la ingeniería industrial. Se utilizó el software SAS para analizar estadísticamente los tiempos entre llegadas y atención al cliente (SAS Institute, 2025). Posteriormente, se desarrolló una simulación en Python (Google Colab) para estimar indicadores como tiempo promedio de espera, tasa de servicio y saturación del sistema. Finalmente, se integró una simulación en FlexSim para visualizar el proceso, detectar cuellos de botella y evaluar escenarios operativos alternativos.

## Metodología

La presente investigación se desarrolló bajo un enfoque cuantitativo, utilizando técnicas de análisis estadístico y simulación computacional, con el fin de evaluar y proponer mejoras en el rendimiento operativo del punto de venta exclusivo de postres en McDonald's Chapinero, Bogotá. La simulación, como herramienta clave en la investigación de operaciones, permitió modelar la complejidad del

sistema de atención al cliente, siguiendo los principios descritos por Banks et al. (2010), quienes destacan su capacidad para estudiar interacciones internas y experimentar con cambios operativos. De acuerdo con Law (2019), el diseño de modelos de simulación requiere una estructura sistemática que integre la recolección de datos, el análisis estadístico y la experimentación controlada, aspectos que guiaron el desarrollo del presente estudio. En esta misma línea, Portilla et al. (2010) subrayan que el diseño metodológico debe considerar la variabilidad del entorno operativo para garantizar que los resultados de la simulación sean representativos y útiles para la toma de decisiones.

El trabajo se desarrolló en tres fases principales: recolección de datos, ajuste estadístico y simulación computacional, descritas a continuación.

### ***Recolección y análisis de datos***

La fase inicial consistió en una observación directa del proceso de atención, recolectando datos sobre los tiempos entre llegadas y los tiempos de servicio (en caja y máquina de postres). La recolección se llevó a cabo durante el mes de marzo de 2025, con la participación de dos observadores capacitados que registraron manualmente los tiempos utilizando cronómetros digitales y formatos pre estructurados diseñados para el estudio. Dichos formatos incluían columnas para registrar: hora exacta de llegada del cliente, inicio y fin de atención en caja, inicio y fin de preparación en máquina de postres, y hora de salida del sistema.

Las jornadas de observación incluyeron tres días laborales, con el objetivo de capturar tanto el comportamiento regular del sistema como las fluctuaciones de demanda en días de alta afluencia. Los horarios monitoreados abarcaron desde las 11:00 a.m. hasta las 4:00 p.m., priorizando las horas pico (12:00 p.m. a 2:00 p.m.), identificadas previamente como las más críticas por la gerencia del local. Según Raj (2025), incorporar picos horarios en modelos de colas permite representar de manera más realista la variabilidad de la demanda y proponer intervenciones temporales (*staggered start times*) que mejoran la precisión y aplicabilidad del análisis. Esta evidencia respalda la importancia de incluir franjas horarias intercaladas y contextos de alta demanda en los estudios de líneas de espera.

Para garantizar la calidad de los datos, se adoptó un protocolo de doble registro, en el cual ambos observadores anotaron de manera independiente los mismos eventos y posteriormente los registros fueron validados mediante comparación cruzada, corrigiendo discrepancias mayores a más o menos 2 segundos. Además, los observadores fueron capacitados para minimizar sesgos, evitando interacciones con el personal y los clientes durante las mediciones, garantizando que los datos reflejaran el funcionamiento natural del sistema.

Este proceso permitió capturar promedios, variabilidad y patrones temporales en los datos, elementos indispensables para una simulación precisa. Como subrayan Villarreal et al. (2021), “un problema central en muchos contextos de servicios es la administración del tiempo de espera”,

razón por la cual fue clave medir con precisión variables como tiempo en cola, tiempo en el sistema y tiempos de servicio en cada etapa.

### **Ajuste estadístico mediante SAS**

En la segunda fase, los datos recolectados fueron analizados con el software SAS 9.4, siguiendo un flujo estructurado:

- Carga y depuración de datos: Los registros fueron importados y se verificó la ausencia de datos atípicos o inconsistentes.
- Estimación de parámetros: Para cada conjunto de datos (tiempos entre llegadas y tiempos de servicio), se estimaron los parámetros de las distribuciones candidatas utilizando el método de máxima verosimilitud.
- Pruebas de bondad de ajuste: Se aplicaron tres pruebas ampliamente utilizadas en investigación operativa:
  - *Kolmogorov-Smirnov (K-S)*: Detecta diferencias máximas entre la distribución empírica y la teórica.
  - *Anderson-Darling ( $A^2$ )*: Da mayor peso a los extremos, útil en distribuciones con colas pesadas.
  - *Cramér-von Mises ( $W^2$ )*: Mide discrepancias promedio entre las dos distribuciones.

Se evaluaron las distribuciones Poisson, Beta, Exponencial, Gamma, Lognormal y Weibull, tanto para los tiempos entre llegadas como para los tiempos de servicio. Este conjunto de pruebas permitió documentar con precisión el ajuste estadístico de los datos y sentar las bases para la selección posterior de las distribuciones más representativas. Tal como indican Walpole et al. (2012), la correcta identificación de las distribuciones subyacentes es fundamental para la validez de los modelos de simulación.

Cabe destacar que la interpretación de los valores p y la selección final de las distribuciones se abordaron en la sección de resultados, manteniendo esta fase como puramente metodológica y libre de inferencias.

### **Simulación computacional**

Con los parámetros derivados del análisis estadístico, se desarrollaron dos modelos complementarios de simulación:

- Simulación en Python (Google Colab): Se programó un sistema de colas para estimar métricas como utilización del servidor, longitud promedio de cola y tiempo promedio en el sistema. Se modelaron dos escenarios: M/G/1 (un servidor) y M/G/2 (dos servidores), incorporando las distribuciones obtenidas en la fase estadística. Este enfoque permitió realizar cálculos analíticos rápidos y comparar configuraciones del recurso humano.

- Simulación en FlexSim: Se creó un modelo tridimensional que replicó el proceso real de atención. El modelo incluyó elementos cómo:
  - o Source: Para representar las llegadas de clientes según los tiempos de interarribo observados.
  - o Queue: Para las colas de caja y máquina de postres, midiendo longitud y tiempo de espera.
  - o Processor: Uno para la caja (registro y cobro de pedidos) y otro para la máquina de postres (preparación y entrega).
  - o Sink: Para la salida del sistema tras la atención.

El operador principal fue configurado con funciones multitarea, simulando su desplazamiento entre caja y máquina, mientras que un segundo operador fue incorporado dinámicamente en las horas pico mediante un trigger de cambio de estado que modificaba su disponibilidad entre las 12:00 p.m. y las 2:00 p.m. Esta decisión metodológica se fundamenta en el enfoque de programación flexible propuesto por Zhang & Guhathakurta (2020), quienes destacan la eficacia del personal multihabilidad para mitigar cuellos de botella en entornos con demanda fluctuante.

El uso combinado de modelos analíticos en Python y modelos dinámicos en FlexSim responde a lo señalado por Negahban & Smith (2014), quienes recomiendan la integración de simulaciones de diferente naturaleza para obtener una visión más completa del sistema: mientras los modelos analíticos permiten cálculos rápidos y evaluación de escenarios, los modelos visuales facilitan la identificación de cuellos de botella y la validación empírica del flujo operativo.

Con estas tres fases, el estudio estableció una metodología sólida, combinando observación empírica, análisis estadístico y modelado computacional, creando un marco robusto para la evaluación del desempeño operativo del sistema y la formulación de estrategias de mejora.

## **Resultados**

A continuación, se presentan los principales resultados:

### ***Análisis de distribuciones***

Inicialmente se realizó la recolección de datos. Se hizo observación directa durante el mes de Marzo de 2025, de los datos de hora de llegada de un cliente a la fila, del tiempo de inicio de atención en caja, del tiempo de servicio en caja y de despacho del helado.

Con base en esta información se hicieron los cálculos de tiempos entre llegadas y tiempos de servicio. Para el análisis de distribuciones que siguen estos tiempos en el sistema de colas del mostrador de postres de McDonald's Chapinero, se evaluaron varias distribuciones teóricas usando SAS Studio: Beta, Exponencial, Gamma, Lognormal y Weibull. Se aplicaron las pruebas de

bondad de ajuste Kolmogorov-Smirnov, Cramer-von Mises y Anderson-Darling para determinar la mejor adaptación de los datos.

**Tabla 1**

*Distribución Gamma ajustada para le tiempo de servicio*

**Distribución Gamma ajustada para Tiempo de servicio (Tiempo de servicio)**

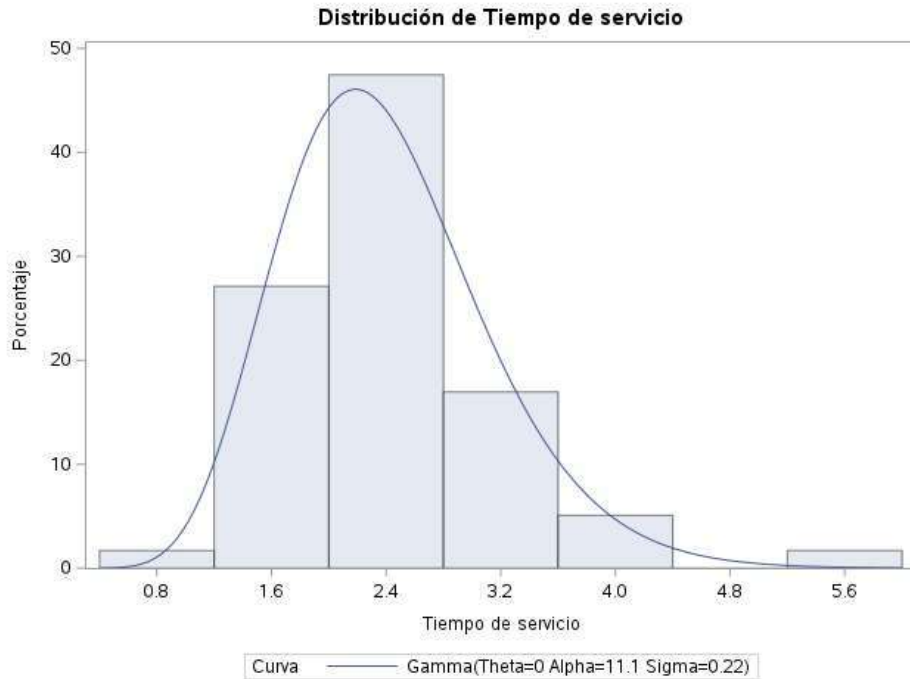
Tests de bondad de ajuste para la distribución Gamma				
Test	Estadístico		P valor	
Kolmogorov-Smirnov	D	0.08859021	Pr > D	>0.250
Cramer-von Mises	W-Sq	0.08991075	Pr > W-Sq	0.156
Anderson-Darling	A-Sq	0.58617865	Pr > A-Sq	0.132

Fuente: Elaboración propia.

La distribución Gamma mostró un buen ajuste, con valores p superiores a 0.05 en todas las pruebas (K-S: 0.250,  $W^2$ : 0.156,  $A^2$ : 0.132), lo que indica que no se rechaza la hipótesis de que los datos provienen de esta distribución. Por lo tanto, la distribución Gamma es adecuada para modelar el tiempo de servicio en este sistema de colas.

**Figura 2**

*Distribución Gamma ajustada para el tiempo de servicio*



Fuente: Elaboración propia.

La distribución Lognormal también presentó un buen ajuste a los datos de tiempo de servicio. Los resultados de las pruebas de bondad de ajuste fueron:

- Kolmogórov-Smirnov:  $D = 0.0863$ ,  $Pr > D = > 0.150$
- Cramer-Von Mises:  $W^2 = 0.0785$ ,  $Pr > W^2 = 0.220$
- Anderson-Darling:  $A^2 = 0.5642$ ,  $Pr > A^2 = 0.142$

Dado que los valores p en todas las pruebas son superiores a 0.05, no se rechaza la hipótesis nula, lo que indica que la distribución Lognormal también modela adecuadamente el tiempo de servicio.

**Tabla 2**

*Distribución Lognormal ajustada para el Tiempo de Servicio*

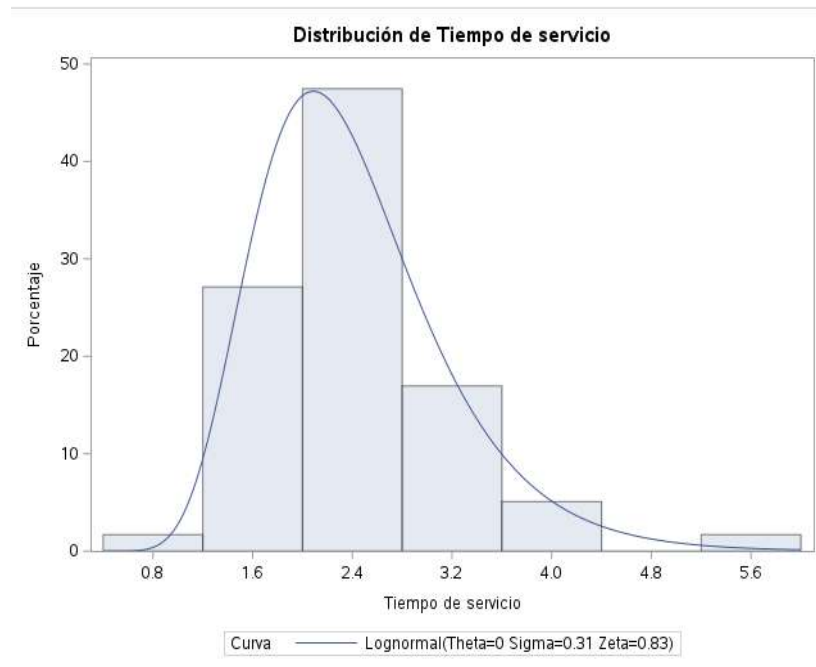
Distribución Lognormal ajustada para Tiempo de servicio (Tiempo de servicio)

Tests de bondad de ajuste para la distribución Lognormal				
Test	Estadístico		P valor	
Kolmogorov-Smimov	D	0.08630562	Pr > D	>0.150
Cramer-von Mises	W-Sq	0.07846358	Pr > W-Sq	0.220
Anderson-Darling	A-Sq	0.56424506	Pr > A-Sq	0.142

Fuente: Elaboración propia.

**Figura 2**

*Distribución Lognormal ajustada para el Tiempo de Servicio*



Fuente: Elaboración propia.

Las distribuciones Exponencial, Beta y Weibull no cumplieron con los criterios mínimos de ajuste para modelar el tiempo de servicio, ya que en todos los casos el test de Anderson-Darling arrojó valores p menores a 0.05, lo que indica diferencias significativas entre los datos reales y las distribuciones teóricas.

En cuanto al tiempo entre llegadas, se evaluaron las mismas distribuciones (Beta, Exponencial, Gamma, Lognormal y Weibull), pero ninguna presentó un ajuste adecuado, ya que todas obtuvieron

valores  $p$  inferiores a 0.05 en el test de Anderson-Darling. Esto sugiere que el comportamiento del tiempo entre llegadas no sigue un patrón probabilístico tradicional.

### **Análisis de la simulación**

Comenzando con las simulaciones, la simulación en Python, implementada en Google Colab, permitió analizar el desempeño del sistema de colas del punto de venta de helados de McDonald's Chapinero, utilizando los datos recolectados durante marzo de 2025. A partir de una muestra de 59 clientes, se calcularon métricas iniciales del sistema: un total de 59 clientes atendidos, un tiempo promedio en cola de 1.87 minutos, un tiempo promedio en el sistema de 4.28 minutos, un tiempo promedio de servicio de 2.40 minutos, y un tiempo promedio entre llegadas de 6.74 minutos. Estas métricas reflejaron la congestión observada en el punto de venta, especialmente durante las horas pico (12:00 p.m. a 2:00 p.m.), motivando la simulación para evaluar mejoras operativas.

### **Figura 3**

*Métricas reales del sistema*

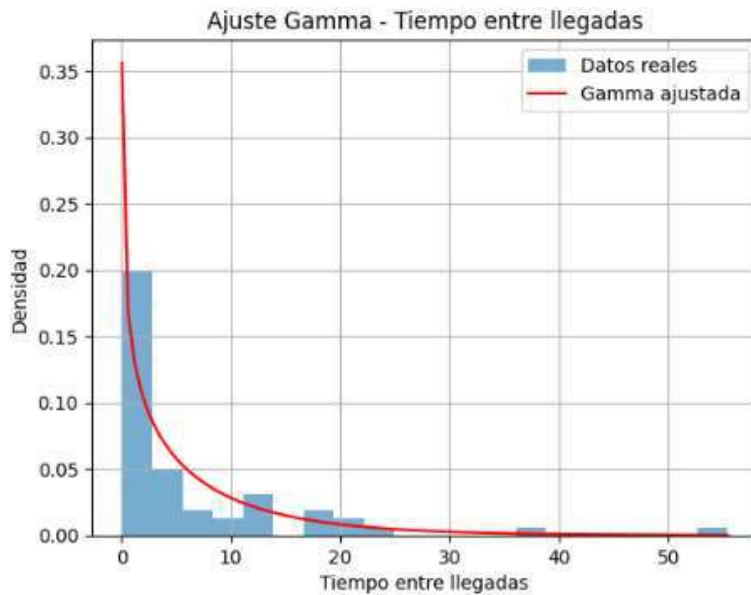
```
Métricas generales del sistema:  
Clientes totales: 59  
Tiempo promedio en cola: 1.87  
Tiempo promedio en el sistema: 4.28  
Tiempo promedio de servicio: 2.40  
Tiempo promedio entre llegadas: 6.74
```

Fuente: Elaboración propia en base a los datos.

El ajuste de la distribución Gamma se realizó para modelar los tiempos entre llegadas y los tiempos de servicio. Para los tiempos entre llegadas, se obtuvieron parámetros de shape ( $k$ ) = 0.7051 y scale ( $\theta$ ) = 9.7210, indicando una alta variabilidad en las llegadas de clientes. La Figura 4 muestra el histograma de los datos reales junto con la distribución Gamma ajustada, confirmando un ajuste visual adecuado, aunque menos robusto que para los tiempos de servicio, como se observó en el análisis con SAS Studio.

### **Figura 1**

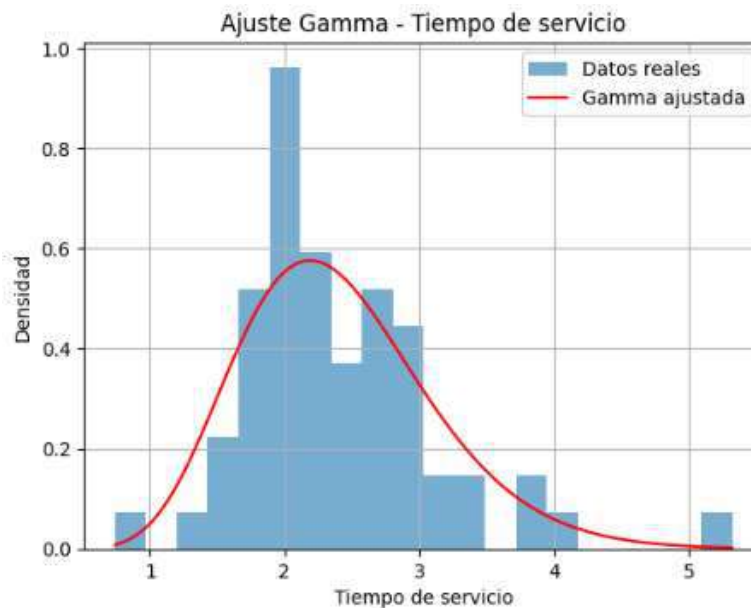
*Distribución Gamma ajustada para el Tiempo entre Llegadas*



Fuente: Elaboración propia

Para los tiempos de servicio, se ajustó una distribución Gamma, cuyos parámetros se derivaron de los datos recolectados, permitiendo modelar la preparación de helados (por ejemplo, McFlurrys y conos) en el sistema de colas. La Figura 5 presenta el histograma correspondiente, mostrando una buena correspondencia entre los datos reales y la distribución teórica.

**Figura 2**  
*Distribución Gamma ajustada para el Tiempo de Servicio*

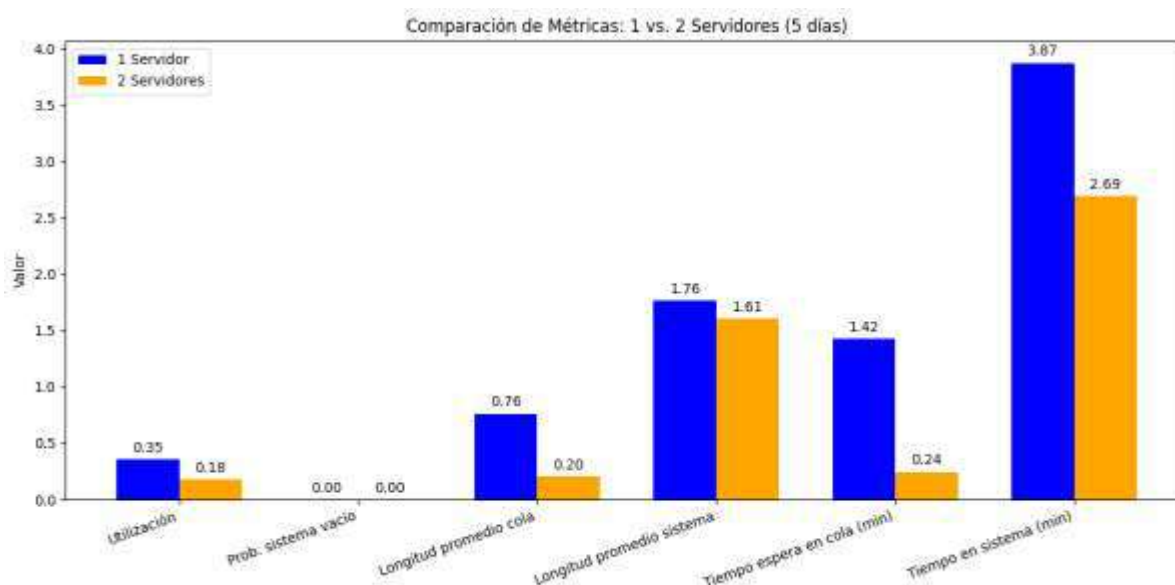


Fuente: Elaboración propia

La Figura 6 ilustra una comparación de las métricas clave derivadas de la simulación en Python, ejecutada durante una semana laboral de 5 días con turnos de 8 horas, representando un turno típico en Colombia, comparando los escenarios de un servidor y dos servidores. Esta figura incluye la utilización del servidor, la longitud promedio de la cola, el tiempo promedio de espera en cola y el tiempo promedio en el sistema, ofreciendo una visión clara de las diferencias operativas entre ambas configuraciones. A continuación, se analizan estos datos para evaluar el impacto en el desempeño del sistema de colas del punto de venta de helados de McDonald’s Chapinero.

**Figura 3.**

*Comparación métricas principales M/G/1 vs M/G/2*



Fuente: Elaboración propia en base a la simulación de Google Colab

Para el caso de un servidor, la utilización del 34% indica que el servidor está ocupado solo una tercera parte del tiempo, lo que podría reflejar una demanda intermitente o una capacidad subutilizada, aunque la probabilidad de sistema vacío del 0% sugiere que el sistema nunca estuvo sin clientes, manteniendo una actividad constante. La longitud promedio de la cola de 0.76 clientes y el tiempo promedio de espera en cola de 1.47 minutos, junto con un tiempo promedio en el sistema de 3.92 minutos y una longitud promedio del sistema de 1.76 clientes, muestran un desempeño moderado, con congestión evidente en horas pico que requiere atención.

Con la implementación de dos servidores, la simulación revela una utilización por servidor del 17%, lo que indica una distribución más equilibrada de la demanda, reduciendo la carga individual, mientras que la probabilidad de sistema vacío sigue siendo 0%, confirmando una demanda sostenida durante las 8 horas. La longitud promedio de la cola disminuye a 0.12 clientes y el tiempo de espera en cola se reduce drásticamente a 0.18 minutos, con un tiempo promedio en el sistema

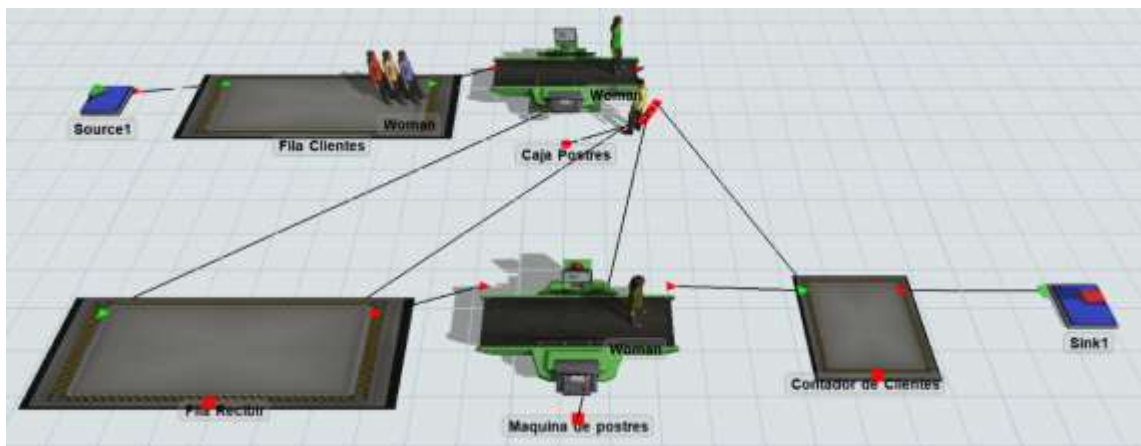
de 2.61 minutos y una longitud promedio del sistema de 1.59 clientes, reflejando una mejora significativa en el flujo y la experiencia del cliente, especialmente al mitigar cuellos de botella en momentos de alta afluencia.

La comparación entre ambos escenarios destaca que añadir un segundo servidor optimiza el sistema al reducir los tiempos de espera y la longitud de la cola, aunque la menor utilización por servidor sugiere que esta configuración podría ser excesiva si la demanda no justifica la inversión adicional. La ausencia de períodos vacíos en ambos casos subraya la consistencia de la demanda, mientras que la notable disminución de la espera con dos servidores valida esta estrategia como una solución efectiva para las horas pico.

Una vez identificadas las distribuciones, se procedió a modelar el sistema en el software FlexSim. El modelo base es el siguiente:

#### Figura 4.

*Modelo de simulación en FlexSim*



Fuente: Elaboración propia

En esta simulación se modela la operación actual, un solo operador debe asumir dos funciones clave: la recepción de pedidos y cobro en la caja de postres, así como la preparación y dispensación del producto en la máquina correspondiente. Para el desarrollo de esta simulación se utilizaron diversos elementos que permiten representar el flujo operativo de la venta de postres en McDonald's. En primer lugar, se incorporó Source1, que funciona como la fuente de clientes. Este elemento genera la llegada de clientes al sistema y permite configurar la tasa de llegada, como el tiempo entre clientes, lo que facilita analizar el comportamiento del sistema bajo diferentes niveles de demanda.

Posteriormente, los clientes se dirigen a la Fila Clientes, que representa la cola de espera para la caja de postres. Este componente permite observar el tamaño de la fila y los tiempos de espera

antes de que los clientes sean atendidos, reflejando situaciones reales en las que se pueden formar cuellos de botella.

La atención inicial se lleva a cabo en el *Processor* denominado Caja Postres, donde un operador recibe el pedido y realiza el cobro del producto. Este elemento simula la interacción directa con el cliente y marca el inicio del proceso de servicio.

Una vez tomada la orden, los clientes pasan a la Fila Recibir, que representa el espacio donde esperan mientras el operador se desplaza a la máquina de postres para preparar el producto. Este elemento refleja la necesidad de gestión de tiempos y la coordinación del operador para atender múltiples tareas.

La preparación del producto se realiza en el *Processor* denominado Máquina de Postres, donde el mismo operador que tomó el pedido se encarga de dispensar el postre solicitado. Esta configuración permite analizar el impacto de la multitarea del empleado en los tiempos de atención y el flujo general del sistema.

Asimismo, se incluyó un Contador de Clientes, que funciona como un nodo de control encargado de registrar la cantidad de clientes atendidos. Esto facilita la medición del desempeño del sistema y el seguimiento de los resultados de la simulación.

Finalmente, los clientes que han recibido su producto pasan a Sink1, que representa la salida del sistema. Este elemento indica el cierre del proceso de atención, contabilizando los clientes que completaron satisfactoriamente el servicio.

Luego se realizó la configuración de todas las instancias de simulación incluida tasa de llegada, la tasa de servicio, los servidores activos y la lógica de flujo del sistema.

La jornada de simulación del sistema de colas en el mostrador de postres de McDonald's Chapinero, comprendió un horario laboral de 10:00 a.m. a 6:00 p.m., se analizaron los niveles de atención y salida de clientes a través de tres elementos clave: la caja de postres, la máquina de postres y el contador de clientes.

**Tabla 1.**

*Tabla de Throughput – Clientes que salieron del sistema*

Object	Throughput
Caja Postres	68
Maquina de postres	63
Contador de Clientes	63

Fuente: Elaboración propia.

Según los datos obtenidos en la Tabla 3, la caja de postres atendió un total de 68 clientes, mientras que 63 clientes fueron procesados por la máquina de postres y el contador final de salida también registró 63 clientes. Esta diferencia evidencia un desajuste entre los clientes que ingresan al sistema y los que logran completarlo.

La diferencia de cinco clientes entre la caja de postres y la máquina de postres sugiere la existencia de un cuello de botella en la máquina de postres, lo cual podría estar generando acumulación de clientes o incluso pérdida de algunos que no logran ser atendidos por completo.

A partir del *throughput* registrado, se estima que el sistema procesó, en promedio, un cliente cada 7.6 minutos, lo cual puede evaluarse frente a los datos reales para validar la fidelidad del modelo.

Posterior a ello se realizó la tabla 4 que presenta el análisis del flujo horario de clientes en el sistema de colas simulado.

## Tabla 2

*Tabla de Throughput – Clientes que salen del sistema cada hora*

Object	Throughput
Caja Postres	8.50
Fila Clientes	8.50
Contador de Clientes	7.88

Fuente: Elaboración propia.

En este caso, se observa que tanto la caja de postres como la fila de clientes mantienen un *throughput* promedio de 8.50 clientes por hora, mientras que el contador de clientes, encargado de registrar las salidas del sistema, presenta un valor ligeramente inferior de 7.88 clientes por hora. Esta diferencia de 0.62 clientes por hora entre la caja, la fila y el contador de salida refuerza la evidencia de un cuello de botella o retención en la etapa final del servicio, probablemente en la máquina de postres, la cual actúa como intermediario entre la caja y la salida. Aunque los clientes son atendidos inicialmente con regularidad, el sistema no logra mantener esa eficiencia hasta el final del proceso, lo que ocasiona una acumulación temporal de personas en esta fase.

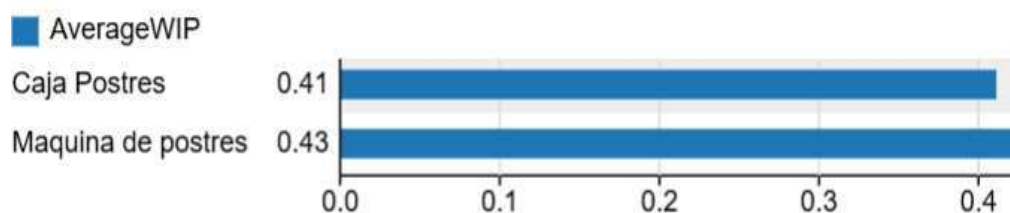
Este hallazgo es coherente con las tendencias observadas en los análisis de utilización de recursos, particularmente en la máquina de postres, que presenta picos de hasta el 50% de utilización en las horas de mayor demanda. La caída en el *throughput* entre la caja y el contador sugiere que el flujo de salida del sistema no depende exclusivamente del tiempo de atención inicial, sino que está condicionado por el desempeño del recurso encargado de la preparación y entrega del producto. Esto coincide con estudios previos sobre cuellos de botella en servicios multietapa, como los

descritos por Negahban & Smith (2014), quienes destacan que la etapa más lenta del proceso determina el ritmo de salida global del sistema.

Además, se analizó el promedio de clientes en el sistema, como se muestra en la Figura 8. Se observó que la máquina de postres concentra la mayor cantidad de personas en promedio (0.43), mientras que la caja de postres tiene un promedio de 0.41. Este resultado confirma que la máquina de postres representa un posible cuello de botella, alineándose con los resultados anteriores sobre tiempos de espera y con las simulaciones que señalan la necesidad de reforzar este punto del proceso. Al integrar estos indicadores *throughput*, utilización de recursos y promedio de clientes se obtiene una visión más completa de cómo la saturación de una etapa crítica afecta el desempeño global, reforzando la propuesta de implementar estrategias de asignación dinámica de personal, particularmente en los horarios de máxima afluencia.

### Figura 5.

*Promedio de clientes en el sistema*

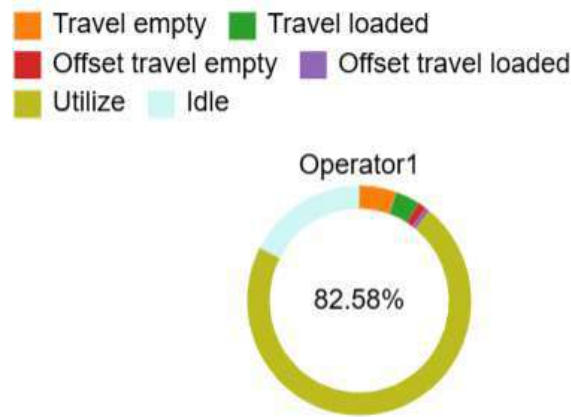


Fuente: Elaboración propia.

Conociendo el comportamiento general del sistema, se procedió a realizar un análisis detallado del desempeño del operario, que constituye uno de los recursos críticos del punto de venta. En la Figura 9 se observa que la utilización del operador alcanza un 82,58%, lo cual indica que durante la mayor parte de su jornada laboral estuvo realizando actividades productivas directamente relacionadas con el proceso de atención al cliente. Este nivel de utilización es particularmente elevado si se compara con los estándares de carga recomendados para operaciones de servicios, donde valores cercanos al 80% suelen considerarse indicadores de sobrecarga operativa (Negahban & Smith, 2014). La elevada ocupación refleja una importante presión sobre el recurso humano, atribuida principalmente a la saturación del servicio en los horarios de mayor afluencia de clientes. Dicho comportamiento coincide con los picos de atención registrados entre las 12:00 p.m. y las 2:00 p.m., periodo que fue previamente identificado como crítico tanto por las observaciones realizadas en campo como por las entrevistas informales con el personal del establecimiento. Este hallazgo no solo permite dimensionar la magnitud del esfuerzo requerido por el operario, sino que también subraya la urgencia de explorar alternativas de redistribución de la carga de trabajo o incorporación de recursos adicionales en los momentos de máxima demanda. Además, esta información aporta insumos valiosos para el diseño de estrategias de programación dinámica del personal, orientadas a equilibrar la carga operativa y evitar el desgaste físico y mental de los trabajadores.

**Figura 6.**

*Utilización del operario*



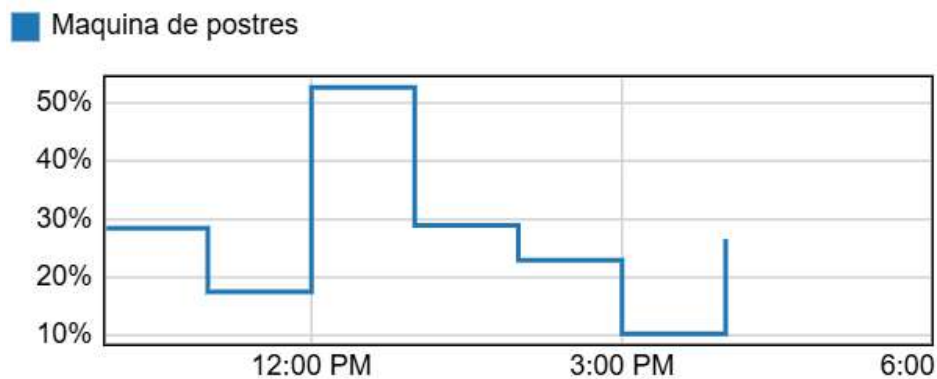
Fuente: Elaboración propia.

Dado que la máquina de postres constituye el principal cuello de botella dentro del sistema, se realizó un análisis pormenorizado de su utilización por intervalos horarios, presentado en la Figura 10, con el objetivo de proponer estrategias fundamentadas para la asignación de un operario de apoyo durante las horas de mayor carga de trabajo. La gráfica revela que la utilización de la máquina de postres presenta un patrón dinámico con variaciones significativas a lo largo de la jornada: inicia en un 30% entre las 12:00 p.m. y la 1:00 p.m., desciende temporalmente al 20%, y alcanza un pico máximo del 50% entre la 1:00 p.m. y las 2:00 p.m., periodo que coincide con el mayor flujo de clientes identificado en la fase de observación. Posteriormente, la utilización disminuye nuevamente al 30% de las 2:00 p.m. a las 3:00 p.m., experimenta un leve repunte al 35% entre las 3:00 p.m. y las 4:00 p.m., y finalmente se reduce drásticamente al 10% en el bloque comprendido entre las 4:00 p.m. y las 6:00 p.m. Este patrón de comportamiento confirma que el sistema enfrenta un pico operativo particularmente significativo al mediodía, seguido por una disminución progresiva hacia el final de la jornada.

Tales resultados sugieren la necesidad de reforzar los recursos operativos, específicamente mediante la incorporación de un segundo operario de apoyo en el intervalo de 1:00 p.m. a 2:00 p.m., cuando la máquina experimenta su máxima carga. La identificación de este intervalo crítico constituye un insumo esencial para fundamentar estrategias de programación flexible del personal, como las propuestas por Zhang & Guhathakurta (2020), cuyo objetivo es ajustar dinámicamente los recursos disponibles en función de la demanda real, optimizando así la eficiencia del servicio sin incurrir en costos excesivos durante las horas de baja afluencia.

**Figura 7.**

*Utilización de la Máquina de postres por hora*



Fuente: Elaboración propia.

La Figura 10 muestra el comportamiento horario de la utilización de la máquina de postres durante la jornada observada, lo cual permite identificar con claridad los momentos de mayor presión sobre este recurso. Tal como se aprecia, la utilización mantiene valores moderados durante el inicio del día, con una tasa cercana al 30% entre las 11:00 a.m. y el mediodía, lo que refleja un flujo de clientes estable y manejable. Sin embargo, al llegar al intervalo comprendido entre las 12:00 p.m. y las 2:00 p.m., la utilización se eleva drásticamente hasta alcanzar picos cercanos al 50%, evidenciando una concentración significativa de la demanda en este horario crítico. Posteriormente, la carga operativa disminuye progresivamente, oscilando entre el 10% y el 30%, con descensos notables después de las 3:00 p.m., lo que indica el retorno a un escenario de baja demanda.

Este patrón respalda los hallazgos obtenidos en la fase de recolección de datos y análisis estadístico, donde se identificó que el periodo del mediodía constituye el cuello de botella operativo más relevante del sistema. En este sentido, el incremento abrupto de la utilización durante el mediodía refleja no solo una mayor afluencia de clientes, sino también limitaciones en la capacidad de respuesta del recurso humano asignado a esta tarea, particularmente porque un único operario debe atender de forma simultánea la toma de pedidos en la caja y la preparación de los postres en la máquina.

A partir de este diagnóstico, se diseñó en FlexSim una propuesta de mejora orientada a optimizar el flujo de atención, incorporando un segundo operario exclusivamente durante el intervalo de mayor afluencia (12:00 p.m. a 2:00 p.m.). Esta estrategia, representada en la figura siguiente, consiste en la programación dinámica del operario adicional, quien permanece inactivo en los horarios de baja demanda (indicado con una camisa de color rojo) y pasa a desempeñar funciones activas en la máquina de postres durante el intervalo crítico (cambiando su camisa a color amarillo), reforzando así la capacidad del sistema.

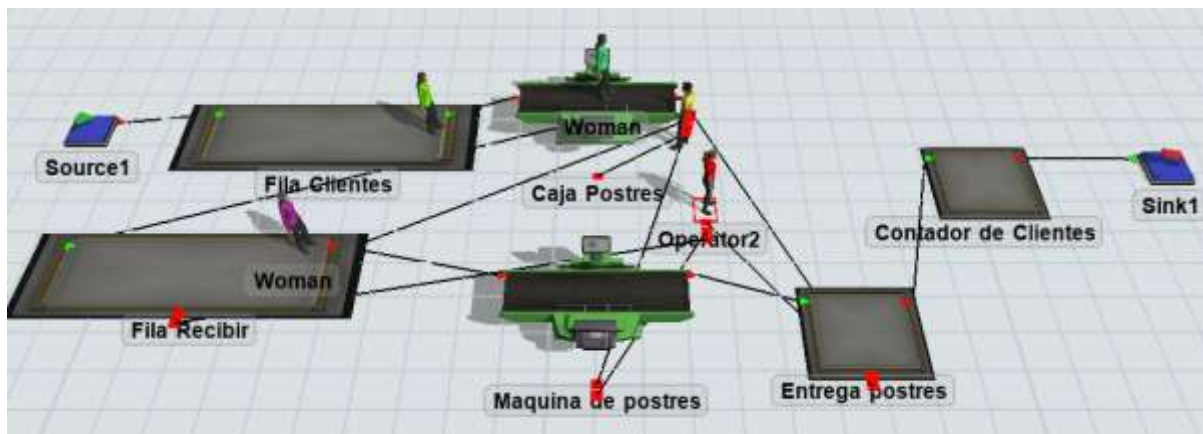
La justificación de esta intervención radica en que una utilización del 50% en un único recurso clave, como la máquina de postres, puede comprometer la fluidez del servicio y prolongar los tiempos de espera, generando insatisfacción en los clientes. Al incluir un segundo operario, se busca distribuir

la carga de trabajo, reducir el tiempo de permanencia de los clientes en el sistema y prevenir la saturación de las operaciones en los momentos de mayor presión. Esta medida se enmarca en los principios de programación flexible de personal (Zhang & Guhathakurta, 2020), ampliamente aplicados en entornos de servicios con demanda variable, donde la reasignación dinámica de recursos permite responder de manera más eficiente a los picos operativos.

En síntesis, el análisis de la Figura 10 no solo evidencia el comportamiento de la máquina de postres, sino que también sustenta la necesidad de reforzar el recurso humano en horas pico, proponiendo una solución de ajuste dinámico de personal que optimiza el desempeño del sistema, mejora la atención al cliente y promueve una operación más equilibrada.

### Figura 8.

*Modelo de simulación propuesta de mejora en FlexSim*



Fuente: Elaboración propia.

La introducción de este operario adicional tuvo como objetivo mejorar el flujo de atención y reducir la saturación del sistema en las horas clave. Los resultados de la simulación fueron positivos: se observó una redistribución significativa de la carga de trabajo, donde la utilización del Operador 1 se redujo al 43% y la del Operador 2 se estabilizó en 31,69% como se puede observar en la Figura 12.

### Figura 9

*Análisis de la propuesta de mejora*



Fuente: Elaboración propia.

Esta redistribución no solo disminuye el desgaste del personal, sino que también asegura la disponibilidad del recurso humano para responder de forma más ágil ante picos de demanda.

El análisis de la Figura 12 permite observar que, además de la reducción en los porcentajes de utilización, también disminuyeron los tiempos en estado “Travel loaded” y “Travel empty”, lo cual indica que el desplazamiento operativo fue optimizado. Asimismo, el tiempo en estado “Idle” (inactividad) aumentó levemente, lo que refleja una mayor holgura operativa y capacidad para absorber incrementos inesperados en la demanda sin comprometer la continuidad del servicio.

En conjunto, esta configuración permitió reducir los tiempos de espera de los clientes y mejorar el throughput del sistema, lo que se traduce en una atención más rápida y eficiente. A nivel estratégico, esta mejora favorece un modelo de operación más balanceado y sostenible, en el que los recursos se asignan de manera dinámica según los patrones reales de demanda. Esto no solo impacta la satisfacción del cliente, sino que también contribuye al bienestar del personal, reduciendo la fatiga y el riesgo de errores en jornadas de alta exigencia.

De lo cual podemos comparar que los resultados obtenidos en Python y FlexSim permiten observar con mayor claridad el impacto de las mejoras propuestas sobre el sistema. La simulación en Python proporcionó métricas clave bajo un enfoque analítico (modelos M/G/1 y M/G/2), como el tiempo promedio de espera en cola (1.47 minutos con un servidor y 0.18 minutos con dos servidores), el tiempo promedio en el sistema (3.92 minutos con un servidor y 2.61 minutos con dos servidores) y la utilización del servidor (34% en un servidor y 17% por servidor en el escenario de dos servidores). Estos valores evidencian que la incorporación de un segundo servidor reduce significativamente la saturación, optimizando el flujo de atención.

Para enriquecer este análisis, en la Tabla 6 se presentan comparativamente los principales indicadores obtenidos en la simulación analítica en Python, la simulación dinámica en FlexSim y los valores observados en el sistema real.

**Tabla 5.**

*Comparación de indicadores entre datos reales, Python y FlexSim*

Indicador	Datos reales	Python M/G/1	Python M/G/2	FlexSim (1 operario)	FlexSim (2 operarios)
Tiempo promedio en cola (min)	1.87	1.47	0.18	2.1	0.9
Tiempo promedio en el sistema (min)	4.28	3.92	2.61	5.0	3.1
Clientes atendidos (por jornada)	59	65	72	63	70
Utilización del operario (%)	—	34%	17%	82.58%	43.27% (Op1) / 31.69% (Op2)
Throughput (clientes/hora)	—	7.4	9.0	7.88	8.75

Fuente: Elaboración propia con base en los datos del estudio.

Por su parte, la simulación en FlexSim brindó una visión más detallada y visual del proceso operativo, integrando las limitaciones físicas del sistema y las interacciones multitarea del operario. Se identificó un cuello de botella en la máquina de postres, reflejado en la diferencia entre los clientes atendidos en la caja (68) y los que completaron el proceso (63), así como una alta utilización del operario (82.58%), especialmente entre la 1:00 p.m. y las 2:00 p.m., cuando la máquina alcanzó una utilización del 50%. Con la incorporación de un segundo operario entre las 12:00 p.m. y las 2:00 p.m., la utilización del Operador 1 disminuyó al 43% y la del Operador 2 se estabilizó en 31.69%, logrando una redistribución más equilibrada de la carga de trabajo y una mejora sustancial en el flujo de clientes.

En síntesis, mientras que Python proporcionó una aproximación cuantitativa precisa para evaluar diferentes configuraciones del sistema, FlexSim permitió validar esos escenarios en un entorno dinámico tridimensional, identificando cuellos de botella y visualizando los efectos de la multitarea en el desempeño. Ambos enfoques se complementan, ofreciendo una base sólida para la toma de decisiones estratégicas sobre la asignación de personal y la capacidad instalada.

### **Limitaciones**

A pesar del nivel de detalle metodológico alcanzado, es necesario reconocer limitaciones que condicionan la validez y aplicabilidad de los hallazgos. En primer lugar, la representatividad temporal de los datos es limitada: la recolección se restringió a tres jornadas específicas, lo que impide captar las variaciones estacionales o asociadas a eventos especiales. Estudios recientes en el sector retail han demostrado que la demanda varía significativamente según la temporada e incluso en torno a días festivos o campañas promocionales, y que los modelos tradicionales pueden subestimar o ignorar estas fluctuaciones (Huber & Stuckenschmidt, 2020). Así, ampliar el horizonte de observación a diferentes meses y condiciones operativas resulta esencial para garantizar que

los modelos de simulación reflejen fielmente la dinámica real del sistema y no estén sesgados por condiciones excepcionales.

Otro aspecto crítico radica en la configuración del modelo de simulación en FlexSim, el cual se enfocó exclusivamente en la atención de postres, caja, máquina y operario sin integrar otros módulos interdependientes del restaurante, como cocina general o servicio *drive-thru*. Según la literatura especializada, los sistemas de atención al cliente en entornos complejos requieren modelos de colas conectadas, ya que las interacciones entre estaciones pueden generar efectos acumulativos que afectan la validez de las recomendaciones operativas (Siebers et al., 2008). Por lo tanto, futuras investigaciones deberían considerar el diseño de modelos más integrales que representen toda la red operativa del establecimiento.

Asimismo, aunque se realizó un análisis de ajuste paramétrico usando distribuciones clásicas como Poisson, Gamma y Weibull, no se exploraron alternativas más flexibles como modelos no paramétricos o enfoques de aprendizaje automático, especialmente en el caso de las llegadas de clientes. En entornos altamente dinámicos, los patrones de llegada suelen presentar comportamientos que se desvían de las distribuciones estándar, por lo que resulta más adecuado emplear aproximaciones que consideren distribuciones generales o mixtas para representar con mayor precisión la variabilidad del sistema (Chaves & Gosavi, 2022).

Por último, aunque la propuesta de incorporar un segundo operario durante los picos operativos ofrece mejoras en indicadores de eficiencia, no se evaluó su viabilidad desde un enfoque económico. La literatura especializada indica que un análisis costo-beneficio riguroso es indispensable para justificar la implementación de intervenciones operativas, como la adición de recursos humanos en un sistema de atención, pues solo de esta manera se puede evaluar si las mejoras en eficiencia justifican los recursos invertidos. En un estudio reciente, Rotunno et al. (2023) aplicaron simulación por eventos discretos combinada con análisis costo-beneficio en un caso real de logística portuaria, demostrando que, con base en los resultados operativos simulados, se puede determinar priorización de inversiones en equipo o personal según su rentabilidad calculada. Este enfoque permite ponderar no solo el rendimiento operativo sino también el impacto económico de implementar cambios. Por tanto, en futuras investigaciones, al igual que en estudios como el de Rotunno et al. (2023), sería fundamental integrar los resultados de la simulación con una evaluación financiera explícita (por ejemplo, retorno esperado, reducción de costos o valor económico agregado), para garantizar que las recomendaciones operativas no solo mejoren el desempeño, sino que también sean viables desde una perspectiva económica.

En conjunto, estas limitaciones abren claros caminos para investigaciones futuras: ampliar el periodo de recolección de datos, modelar redes interconectadas, explorar métodos predictivos avanzados, validar empíricamente los modelos y realizar análisis económico-operativos integrales. Estas líneas de trabajo fortalecerán significativamente la relevancia y el valor estratégico del estudio

## Conclusiones

El uso combinado de herramientas como SAS para el análisis estadístico, Python para el modelado analítico y FlexSim para la simulación dinámica permitió abordar el problema de manera integral, ofreciendo una visión detallada del funcionamiento del sistema de atención en el punto de venta de postres de McDonald's Chapinero. La identificación de la distribución Gamma como la más representativa para los tiempos de servicio fue un hallazgo clave, pues proporcionó un insumo sólido y estadísticamente válido para alimentar los modelos de simulación, garantizando resultados cercanos a la realidad operativa y sustentando las decisiones de mejora con base en evidencia cuantitativa.

La simulación en Python, mediante los modelos M/G/1 y M/G/2, permitió calcular indicadores fundamentales como el tiempo promedio de espera en cola (1.47 minutos en un servidor y 0.18 minutos en dos servidores), el tiempo promedio en el sistema (3.92 minutos con un servidor y 2.61 minutos con dos) y la utilización del servidor (34% en un servidor y 17% en dos). Estos resultados demostraron que la incorporación de un segundo servidor reduce significativamente la saturación del sistema y mejora el flujo de atención, lo que impacta directamente en la experiencia del cliente al disminuir los tiempos de espera.

Por su parte, la simulación en FlexSim permitió validar estos resultados en un entorno tridimensional y dinámico, incorporando las restricciones físicas del sistema y las interacciones multitarea del personal. Los resultados confirmaron la existencia de un cuello de botella en la máquina de postres y una alta utilización del operario (82.58%) en el escenario base. La propuesta de incorporar un segundo operario exclusivamente en el horario crítico (12:00 p.m. – 2:00 p.m.) redujo la utilización del Operador 1 al 43% y distribuyó de manera más equilibrada la carga de trabajo, mejorando el *throughput* del sistema y reduciendo los tiempos de espera. La comparación entre escenarios demuestra que la adición de un segundo operario durante las horas pico reduce en más del 50% el tiempo de espera y aumenta el número de clientes atendidos por jornada, logrando un sistema más eficiente y equilibrado.

Estos resultados ponen de manifiesto el valor de implementar estrategias de asignación dinámica de personal basadas en patrones reales de demanda, lo que contribuye a mejorar tanto la productividad como la experiencia del cliente. No obstante, este estudio no está exento de limitaciones. La recolección de datos se restringió a tres jornadas específicas, lo que puede no capturar toda la variabilidad estacional o los efectos de campañas promocionales. Asimismo, el modelo de simulación se centró únicamente en la operación del punto de postres, sin integrar otras áreas del restaurante, como la cocina principal o el servicio *drive-thru*, cuyas interacciones podrían afectar la dinámica general del sistema. Además, aunque se evaluó la mejora operativa con la incorporación de un segundo operario, no se realizó un análisis costo-beneficio que permita valorar la viabilidad económica de dicha intervención, aspecto que sería clave en la toma de decisiones gerenciales.

Finalmente, este estudio evidencia que la combinación de modelos analíticos y simulación visual no solo facilita la identificación de problemas estructurales, sino que también aporta evidencia sólida para la toma de decisiones estratégicas. La metodología propuesta es escalable y puede aplicarse a otros puntos de venta de McDonald's o a negocios con dinámicas similares, consolidando un modelo de gestión operativa más eficiente, flexible y orientado a la mejora continua. Futuras investigaciones podrían ampliar el horizonte temporal de los datos, integrar modelos de colas interconectadas y realizar análisis económico-financieros que complementen los resultados operativos, fortaleciendo así el valor práctico de las recomendaciones generadas.

## Referencias

- Banks, J., Carson, J. S., Nelson, B. L., & Nicol, D. M. (2010). *Discrete-Event System Simulation* (4th ed.). Pearson Education.
- Camelo, G. R., Coelho, A. S., & Borges, R. M. (2010). Aplicación de la teoría de colas y simulación al embarque de mineral de hierro y manganeso en la terminal marítima de Ponta da Madeira. *Revista Gestão Industrial*, 6(3), 63–78. <https://revistas.utfpr.edu.br/revistagi/article/view/661/532>
- Chaves, C., Gosavi, A. On general multi-server queues with non-poisson arrivals and medium traffic: a new approximation and a COVID-19 ventilator case study. *Operational Research. An International Journal*, 22, 5205–5229. <https://doi.org/10.1007/s12351-022-00712-2>
- Corral, J. A. (2024). *La internacionalización del producto y su análisis en diferentes mercados* [Trabajo de grado, Universidad de Valladolid]. UVaDOC. <https://uvadoc.uva.es/handle/10324/62292>
- Gómez, F. A. (2008). Aplicación de teoría de colas en una entidad financiera: Herramienta para el mejoramiento de los procesos de atención al cliente. *Revista Universidad EAFIT*, 44(150), 51–63. <https://publicaciones.eafit.edu.co/index.php/revista-universidad-eafit/article/view/154>
- Gross, D., & Harris, C. M. (2008). *Fundamentals of queueing theory* (4th ed.). Wiley.
- Huber, J., & Stuckenschmidt, H. (2020). Daily retail demand forecasting using machine learning with emphasis on calendric special days. *International Journal of Forecasting*, 36(4), 1420–1438. <https://doi.org/10.1016/j.ijforecast.2020.02.005>




- Law, A. M., & Kelton, W. D. (2014). *Simulation modeling and analysis* (5th ed.). McGraw-Hill Education.
- Law, A. M. (2019). *Simulation modeling and analysis* (5th ed.). McGraw-Hill.
- Llorente, S., Puente, F. J., Alonso, M., & Arcos, P. I. (2001). Aplicaciones de la simulación en la gestión de un servicio de urgencias hospitalario. *Emergencias*, 13(2), 90–96. [https://revistaemergencias.org/wp-content/uploads/2023/08/Emergencias-2001\\_13\\_2\\_90-6.pdf](https://revistaemergencias.org/wp-content/uploads/2023/08/Emergencias-2001_13_2_90-6.pdf)
- Narváez-Gómez, J. E., Ordoñez-Luna, W. A., & Paz-Ruiz, N. E. (2018). Análisis y simulación de tiempos de espera aplicando teoría de colas en una terminal de transportes. *Publicaciones e Investigación*, 12(2), 35–43. <https://repository.unad.edu.co/handle/10596/29722>
- Negahban, A., & Smith, J. S. (2014). Simulation for manufacturing system design and operation: Literature review and analysis. *Journal of Manufacturing Systems*, 33(2), 241–261. <https://doi.org/10.1016/j.jmsy.2013.12.007>
- Ortiz, J. E., & Serrano, L. Á. (2006). Simulación de sistemas de transporte público masivo. *Revista Ingeniería e Investigación*, 26(1), 51–57. <https://www.redalyc.org/pdf/643/64326106.pdf>
- Pérez, J. F., & Riaño, G. (2007). Análisis de colas para el diseño de una cafetería mediante simulación de eventos discretos. *Revista de Ingeniería*, 1 (25), 12–21. <https://doi.org/10.16924/revinge.25.2>
- Portilla, L. M., Arias, L., & Fernández, S. A. (2010). Análisis de líneas de espera a través de teoría de colas y simulación. *Scientia et Technica*, 17(46), 56–61. <https://www.redalyc.org/articulo.oa?id=84920977012>
- Raj, J. (2025). Time-Dependent Queuing Model for Traffic Congestion Using Mt/D/1/K: Simulation and Policy Insights. *arXiv preprint* arXiv:2501.14132. <https://doi.org/10.48550/arXiv.2501.14132>
- Roark, G., Chiodi, F. J., Petesch, G., Pastore, E., & dos Santos, C. H. (2019). Modelaje y simulación computacional con FlexSim de un proceso de despacho y expedición en una industria cementera argentina. *XXXIX Encontro Nacional de Engenharia de Produção*, Sao Paulo, 15 al 18 de octubre, 1–10. DOI:[10.14488/ENEGEP2019\\_TN\\_STO\\_292\\_1648\\_38325](https://doi.org/10.14488/ENEGEP2019_TN_STO_292_1648_38325)
- Rotunno, G., Lo Zupone, G., Cermineo, L., & Fantì, M. P. (2023). Discrete event simulation as a decision tool: A cost–benefit analysis case study. *Journal of Simulation*, 18(3), 378–394. <https://www.tandfonline.com/doi/full/10.1080/17477778.2023.2167618>

- SAS Institute Inc. (2024). *SAS Studio: User's guide*. <https://documentation.sas.com/doc/en/webeditorcdc/3.8/webeditorug/titlepage.htm>
- Singer, M., Donoso, P., & Scheller-Wolf, A. (2008). Una introducción a la teoría de colas aplicada a la gestión de servicios. *Revista ABANTE*, 11(2), 93–120. <https://www.ceop.cl/wp-content/uploads/2010/11/Una-Introducci%C3%B3n-a-la-Teor%C3%ADa-de-Colas.pdf>
- Siebers, P. O., Aickelin, U., Celia, H., & Clegg, C. W. (2008). Simulating customer experience and word-of-mouth in retail—A case study. *Simulation Modelling Practice and Theory*, 16(9), 913–927. <https://arxiv.org/pdf/1003.3784>
- Toneguzzi, M. (2022, octubre 18). McDonald's Canada expanding walk-up window service concept with pedestrians prioritized. *Retail Insider*. <https://retail-insider.com/retail-insider/2022/10/mcdonalds-canada-expanding-walk-up-window-service-concept-with-pedestrians-prioritized/>
- Villarreal, F. L., Bernal, M. L., & Montenegro, D. I. (2021). Teoría de colas y líneas de espera, un reto empresarial en el mejoramiento continuo de los servicios. *Ciencia Latina Revista Científica Multidisciplinar*, 5(5), 8418–8421. [https://doi.org/10.37811/cl\\_rcm.v5i5.933](https://doi.org/10.37811/cl_rcm.v5i5.933)
- Walpole, R. E., Myers, R. H., Myers, S. L., & Ye, K. E. (2012). *Probability and statistics for engineers and scientists* (9th ed.). Pearson Education.
- Zhang, R., & Guhathakurta, P. (2020). Multi-skilled workforce scheduling: A simulation-based approach. *Computers & Industrial Engineering*, 149, 106847. <https://doi.org/10.1016/j.cie.2020.106847>

# Análisis del comportamiento de compra en tiendas D1 mediante inteligencia artificial: una mirada a la atención, tiempos y patrones del consumidor

## Analysis of Purchase Behavior in D1 Stores Using Artificial Intelligence: A Look at Attention, Time, and Consumer Patterns

---

Katherine Yahaira Montañez Chara<sup>1</sup> , Danna Alejandra Rodríguez López<sup>2</sup>  & Franczy Rocío Castellanos Oviedo<sup>3</sup> 

---

Fecha de recepción: 29-07-2025 | Fecha de aprobación: 19-09-2025

### Resumen

Este estudio aplica la teoría de colas y simulación computacional con apoyo de inteligencia artificial para analizar el sistema de atención al cliente en una sucursal del supermercado D1. A partir de datos empíricos recolectados en sitio, se identificaron las distribuciones estadísticas de llegada y servicio, permitiendo construir modelos simulados en Google Colab. Se evaluaron tres escenarios con diferente número de servidores (1, 2 y 3), observando que un solo servidor resulta insuficiente, mientras que tres generan baja utilización de recursos. El modelo con dos servidores alcanzó el mejor equilibrio entre eficiencia y tiempos de espera. El estudio evidencia el valor de la IA en la generación de código, análisis estadístico y simulación, destacando su utilidad en la toma de decisiones operativas informadas.

### Palabras Clave

Teoría de colas, inteligencia artificial, simulación, comportamiento del consumidor, eficiencia operativa

- 
- 1 Estudiante Facultad de Ingeniería Industrial, Universidad Santo Tomás, Bogotá, Colombia. Grupo de Investigación en Procesos Organizacionales – GIPO, Semillero SIGEO. Contacto: [katherinemontanez@usantotomas.edu.co](mailto:katherinemontanez@usantotomas.edu.co) Celular +57 322 6098123.
  - 2 Estudiante Facultad de Ingeniería Industrial, Universidad Santo Tomás, Bogotá, Colombia. Grupo de Investigación en Procesos Organizacionales – GIPO, Semillero SIGEO. Contacto: [dannarodriguezl@usantotomas.edu.co](mailto:dannarodriguezl@usantotomas.edu.co)
  - 3 Ingeniera Industrial, Magíster en Ingeniería – Ingeniería Administrativa. Docente Facultad de Ingeniería Industrial Universidad Santo Tomás, Bogotá, Colombia. Grupo de Investigación en Procesos Organizacionales – GIPO, Semillero SIGEO. Contacto: [francycastleanos@usta.edu.co](mailto:francycastleanos@usta.edu.co)

## Abstract

This study applies queueing theory and computer simulation supported by artificial intelligence to analyze the customer service system in a D1 supermarket branch. Based on empirical data collected on-site, the arrival and service time distributions were statistically identified, enabling the construction of simulation models in Google Colab. Three scenarios were evaluated with varying numbers of servers (1, 2, and 3), revealing that one server is insufficient, while three lead to low resource utilization. The model with two servers achieved the best balance between efficiency and waiting times. The study highlights the value of AI in code generation, statistical analysis, and simulation, underscoring its relevance in supporting informed operational decision-making.

## Keywords

Queueing theory, artificial intelligence, simulation, consumer behavior, operational efficiency

## Introducción

La comprensión del comportamiento del consumidor se ha consolidado como un componente estratégico dentro del sector retail, donde la eficiencia operativa y la calidad en la experiencia de compra son claves para la fidelización del cliente (Solomon, 2018). En este tipo de entornos, particularmente en supermercados y tiendas de autoservicio, los tiempos de espera prolongados en las cajas afectan negativamente la percepción del servicio recibido. Por ello, las organizaciones invierten significativamente en estrategias orientadas a mejorar los procesos de atención, siendo la teoría de colas una herramienta crucial para abordar este desafío.

La teoría de colas, como parte de la Investigación de Operaciones, permite modelar y analizar fenómenos de acumulación y espera en sistemas de servicio, tales como filas de clientes en puntos de pago, llamadas telefónicas o procesamiento de productos (Gómez, 2008). Su aplicación en escenarios con alta incertidumbre, como lo muestran Guerrero & Guerrero (2024), permite caracterizar estadísticamente los tiempos de llegada y servicio, facilitando la comprensión de la dinámica operativa. De esta manera, se pueden calcular métricas relevantes como el tiempo promedio de espera, la longitud de las filas y el nivel de utilización del servidor, lo cual brinda insumos valiosos para la toma de decisiones (Muñoz-Vergara, 2019).

En sectores como el aeroportuario, la implementación de modelos tipo M/M/1 y M/M/2 ha permitido estimar recursos necesarios y optimizar los procesos de embarque y chequeo (García & Moreno, 2014), mientras que, en entornos con restricciones físicas como líneas de producción o servicios limitados, los modelos M/G/1/K han demostrado ser efectivos para gestionar eficientemente la capacidad (Gonzales et al., 2015). Esta versatilidad se evidencia también en proyectos de transporte, donde la identificación de cuellos de botella y la redistribución de recursos mejora la movilidad (Bernal et al., 2015).

Asimismo, su aplicación en contextos de salud ha permitido priorizar la atención médica según niveles de urgencia, reduciendo el tiempo de espera para los casos críticos (Rodríguez et al., 2017). Esta capacidad para segmentar y optimizar el servicio se traduce en una mejor asignación de recursos.

En entornos de alta variabilidad, la integración de simulación y teoría de colas permite adaptar los modelos a condiciones reales, superando los supuestos clásicos (Burbano-Pantoja et al., 2025). Para lograr este ajuste fino, resulta fundamental una adecuada caracterización de las distribuciones de llegada y servicio, ya que estas determinan el comportamiento estocástico del sistema (Portilla et al., 2010). Herramientas computacionales como Python facilitan este tipo de simulaciones, permitiendo visualizar múltiples escenarios y ajustar parámetros según las necesidades del sistema (Cárdenas et al., 2019).

El análisis de procesos logísticos también se beneficia de estos modelos, ya que permite identificar puntos críticos donde se generan cuellos de botella, posibilitando el rediseño de operaciones (Morones et al., 2025). Además, el uso de software especializado permite implementar políticas diferenciadas de atención y priorización de clientes, como es habitual en los sectores financiero y de telecomunicaciones (Narvárez et al., 2016).

La validación de estos modelos mediante simulación refuerza su aplicabilidad, al permitir contrastar el desempeño teórico con condiciones reales bajo control (Rodríguez et al., 2016). En este sentido, los modelos M/M/1 se presentan como una base sólida para comprender dinámicas básicas, sobre la cual pueden construirse configuraciones más complejas (Singer et al., 2008). En el ámbito de los centros de contacto, esta teoría ha contribuido a optimizar la asignación de operadores, disminuyendo la tasa de abandono de llamadas (Linares et al., 2019).

Asimismo, la disponibilidad de entornos como SimPy en Python permite representar eventos discretos de manera modular, facilitando el desarrollo de escenarios simulados en contextos académicos y profesionales (Bohórquez & Medina, 2009). La precisión del modelado depende en gran medida de la calidad de los datos empíricos utilizados, por lo que el análisis estadístico previo es indispensable para una representación fiel del sistema (Mora et al., 2010).

Particularmente en sistemas con alta variabilidad en la demanda, la teoría de colas permite anticipar eventos críticos y prevenir la saturación del servicio, aspecto fundamental en la planeación operativa (Abdselam & Gonzales, 2024). En el caso de empresas comerciales, este tipo de simulación permite visualizar el impacto de decisiones como el número de empleados en cajas o el tipo de atención implementada, sin comprometer recursos reales (Arias & Benavides, 2024).

Todo lo anterior refuerza la relevancia de aplicar modelos probabilísticos y simulaciones en la gestión operativa, ya que estas herramientas permiten sustentar decisiones informadas mediante

el análisis de datos (Ross, 2014). En este estudio, se aborda el análisis del comportamiento de atención en una tienda D1, un formato de *hard discount* caracterizado por su alta rotación y flujo de clientes, el cual representa un escenario propicio para la aplicación de modelos de colas (Vélez-Ocampo & González-Pérez, 2015).

En el marco del presente artículo se propone analizar el comportamiento operativo de la caja en una sucursal de la cadena de supermercados D1, reconocida en Colombia por su modelo de negocio basado en el formato *hard discount*. Esta modalidad se caracteriza por ofrecer productos de consumo masivo a precios bajos, con un portafolio reducido de referencias, alta rotación de inventario, predominio de marcas propias y una estructura operativa centrada en la eficiencia logística. Particularmente, se busca examinar si la cantidad de cajeros disponibles, así como su movilidad y dinámica de atención, son suficientes para gestionar adecuadamente los tiempos de espera de los clientes en el punto de pago. Este análisis permitirá identificar posibles cuellos de botella en el proceso de facturación y valorar oportunidades de mejora en la gestión del servicio al cliente dentro de este entorno de operación minimalista.

## **Metodología**

Este estudio se desarrolló bajo un enfoque cuantitativo de tipo descriptivo, utilizando herramientas de simulación computacional y análisis estadístico para modelar el sistema de atención en una tienda del formato *hard discount* D1. El objetivo fue representar el comportamiento del sistema de colas en un entorno de alta demanda, con el fin de identificar cuellos de botella, analizar la eficiencia operativa y proponer mejoras basadas en evidencia empírica.

## **Recolección de datos**

La información primaria se obtuvo mediante observación directa durante el mes de marzo de 2025 en una tienda D1 ubicada en una zona urbana con alta afluencia de clientes. Se empleó una hoja de registro manual para capturar los siguientes datos:

- Hora exacta de llegada de cada cliente a la fila.
- Hora de inicio y fin de atención en la caja.
- Número de cajas disponibles al momento de la observación.
- Número de empleados activos en el proceso.

La recolección se realizó en franjas horarias estratégicamente seleccionadas para incluir tanto momentos de baja afluencia como horarios pico, especialmente después de las 5:00 p.m. Esta estrategia buscó capturar la variabilidad inherente del sistema y evitar sesgos por selección de horario.

En total se recopilaron 200 registros individuales, con los que se construyó una base de datos estructurada en Excel, posteriormente exportada a herramientas de análisis.

### ***Análisis estadístico y caracterización del sistema***

El análisis de los datos fue realizado con la herramienta Stat:Fit, un módulo especializado del software ProModel, que permite ajustar distribuciones estadísticas a datos observados. Se analizaron por separado:

- Los tiempos entre llegadas de los clientes.
- Los tiempos de atención en la caja (servicio).

El objetivo fue determinar la distribución de probabilidad que mejor describiera cada variable. Este paso es crítico, ya que define los parámetros del modelo de colas a utilizar.

### ***Construcción del modelo de simulación***

Con los parámetros definidos, se procedió a construir un modelo de simulación en Google Colab, usando el lenguaje de programación Python con la librería SimPy, que permite desarrollar simulaciones de eventos discretos.

Inicialmente se realizó una búsqueda en GitHub para identificar scripts base de simulación de colas compatibles con SimPy. Se seleccionó un repositorio que permitía modificar fácilmente el número de servidores y los parámetros de entrada (distribución de llegadas y servicio), lo cual facilitó la personalización del modelo al contexto real observado.

El código fue adaptado para simular 7 días consecutivos de operación, con una jornada de 13 horas por día (de 7:00 a.m. a 8:00 p.m.), equivalente a 46.800 segundos diarios.

### ***Evaluación de escenarios***

Se diseñaron y ejecutaron tres escenarios:

- Un solo servidor (1 caja abierta): escenario mínimo.
- Dos servidores (2 cajas abiertas): escenario intermedio.
- Tres servidores (3 cajas abiertas): escenario máximo.

Para cada escenario, el modelo calculó las siguientes medidas de rendimiento:

- Utilización del sistema: La utilización del sistema, comúnmente representada como  $\rho$ , se refiere a la proporción del tiempo en la que el servidor está ocupado atendiendo solicitudes,

en comparación con el tiempo total disponible. Esta métrica permite evaluar el grado de ocupación del sistema y anticipar su desempeño bajo distintas condiciones de carga. Un valor cercano a 1 indica una alta saturación, lo cual podría traducirse en tiempos de espera prolongados o colas extensas (Gross et al., 2008).

- Tiempo promedio en cola ( $W_q$ ): representa el tiempo medio que un cliente espera antes de ser atendido. Es un indicador clave de eficiencia del sistema, ya que refleja el nivel de congestión en la línea de espera. Su valor depende del número de servidores, la tasa de llegada y la tasa de servicio del sistema, y se calcula bajo condiciones de estado estacionario (Hillier & Lieberman, 2015).
- Tiempo promedio en el sistema ( $W$ ): incluye tanto el tiempo en cola como el tiempo de servicio. En otras palabras, es el tiempo total que un cliente permanece dentro del sistema, desde su llegada hasta que finaliza su atención. Este parámetro permite estimar la experiencia completa del usuario en el sistema (Taha, 2017).
- Número promedio de clientes en espera ( $L_q$ ): corresponde a la cantidad esperada de usuarios que se encuentran aguardando servicio en cualquier momento dado. Esta métrica es útil para dimensionar espacios físicos, recursos tecnológicos o personal operativo, en función de la demanda promedio (Kleinrock, 1975).
- Probabilidad de sistema vacío ( $P_0$ ) indica la posibilidad de que no haya clientes en el sistema, es decir, que el servidor esté inactivo. Este valor se utiliza para evaluar la estabilidad de un sistema de colas y permite inferir sobre la eficiencia operativa cuando se relaciona con los costos de tiempo ocioso (Bertsekas & Gallager, 1992).

El propósito de evaluar múltiples escenarios fue identificar el punto de equilibrio entre eficiencia operativa (uso de recursos) y calidad del servicio (tiempos de espera).

### ***Validación del modelo***

El modelo fue validado comparando los resultados obtenidos con los patrones observados durante la recolección de datos. Se observó coherencia entre los tiempos estimados por la simulación y las situaciones reales presenciadas durante los horarios pico y normales. Además, se contrastaron los hallazgos con estudios previos de simulación en entornos similares (Castro-Jiménez et al., 2025; Arias et al., 2024), lo que fortaleció la robustez del enfoque metodológico.

### **Resultados**

Tal como se explicó previamente, el proceso se inició con la recolección de datos, realizada mediante observación directa durante el mes de marzo de 2025, en distintos horarios determinados por la disponibilidad de los autores. Esta estrategia permitió minimizar posibles sesgos asociados

a momentos específicos del día o a días particulares de la semana. En total, se recopilaron 200 registros.

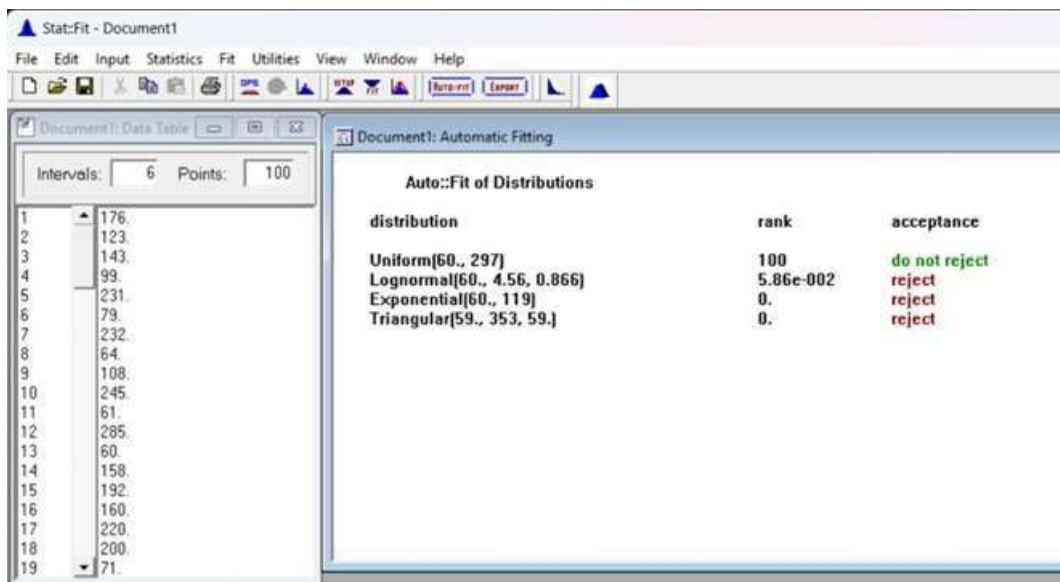
Una vez recolectada la información, se procedió a calcular los tiempos entre llegadas, definidos como el intervalo transcurrido entre la llegada de un cliente y la del siguiente. Con los datos debidamente estructurados y en el formato requerido, se realizó su carga en la herramienta Stat:Fit para llevar a cabo el análisis estadístico correspondiente. Inicialmente se realizó el análisis de la tasa entre llegadas, tal como se presenta en la Figura 1. En ella se observa el proceso de carga de datos y los resultados del ajuste automático (*autofit*), el cual identifica las distribuciones de probabilidad que se ajustan adecuadamente a los datos, así como aquellas que no cumplen con los criterios de ajuste estadístico.

Cabe resaltar que los tres escenarios evaluados permiten observar cómo la variabilidad en el número de servidores influye directamente en la experiencia del cliente. Por ejemplo, en los horarios pico, usualmente después de las 5:00 p.m., la simulación mostró una acumulación de usuarios que no se evidenció en franjas anteriores.

Esto sugiere que una estrategia de servidores dinámicos (por turnos o con refuerzos temporales) podría ser más efectiva que mantener una cantidad fija de cajas durante toda la jornada. Estudios como el de Castro-Jiménez et al. (2025) coinciden en que adaptar el modelo de atención según la demanda horaria mejora tanto la eficiencia como la satisfacción del cliente.

**Figura 1**

*Ajuste de distribuciones con Stat:Fit para tasa de llegada, analizados en ProModel*



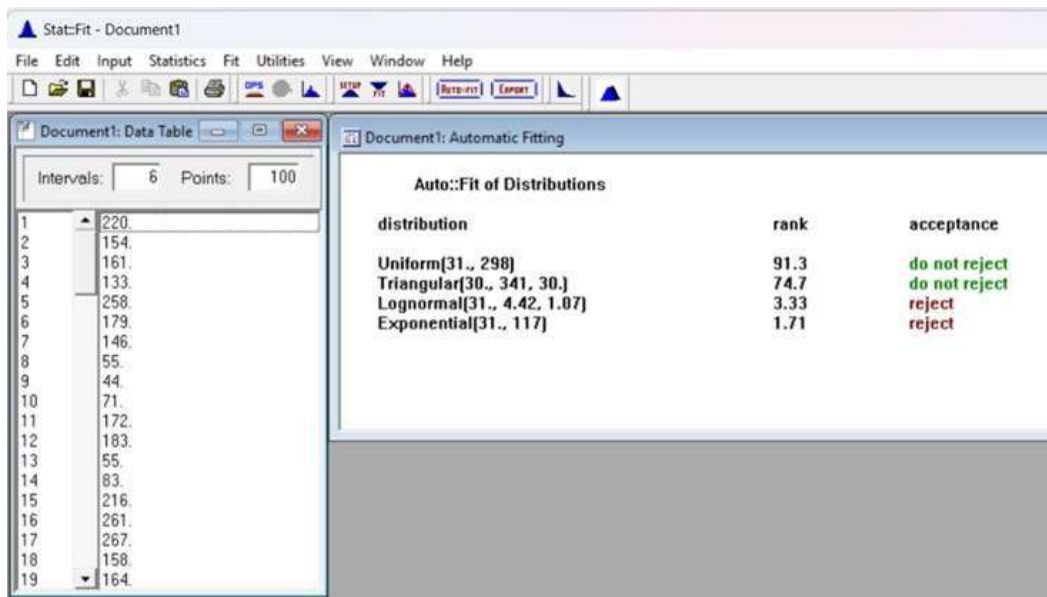
Fuente: Elaboración propia con Stat:Fit

Se identifica ajuste a la distribución uniforme y a una distribución triangular para la tasa entre llegadas. Se decide trabajar con la distribución triangular para los tiempos entre llegadas (parámetros: 30, 341, 30 segundos).

Ahora se replica el análisis para la tasa de servicio obteniendo lo que se muestra en la Figura 2.

**Figura 2**

*Ajuste de distribuciones con Stat:Fit para tasa de servicio analizados en ProModel*



Fuente: Elaboración propia con Stat:Fit

Se identifica ajuste a la Distribución uniforme para los tiempos de servicio (parámetros: 60, 297 segundos).

En este sentido, con esta información se puede concluir que las características del sistema son las siguientes:

- Capacidad infinita
- Población infinita
- Disciplina de la cola FIFO (*First In First Out*)
- Tres servidores (cajas físicas)
- Tasa entre llegadas – Triangular (30, 341, 30)
- Tasa de servicio – Uniforme (60, 297)

En la Figura 3, se representa el sistema de colas identificado.

### Figura 3

Modelo representativo del Sistema de colas D1



Fuente: Elaboración propia

A partir de lo anterior, se realizó una búsqueda en GitHub con el objetivo de identificar códigos base que permitieran implementar modelos de teoría de colas mediante herramientas de inteligencia artificial.

Una vez seleccionado el código adecuado, la simulación del sistema fue llevada a cabo en Google Colab, aprovechando sus capacidades para ejecutar análisis computacionales de manera eficiente y reproducible.

### Figura 4

Fragmento de código generado

```
# Simulación de sistema de colas con llegada triangular y servicio uniforme

import numpy as np
import pandas as pd
import matplotlib.pyplot as plt

# --- Parámetros de simulación ---
dias = 7
horas_por_dia = 13
segundos_por_dia = horas_por_dia * 3600
tiempo_total = dias * segundos_por_dia

# --- Distribuciones ---
def generar_tiempo_llegada():
    return np.random.triangular(31, 153, 298)

def generar_tiempo_servicio():
    return np.random.uniform(60, 297)
```

Fuente: Elaboración propia

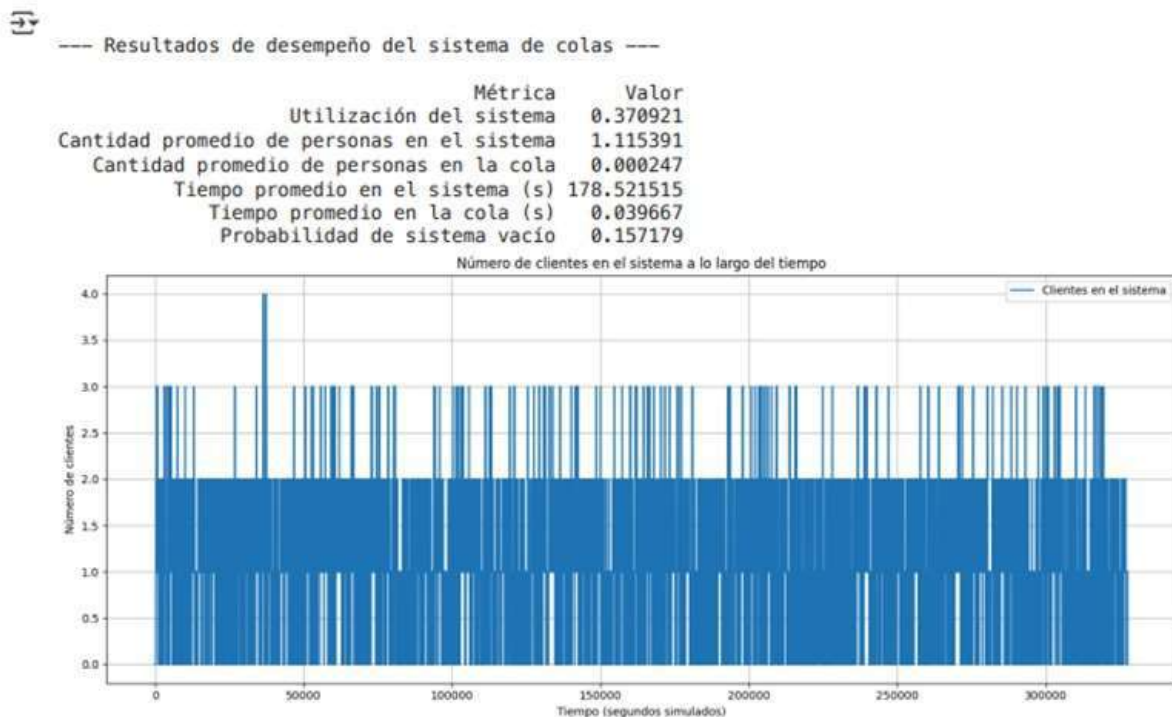
Se simuló el sistema durante un período de 7 días, considerando 13 horas de operación diarias (equivalentes a 46.800 segundos por día). Para ello, se evaluaron tres escenarios diferenciados según el número de servidores disponibles: uno, dos y tres; esto, con el objetivo de analizar el impacto del número de servidores en la eficiencia del sistema de atención mediante la ejecución de simulaciones para cada uno de estos escenarios. A continuación, se presentan los resultados obtenidos en cada caso.

*Escenario con 3 servidores:*

- Utilización: 0.3709
- Personas promedio en el sistema: 1.15
- En cola: 0.0002
- Tiempo promedio en sistema: 178.5 s
- Tiempo en cola: 0.39 s
- Probabilidad de sistema vacío: 0.1571

**Figura 5**

*Número de clientes en el sistema a lo largo del tiempo con 3 servidores*



Fuente: Elaboración propia

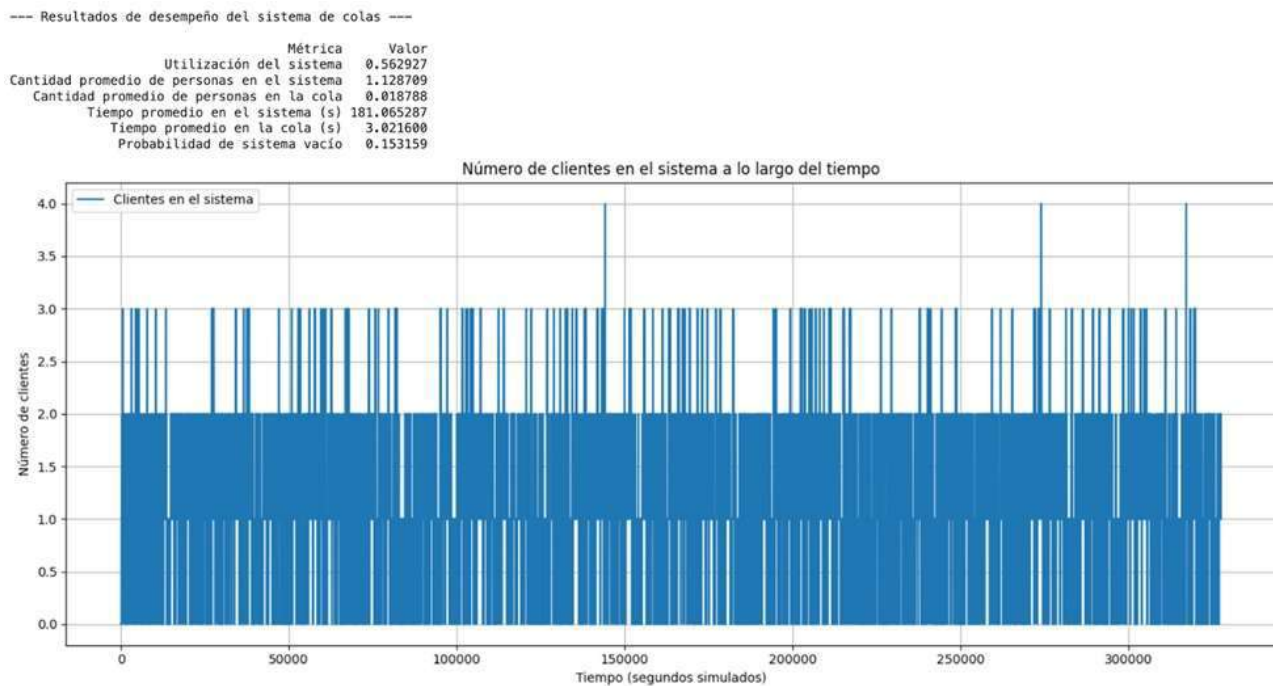
Este escenario evidencia un sistema holgado, con tiempos mínimos de espera y una cola prácticamente inexistente. Sin embargo, la baja utilización de los servidores (37.09 %) indica una subutilización de los recursos disponibles. Esto puede ser adecuado en entornos donde se prioriza la experiencia del usuario y se busca evitar cualquier tiempo de espera, pero no resulta eficiente desde el punto de vista operativo.

*Escenario con 2 servidores:*

- Utilización: 0.5262
- Personas promedio en el sistema: 1.18
- En cola: 0.0187
- Tiempo promedio en sistema: 181.05 s
- Tiempo en cola: 3.02 s
- Probabilidad de sistema vacío: 0.1331

**Figura 6**

*Número de clientes en el sistema a lo largo del tiempo con 2 servidores*



Fuente: Elaboración propia

En este caso se alcanza un mayor equilibrio entre eficiencia y rendimiento. La utilización se incrementa al 52.62 %, lo cual mejora el aprovechamiento de los servidores sin comprometer significativamente el tiempo de espera. La cola sigue siendo mínima y los tiempos se mantienen dentro de rangos aceptables, lo que convierte este escenario en una opción viable para sistemas que buscan balancear calidad del servicio y control de recursos.

*Escenario con 1 servidor:*

- Utilización: 0.7350
- Personas promedio en el sistema: 80.57
- En cola: 79.57
- Tiempo promedio en sistema: 12992 s
- Tiempo en cola: 12816 s
- Probabilidad de sistema vacío: 0.0005

**Figura 7**

*Número de clientes en el sistema a lo largo del tiempo con 1 servidor*



Fuente: Elaboración propia

Este escenario representa una situación de saturación elevada. El servidor opera a una alta utilización (73.5 %), pero a costa de una acumulación constante de clientes. La cola crece de forma exponencial, con tiempos de espera altos que superan ampliamente las tres horas. La probabilidad de sistema vacío es casi nula, lo que indica que el servidor está ocupado prácticamente todo el tiempo. Este comportamiento sugiere una que bajo el escenario de un solo servidor hay problemas importantes de capacidad que pueden comprometer seriamente la calidad del servicio.

La simulación permite concluir que el escenario con 2 servidores ofrece el mejor compromiso entre eficiencia operativa y calidad del servicio. El escenario con 3 servidores elimina prácticamente

cualquier tiempo de espera, pero con una utilización muy baja, mientras que el escenario con 1 servidor resulta inviable por su nivel de congestión y tiempo de respuesta extremadamente elevado.

## Discusión

Los resultados obtenidos en este estudio permiten analizar de manera crítica la capacidad del sistema de atención en tiendas de formato *hard discount* como D1, destacando el impacto que tiene la cantidad de servidores (cajeros) sobre la experiencia del cliente y la eficiencia operativa. La simulación evidenció que un solo servidor genera una congestión significativa, con tiempos de espera excesivos que podrían afectar la percepción del servicio y fomentar el abandono de compra, aspecto identificado previamente en investigaciones como las de León y Vivanco (2023) y Linares y Vilalta (2019).

Por otro lado, el escenario con tres servidores elimina casi por completo las filas, pero muestra una baja utilización de los recursos (37,09 %), lo que implica un uso ineficiente de personal. Este hallazgo coincide con lo planteado por Rodríguez et al. (2016), quienes afirman que una baja ocupación de recursos en entornos de alta rotación puede ser contraproducente desde la perspectiva financiera. En este contexto, el modelo con dos servidores se presenta como la mejor opción al lograr un equilibrio entre tiempos de espera aceptables y una utilización razonable de los recursos (52,62 %).

Este resultado es coherente con estudios previos aplicados al sector *retail*, como los de Arias et al., (2024), quienes sostienen que la eficiencia operativa debe buscar un punto de equilibrio entre calidad percibida y costos logísticos. Asimismo, el uso de inteligencia artificial para la generación de código y ejecución de simulaciones representa un valor agregado en términos de agilidad y precisión del análisis, alineándose con lo planteado por Cárdenas et al. (2020) y Bohórquez & Medina (2009).

Otro aspecto relevante es la variabilidad de la demanda según la franja horaria. La acumulación de clientes observada en los horarios pico sugiere la necesidad de implementar estrategias dinámicas de asignación de personal, como lo plantean Castro-Jiménez et al. (2025). Esta medida podría maximizar la eficiencia sin incurrir en el sobredimensionamiento de recursos en momentos de baja afluencia.

Además, los hallazgos del presente estudio reafirman la utilidad de la teoría de colas como herramienta para modelar situaciones reales y proponer soluciones fundamentadas en evidencia empírica. La caracterización estadística de los tiempos de llegada y atención permitió adaptar los modelos a las condiciones específicas del entorno analizado, lo cual representa un avance respecto a los enfoques tradicionales que suelen asumir distribuciones estándar como Poisson o exponencial (Montgomery & Runger, 2018).

Es importante señalar algunas limitaciones del estudio. La recolección de datos fue realizada en una sola tienda y durante un periodo corto de tiempo, lo cual restringe la generalización de los resultados. A futuro, se recomienda ampliar el muestreo a múltiples sucursales y considerar variables adicionales como el tipo de cliente, el número de productos por compra y el comportamiento de los empleados. Asimismo, sería pertinente evaluar el impacto de tecnologías complementarias como cajas de autocobro o asistentes digitales.

En síntesis, los resultados obtenidos respaldan la aplicación de modelos de simulación con soporte de inteligencia artificial como herramientas efectivas para optimizar procesos en el sector minorista. La evidencia sugiere que ajustes relativamente simples, como el número de cajeros o la asignación por turnos, pueden generar mejoras significativas en la experiencia del cliente y en el uso de los recursos operativos.

### ***Limitaciones y recomendaciones***

A pesar de los hallazgos significativos obtenidos en este estudio, es importante reconocer una serie de limitaciones que pueden haber influido en los resultados y que deben ser consideradas tanto para la interpretación como para futuras investigaciones.

En primer lugar, la recolección de datos se realizó durante un periodo limitado (marzo de 2025) y en una única tienda D1 ubicada en un contexto urbano específico. Esto restringe la generalización de los resultados a otras tiendas de la cadena con dinámicas distintas, como las ubicadas en zonas rurales o en áreas de menor densidad poblacional. La operación, demanda y comportamiento de los consumidores pueden variar sustancialmente en función de factores geográficos, socioeconómicos y culturales.

Además, los datos se capturaron manualmente por observación directa, lo que conlleva el riesgo de errores humanos y sesgos de medición, especialmente en la estimación de tiempos de llegada y atención. Si bien se aplicaron estrategias de validación y consistencia, el uso de herramientas automáticas de captura (como sensores o cámaras con reconocimiento de eventos) podría mejorar la precisión y confiabilidad de los registros en investigaciones futuras.

Otra limitación importante es que el modelo de simulación se construyó sobre ciertos supuestos simplificadores, como la disciplina de cola FIFO, la independencia de las llegadas, la distribución triangular o uniforme de tiempos, y la atención constante por parte de los cajeros. En la práctica, existen múltiples factores dinámicos que no fueron modelados, como interrupciones del personal, variabilidad por tipo de producto, uso de medios de pago, o cambios en el comportamiento del cliente durante promociones. Estas simplificaciones, aunque necesarias para la viabilidad técnica del modelo, reducen la complejidad del sistema real.

Desde el punto de vista computacional, si bien se utilizó un modelo basado en Python con la librería SimPy y se personalizó un script proveniente de GitHub, este no fue sometido a validaciones cruzadas con otros lenguajes o plataformas de simulación como Arena, AnyLogic o ProModel. Esto limita la comparación de resultados entre diferentes motores de simulación y podría generar diferencias en la precisión de los indicadores estimados.

En términos de recomendaciones prácticas, se sugiere a la cadena D1 explorar alternativas de optimización como la rotación dinámica del personal de caja, según la afluencia por hora, o la implementación de sistemas de autoservicio en puntos específicos. El modelo desarrollado puede ser fácilmente adaptado para probar dichas configuraciones, lo que representa una oportunidad para avanzar hacia una toma de decisiones más basada en datos. Asimismo, la empresa podría establecer mecanismos de monitoreo continuo de sus tiempos de atención y filas, integrando sensores de conteo de personas o sistemas de gestión de turnos, para retroalimentar y ajustar sus operaciones en tiempo real.

Para investigaciones futuras, se recomienda ampliar el estudio a otras tiendas de la misma cadena y considerar distintas franjas horarias y días de la semana. También sería valioso incorporar variables adicionales, como la segmentación por tipo de cliente, cantidad de productos adquiridos o nivel de ocupación de las cajas por hora. Finalmente, podría evaluarse la viabilidad de modelos híbridos que integren inteligencia artificial para predecir la demanda en tiempo real y ajustar automáticamente los recursos asignados en función de la predicción.

## **Conclusiones**

Este estudio demuestra que la integración de técnicas de simulación con herramientas de inteligencia artificial constituye una estrategia eficaz para analizar y mejorar procesos operativos en entornos comerciales. En particular, la aplicación de la teoría de colas en un supermercado de formato hard discount como D1 permitió evaluar de manera precisa el desempeño del sistema de atención al cliente en función del número de servidores disponibles.

El uso de software, en este caso ProModel, facilitó la caracterización estadística de las tasas de llegada y servicio, lo que posibilitó la construcción de modelos ajustados a la realidad observada y la utilización de distribuciones de probabilidad que no se usan en los modelos manuales que sugiere la Investigación de operaciones. La implementación de estos modelos en Google Colab permitió simular múltiples escenarios y evaluar sus implicaciones operativas sin comprometer recursos físicos.

Los resultados obtenidos muestran que operar con un solo servidor conlleva altos niveles de saturación y tiempos de espera excesivos, lo cual afecta negativamente la percepción del servicio. Por otro lado, el uso de tres servidores minimiza los tiempos de espera, pero incurre en una baja utilización de recursos, lo que podría considerarse ineficiente. En contraste, el escenario con dos

servidores alcanza un balance óptimo entre eficiencia y calidad del servicio, presentándose como la alternativa más adecuada para el contexto analizado.

Estos hallazgos ratifican la pertinencia de utilizar modelos de simulación y algoritmos inteligentes como soporte para la toma de decisiones estratégicas en la gestión de operaciones, especialmente en sectores sensibles al tiempo de atención y a la experiencia del cliente.

Finalmente, se propone que este enfoque metodológico sea replicado en otros formatos comerciales con dinámicas similares, como Tiendas Ara o Justo & Bueno. La integración de simulación, inteligencia artificial y análisis estadístico representa una alternativa poderosa y de bajo costo para empresas que buscan optimizar sus operaciones a través de la analítica de datos y la transformación digital.

## Referencias

- Abdselam, G., & Gonzales, M. (2024). *La importancia de los Cuellos de Botella para la optimización de los procesos productivos*. <https://biblus.us.es/bibing/proyectos/abreproy/94916/fichero/TFG-4916+Abdselam+Miró.pdf>
- Arias, D., Nieto, E., Troncoso, A & Roncallo, L. (2024). Análisis de líneas de espera en una comercializadora mediante simulación. *Revista de Simulación y Negocios*, 11(1), 30–40. <https://doi.org/10.17981/bilo.6.1.2024.01>
- Bernal, F., Sandoval, A., & Muñoz, C. (2015). Estimación del potencial de los trenes de media y alta velocidad en Colombia. *Ingeniería e Investigación*, 16(2), 262-277. <https://www.scielo.org.mx/pdf/iit/v16n2/v16n2a10.pdf>
- Bertsekas, D., & Gallager, R. (1992). *Data networks* (2nd ed.). Prentice Hall.
- Bohórquez, P., & Medina, S. (2009). *Simulación de líneas de producción y servicios mediante el uso de Python-Simpy*. [http://bdigital.ula.ve/storage/pdftesis/pregrado/tde\\_arquivos/8/TDE-2012-07-28T22:00:01Z-1566/Publico/bohorquezpedro.pdf](http://bdigital.ula.ve/storage/pdftesis/pregrado/tde_arquivos/8/TDE-2012-07-28T22:00:01Z-1566/Publico/bohorquezpedro.pdf)
- Burbano-Pantoja, C., Ángel, M. A., Valdivieso-Miranda, J., & Burbano, H. (2025). Estrategia pedagógica para modelar sistemas de colas a partir de datos reales. *Saber, Ciencia y Libertad*, 20(1), 211–230. <https://revistas.unilibre.edu.co/index.php/saber/article/view/12614>
- Cárdenas, R., Pérez, M., Tejada, A., & Cevallos, L. (2019). Aplicación de un modelo híbrido de teoría de colas y algoritmo evolutivo para medir la optimización en el servicio de atención al cliente

- en un local de comidas rápidas. *Ecuadorian Science Journal*, 3(1), 15–22. <https://doi.org/10.46480/esj.3.1.23>
- Castro-Jiménez, N., Del Campo-de-La-Salas, S., Jay-Archibol, J., Caballero-De-La-Hoz, B., & Troncoso-Palacio, A. (2025). Gestión de colas para mejorar la satisfacción de clientes en supermercados. *BILO: Boletín de Innovación Logística y Operaciones*, 7(2), 1–9. <https://doi.org/10.17981/bilo.7.2.2025.01>
- García, H., & Moreno, Q. (2014). Evaluación de una nueva estrategia de atención a los despegues. *Ingeniería*, 15(1), 51–61. <https://www.redalyc.org/pdf/404/40429649006.pdf>
- Gómez, F. A. (2008). Aplicación de teoría de colas en una entidad financiera: herramienta para el mejoramiento de los procesos de atención al cliente. *Revista Universidad EAFIT*, 44(150), 51–63. <https://publicaciones.eafit.edu.co/index.php/revista-universidad-eafit/article/view/154>
- Gonzales, S., Ripalda, M., Flores, I., & Bernal, A. (2015). Propiedades de sistemas tipo M/G/1/K con estado de espera. *Ingeniería e Investigación*, 16(3), 345–354. [https://www.scielo.org.mx/scielo.php?script=sci\\_arttext&pid=S1405-77432015000300003](https://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1405-77432015000300003)
- Gross, D., Shortle, J. F., Thompson, J. M., & Harris, C. M. (2008). *Fundamentals of queueing theory* (4th ed.). John Wiley & Sons.
- Guerrero, H., & Guerrero, M. (2024). *Teoría de colas y simulación: Un enfoque fundamental y práctico*. [https://www.researchgate.net/publication/394856063\\_Teoria\\_de\\_colas\\_y\\_simulacion\\_Un\\_enfoque\\_fundamental\\_y\\_practico](https://www.researchgate.net/publication/394856063_Teoria_de_colas_y_simulacion_Un_enfoque_fundamental_y_practico)
- Hillier, F. S., & Lieberman, G. J. (2010). *Introduction to operations research* (9th ed.). McGraw-Hill Education.
- Kleinrock, L. (1975). *Queueing systems, volume 1: Theory*. Wiley-Interscience.
- Leon, L., & Vivanco, L. (2023). *Modelación para la mejora del servicio de atención al cliente mediante la teoría de colas*. <https://doi.org/10.26439/interfases2023.n017.6299>
- Linares, J., Vilalta, J., & Garza, R. (2019). La teoría de colas aplicada a una Oficina Comercial de Telecomunicaciones. *Ingeniería Industrial*, 41(2), 118–120. <https://www.redalyc.org/journal/3604/360464740007/html>
- Montgomery, D. C., & Runger, G. C. (2018). *Applied Statistics and probability for engineers* (7th ed.). Wiley.

- Mora, A. (2010). Estimadores del índice de cola y el valor en riesgo. *Cuadernos de Administración*, 44. [http://www.scielo.org.co/scielo.php?script=sci\\_arttext&pid=S0120-46452010000200005](http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S0120-46452010000200005)
- Morones, D., Rodríguez, R., & Castro, D. (2025). Optimización de disciplinas de colas con simulación de eventos discretos: Una revisión teórica y empírica. *Equilibrio Económico*, 21(59), 58–81. <https://revistas.uadec.mx/equilibrioeconomico/article/view/73>
- Muñoz-Vergara, A. (2019). Aplicación de la teoría de líneas de espera en el servicio de biblioteca de una organización educativa en Cartagena – Colombia. *Saber, Ciencia y Libertad*, 14(1), 171–178. <https://doi.org/10.18041/2382-3240/saber.2019v14n1.5236>
- Narváez, J., Ordoñez, W., & Paz, N. (2016). Análisis y simulación de tiempos de espera aplicando teoría de colas en la terminal de transporte de Popayán. *Publicaciones e Investigación*, 12(3), 90–100. <https://hemeroteca.unad.edu.co/index.php/publicaciones-e-investigacion/article/view/2966/5063>
- Portilla, L., Arias, L., & Fernández, S. (2010). Análisis de líneas de espera a través de teoría de colas y simulación. *Revista de Ingeniería*, 17(46), 56–61. <https://www.redalyc.org/pdf/849/84920977012.pdf>
- Rodríguez, L., & Lopez, M. (2016). Aplicación de técnicas de validación de un modelo de simulación de dinámica de sistemas. *Revista Latinoamericana de Ingeniería de Software*, 4 (4), 187-196. <https://doi.org/10.18294/relais.2016.187-196>
- Rodríguez, G., Hernando, S., González, A., & Hernandez, M. (2017). Análisis del servicio de urgencias aplicando teoría de líneas de espera (colas). *Contaduría y Administración*, 62(3), 719–732. <https://www.elsevier.es/es-revista-contaduria-administracion-87-articulo-analisis-del-servicio-urgencias-aplicando-S0186104217300359>
- Ross, S. M. (2014). *Introduction to probability models* (11th ed.). Academic Press.
- Singer, M., Donoso, P., & Scheller, A. (2008). Una introducción a la teoría de colas aplicada a la gestión de servicios. *Revista ABANTE*, 11(2), 70–80. <https://www.ceop.cl/wp-content/uploads/2010/11/Una-Introducción-a-la-Teoría-de-Colas.pdf>
- Solomon, M. R. (2018). *Comportamiento del consumidor: Comprando, poseyendo y siendo* (12.<sup>a</sup> ed.). Pearson.
- Taha, H. A. (2017). *Operations research: An introduction* (10th ed.). Pearson.


Velez-Ocampo, J., & Gonzalez-Perez, M. A. (2015). International expansion of Colombian firms: Understanding their emergence in foreign markets. *Cuadernos de Administración*, 28(51), 189-215. <https://doi.org/10.11144/Javeriana.cao28-51.iecf>



# Cacao For Peace And Region Programme – Procapaz: Una Estrategia que Contribuye al Fortalecimiento para el Desarrollo Comunitario

## Cacao for Peace and Regional Programme – Procapaz: A Strategy Contributing to the Strengthening of Community Development

---

Lady Johanna Peñaloza Farnán<sup>1</sup> 

Fecha de recepción: 27-08-2025 | Fecha de aprobación: 22-10-2025

---

### Resumen

El proyecto “Cacao for Peace and Region Programme” – PROCAPAZ, propone una estrategia participativa para fortalecer la cadena cacaofera de Rovira (Tolima) integrando innovación social y herramientas de la Ingeniería Industrial. Se empleó una metodología mixta de investigación-acción que inició con un diagnóstico participativo (FODA, encuestas a familias y estudiantes, entrevistas a productores y talleres comunitarios), continuó con el codiseño e implementación de estrategias (línea estandarizada de producción desde el precultivo hasta la transformación, Manual de Buenas Prácticas Agrícolas y plan de empalme generacional soportado en mentorías y bitácoras) y culminó con acciones de transferencia, validación y sistematización (Feria y Jornada cacaofera, uso de la bitácora como repositorio de evidencias y diseño/aprobación de un curso electivo de bioeconomía para su oferta en 2026). Entre los principales resultados se destacan la línea estandarizada con puntos críticos de proceso, BPA contextualizadas y una arquitectura de actores para la gobernanza colaborativa; además, se consolidó una línea base, productividad cercana a 600 kg/ha y rendimiento de fermentación alrededor del 70%, que orienta la ruta hacia trazabilidad, certificación y acceso a mercados diferenciados. Asimismo, se fortaleció la identidad cacaofera, aumentó la participación juvenil y se robustecieron capacidades técnicas y organizativas. En conjunto, la articulación universidad-comunidad-sector productivo, acompañada de dispositivos sociopedagógicos y estandarización técnica, habilita mejoras verificables en desempeño y cohesión social y sienta bases para el escalamiento regional.

### Palabras Clave

Cacao, Innovación social, Producción sostenible, Transferencia de conocimiento, Empalme generacional, Bioeconomía

---

1 Universidad de Ibagué, Grupo de investigación EULOGOS. ORCID: Correo: lady.penaloza@unibague.edu.co

## Abstract

The Cacao for Peace and Region Programme - PROCAPAZ advances a participatory strategy to strengthen the cacao value chain in Rovira (Tolima) by integrating social innovation with Industrial Engineering tools. A mixed-methods action-research design was used, beginning with a participatory assessment (SWOT analysis, household and student surveys, producer interviews, and community workshops), moving into the co-design and implementation of strategies (a standardized production line spanning pre-cultivation through processing, a Good Agricultural Practices, GAP Manual, and a generational-bridging plan supported by mentoring and learning logbooks), and culminating in transfer, validation, and systematization activities (a Cocoa Fair and an Experience Day, the use of the logbook as an evidence repository, and the design/approval of a bioeconomy elective course to be offered in 2026). Key results include the standardized line with defined critical control points, context-specific GAP guidelines, and a stakeholder architecture for collaborative governance. A baseline was also consolidated, productivity near 600 kg/ha and fermentation yield around 70%, that guides the pathway toward traceability, certification, and access to differentiated markets. In parallel, the project strengthened the territory's cacao identity, increased youth participation, and reinforced technical and organizational capacities. Overall, the articulation of university-community-productive sector, coupled with sociopedagogical devices and technical standardization, enables verifiable improvements in performance and social cohesion and lays the groundwork for regional scaling.

## Keywords

Cacao, Social Innovation, Sustainable Production, Knowledge Transfer, Generational Bridging, Bioeconomy

## Introducción

La región del Tolima ha enfrentado profundas dificultades derivadas del conflicto armado, la migración juvenil y la debilidad de las estructuras productivas rurales. El municipio de Rovira, históricamente marginado de programas estatales, encuentra en el cacao una alternativa estratégica para la construcción de paz y el desarrollo sostenible. Sin embargo, los productores enfrentan limitaciones en estandarización, apropiación tecnológica y articulación intergeneracional.

El proyecto PROCAPAZ surgió como respuesta a esta problemática, buscando articular la Ingeniería Industrial con la innovación social y la transferencia de conocimiento para fortalecer la cadena productiva del cacao, promover la resiliencia comunitaria y fomentar una identidad cultural cacaotera.

La Universidad de Ibagué, como institución de vocación regional, orienta sus funciones de docencia,

investigación y proyección social al desarrollo regional, impulsando proyectos que articulan conocimiento académico con necesidades del territorio. Desde la Ingeniería Industrial, a través del trabajo con el Grupo de investigación EULOGOS y los semilleros, se contribuye con pensamiento sistémico, diseño y estandarización de procesos, gestión de la calidad, logística y trazabilidad en contextos rurales, y transferencia de conocimiento mediante metodologías participativas. Esta articulación universidad, comunidad favorece la cocreación de soluciones pertinentes y sostenibles, fortalece la asociatividad y mejora la competitividad de las cadenas agroalimentarias del Tolima.

## Metodología

El proyecto se desarrolló bajo un diseño mixto de investigación-acción, adecuado para generar conocimiento aplicable junto con los actores del territorio y, a la vez, orientar mejoras prácticas en la cadena cacaoera de Rovira. La combinación de enfoques cuantitativos y cualitativos permitió triangular datos y construir interpretaciones compartidas, siguiendo criterios de diseño de métodos mixtos (Creswell & Plano, 2018) y de investigación-acción participativa en contextos de desarrollo territorial (Kemmis et al., 2014; Reason & Bradbury, 2008).

### Figura 1

*Proceso Metodológico Mixto*



Fuente: Elaboración propia basada en la metodología presentada por los estudiantes del semillero ARA.

Como productos metodológicos se generaron una bitácora colectiva y una línea estandarizada de producción implementada con estudiantes de la Institución Educativa Técnica Ambiental (I.E.T.A.), La Ceiba, los cuales aportaron a la sostenibilidad de la cadena; en conjunto, la metodología alcanzó el 90% de cumplimiento, consolidando un modelo efectivo de transferencia de conocimiento en Rovira.

A continuación, se describen las tres fases o etapas del desarrollo de la metodología del proyecto.

**Fase 1: Diagnóstico Participativo:** Se aplicó una metodología de investigación-acción con enfoque participativo.

Esta etapa inauguró el ciclo metodológico y tuvo como propósito caracterizar rigurosamente el sistema cacaotero y orientar el codiseño posterior. Se aplicaron encuestas a familias rurales y a estudiantes de la I.E.T.A. La Ceiba, construidas y administradas bajo el enfoque de diseño adaptado (*tailored design*) para asegurar muestreo pertinente, ítems claros y resguardo ético de los datos (Dillman et al., 2014). En complemento, se realizaron entrevistas semiestructuradas a productores con guiones flexibles para reconstruir trayectorias productivas, identificar barreras y reconocer márgenes de posibilidad en la cadena; esta técnica aportó profundidad contextual y permitió contrastar percepciones con los registros cuantitativos. Asimismo, se desarrollaron grupos focales/talleres con actores locales, utilizando protocolos de discusión y registros sistemáticos para garantizar saturación temática y validez interpretativa (Krueger & Casey, 2015). Finalmente, se ejecutó un análisis FODA y un mapeo de actores en sesiones participativas inspiradas en la evaluación rural participativa, con el fin de involucrar a la comunidad en la coproducción de conocimiento y en la priorización de temas (Chambers, 1994).

El uso articulado de estas técnicas, como las encuestas, entrevistas, grupos focales, FODA y mapeo, favoreció la triangulación y la corroboración de hallazgos propios de los métodos mixtos, fortaleciendo la credibilidad y la utilidad de la línea base (Creswell & Plano Clark, 2018). Más que anticipar resultados, la fase produjo insumos metodológicos: un banco de variables técnicas (productividad y desempeño de poscosecha) y socioorganizativas (asociatividad, relevo generacional), criterios y guías para trabajo de campo y categorías preliminares que orientaron el diseño de instrumentos formativos y la siguiente fase de codiseño.

Fase 2: Diseño e Implementación de Estrategias: Codiseño de una línea estandarizada de producción de cacao junto a productores y estudiantes, que incluye desde el cultivo hasta la transformación.

Esta fase traduce el diagnóstico en diseño operativo mediante un proceso participativo que integra saber experto y conocimiento local. El codiseño de la línea estandarizada (del precultivo a la transformación) funciona como un andamiaje metodológico: definir puntos críticos de proceso y criterios de calidad permite planear, ejecutar y auditar prácticas de cultivo y poscosecha con lógica de *plan-act-observe-reflect* propia de la investigación-acción (Kemmis et al., 2014; Reason & Bradbury, 2008). La iteración aula-finca y las pruebas en campo no buscan reportar resultados, sino verificar factibilidad y adecuación contextual (infraestructura, clima, mano de obra) antes de escalar. Además, anclar los parámetros de referencia. Por ejemplo, fermentación controlada y secado lento hasta, 7-8% de humedad, en guías y literatura del sector reduce el riesgo técnico (calidad sensorial e inocuidad) y aporta validez externa al diseño (ICCO, s. f.; Afoakwa, 2014; Dzelagha et al., 2020). En paralelo, el “Manual de Buenas Prácticas Agrícolas” (BPA) opera como objeto de frontera que estandariza lenguaje y procedimientos entre actores; y el plan de empalme generacional, con mentorías y bitácora colectiva es la tecnología social que moviliza el conocimiento tácito, favoreciendo apropiación y futura adopción (Chambers, 1994).

Fase 3: Transferencia, Validación y Sistematización: Se desarrollaron espacios de socialización y validación comunitaria, como la Feria Cacaotera y la Jornada de Experiencias Cacaoteras, con demostraciones, degustaciones y diálogo entre generaciones.

Metodológicamente, esta fase cierra el ciclo de aprendizaje y crea condiciones de sostenibilidad. La transferencia se apoya en enfoques de extensión participativa (p. ej., “Farmer Field Schools” y mentorías entre pares) que la evidencia vincula con mayor adopción de buenas prácticas y agencia juvenil en contextos rurales; por eso se incorporan como dispositivos formativos y no solo como actividades (FAO, 2020, 2021). La validación comunitaria a través de espacios deliberativos (p. ej., ferias/jornadas) funciona como evaluación formativa: contrasta lo diseñado con criterios de uso real, reduce riesgos de implementación y ofrece retroalimentación estructurada para ajustes.

La bitácora actúa como rastro de auditoría (*audit trail*), integrando evidencia de campo para la toma de decisiones pedagógicas y técnicas. Finalmente, la institucionalización mediante el diseño y aprobación de un curso electivo en bioeconomía asegura transferencia curricular y escalamiento temporal de capacidades, alineado con agendas internacionales de formación para la circularidad e innovación (European Commission, s. f.; EUBioNet, s. f.). Todo ello se gobierna bajo la Triple Hélice - universidad, sector productivo y Estado/comunidad- para alinear conocimiento, mercado y política pública alrededor de metas compartidas de competitividad y desarrollo regional (Etzkowitz & Leydesdorff, 2000).

### **Actores Clave**

La arquitectura de actores de PROCAPAZ (véase Tabla 1) se diseñó como un arreglo colaborativo que combina funciones académicas, formativas, productivas, socioculturales y de apoyo técnico-institucional. Este diseño responde al enfoque de Triple Hélice, donde universidad, sector productivo y Estado/comunidad coproducen soluciones y dinamizan la innovación en el territorio (Etzkowitz & Leydesdorff, 2000). La forma de coordinación privilegia la gobernanza colaborativa -reglas, rutinas y capacidades compartidas para alcanzar metas comunes-, condición asociada a mejores resultados colectivos cuando existen objetivos claros, liderazgo distribuido y mecanismos de rendición de cuentas (Emerson & Nabatchi, 2015; Provan & Kenis, 2008).

**Tabla 1***Actores clave y roles en el proyecto PROCAPAZ*

Actor	Rol en el proyecto
Universidad de Ibagué - Grupo EULOGOS	Coordinación académica, investigación, mentoría y transferencia de conocimiento.
Institución Educativa Técnica Agroindustrial La Ceiba	Formación de estudiantes, prácticas de estandarización de procesos.
Productores de cacao de Rovira	Participación en talleres, mentorías y validación de estrategias productivas.
Actores comunitarios y culturales	Rescate de saberes ancestrales, talleres culturales.
Estudiantes del semillero de investigación ARA (Avancemos Región Activa)	Asistencia en diagnóstico, levantamiento de información y construcción de propuestas.
AGROSAVIA	Apoyo institucional, articulación territorial y fortalecimiento ambiental.

Fuente: Elaboración propia

Dentro de esta red, la Universidad de Ibagué-Grupo EULOGOS articula investigación aplicada, mentoría y transferencia; la I.E.T.A. La Ceiba opera como núcleo formativo y puente con familias rurales; las asociaciones y productores aportan conocimiento práctico y validan estándares; los actores comunitarios y culturales fortalecen identidad y apropiación; el semillero ARA cumple el rol de *boundary spanner* estudiantil que conecta aula, laboratorio y finca; y AGROSAVIA contribuye a la articulación técnica y ambiental. Esta configuración se alinea con la literatura sobre acción colectiva y acceso a mercados, que destaca el valor de la asociatividad y de las alianzas para mejorar negociación, financiamiento y servicios a pequeños productores (Markelova et al., 2009), así como con enfoques de participación y evaluación rural que legitiman el protagonismo comunitario en el ciclo de aprendizaje-acción (Chambers, 1994).

A continuación, en la Figura 2, se sintetiza la articulación de actores relevantes situando el cacao como eje organizador de procesos de capacitación, tecnología, comercialización y sostenibilidad. El esquema encarna los principios de *impacto colectivo* agenda compartida, medición conjunta, actividades mutuamente reforzadas, comunicación continua y soporte de una organización articuladora necesarios para sostener cambios a escala territorial (Kania & Kramer, 2011). En este marco, la convergencia entre liderazgo académico, participación productiva y soporte institucional habilita mentorías intergeneracionales, estandarización de procesos y apropiación cultural, generando capacidades duraderas y condiciones para el escalamiento de resultados.

**Figura 2**

*Articulación de Actores Relevantes*



Fuente: creación de la IA GTP con base al desarrollo del proyecto.

La Figura 2 representa gráficamente el modelo de colaboración del proyecto, mostrando cómo los sectores del territorio convergen alrededor de la cadena de valor del cacao. La disposición en red facilita la coordinación interinstitucional, la coproducción de conocimiento y la transferencia desde la universidad hacia el sistema productivo, y viceversa, fortaleciendo una dinámica de aprendizaje continuo y gobernanza compartida (Emerson & Nabatchi, 2015; Provan & Kenis, 2008).

## Resultados

Con base en los resultados del desarrollo del proyecto a continuación, se describe a destalle los logros con la misma estructura metodológica en sus tres fases:

***Diagnóstico Participativo: Se realizó un Análisis FODA (ver Tabla 2), Encuestas a Familias Rurales del Municipio de Rovira, Entrevistas a Productores y Talleres Comunitarios***

En primera instancia y como se muestra en la Tabla 2, la cual presenta la matriz FODA de la producción de cacao en Rovira, y fue elaborada a partir del diagnóstico participativo del proyecto PROCAPAZ con productores, comunidad educativa y actores institucionales, el análisis organiza los factores internos (fortalezas y debilidades) y externos (oportunidades y amenazas) que inciden en la cadena cacaotera

local, con el fin de orientar decisiones técnicas y sociales. Esta síntesis permite reconocer ventajas como el cacao fino de aroma y el conocimiento tradicional, al tiempo que visibiliza brechas en estandarización y apropiación tecnológica, así como riesgos contextuales que condicionan la sostenibilidad.

**Tabla 2**

*Matriz FODA de la producción de cacao en Rovira*

Fortalezas	Oportunidades	Debilidades	Amenazas
Condiciones climáticas favorables para el cacao fino de aroma.	Potencial de acceso a mercados internacionales especializados.	Falta de estandarización en procesos de producción y transformación.	Migración juvenil que reduce la mano de obra local.
Conocimiento ancestral de productores locales.	Posibilidad de asociatividad entre productores y acceso a programas de cooperación.	Escasa apropiación tecnológica.	Variabilidad en precios internacionales del cacao.
Apoyo institucional y académico (Universidad de Ibagué, ADR, CORTOLIMA).	Inclusión del cacao en programas de desarrollo rural sostenible.	Débil articulación entre generaciones de productores.	Problemas ambientales como deforestación y cambio climático.

Fuente: Elaboración propia

La lectura integrada del FODA sugiere un alto potencial de mejora si se aprovechan las fortalezas y oportunidades mediante cuatro frentes: (1) estandarización de procesos y adopción de buenas prácticas; (2) transferencia tecnológica y formación continua para productores y jóvenes; (3) fortalecimiento asociativo para negociar, financiarse y acceder a mercados diferenciados; y (4) gestión de riesgos frente a migración juvenil, variabilidad de precios y retos ambientales. En conjunto, estas líneas de acción convierten el FODA en una hoja de ruta operativa para consolidar la competitividad del cacao de Rovira y su aporte al desarrollo regional.

Por su parte, en cuanto al desarrollo de las encuestas se aplicaron a las familias rurales (padres/madres) de estudiantes de la I.E.T.A. La Ceiba mediante formularios estructurados distribuidos por la institución (registradas como Anexo 4. Encuesta a padres de familia), y se complementaron con una encuesta sociodemográfica a los propios estudiantes a través de Google Forms. Los instrumentos capturaron variables personales, familiares, educativas, de acceso tecnológico, intereses y salud, garantizando la confidencialidad y presentando los resultados en gráficos para facilitar su lectura. En el levantamiento se evidenciaron limitaciones de conectividad del hogar rural, que se registraron como hallazgos del contexto.

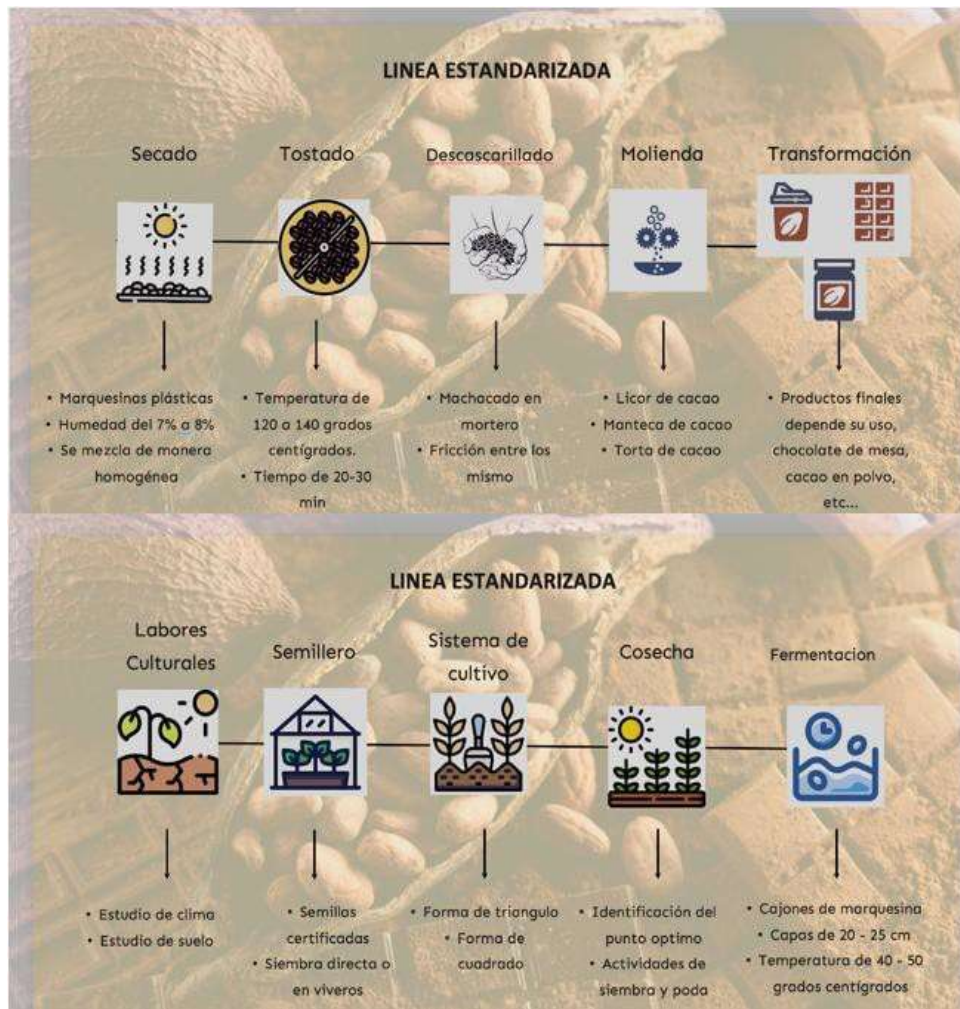
Las entrevistas a productores se desarrollaron con un muestreo intencional de 10 pequeños cacaoteros ubicados en Rovira y, de apoyo, en Ibagué. Se empleó un formato semiestructurado de ocho preguntas aplicado en visitas a finca o por llamada telefónica cuando las distancias o el acceso lo exigían. Previo consentimiento, las conversaciones fueron grabadas con fines académicos y sistematizadas en fichas técnicas para alimentar el diagnóstico y el diseño de la línea estandarizada de procesos.

Los talleres comunitarios se desarrollaron en la institución educativa y con actores del territorio. Además, se realizó un conversatorio (grupo focal) con productores, funcionarios y comunidad académica; aunque fue planificado presencial, se adaptó a modalidad remota (Google Meet) para ampliar la participación intermunicipal. Estas actividades evidenciaron alta participación estudiantil y permitieron identificar categorías para apropiación cultural y fortalecimiento de capacidades.

### ***Diseño de Estrategias: Construcción de una Línea Estandarizada de Producción de Cacao***

A partir de los hallazgos del diagnóstico, se desarrolló un proceso de co-diseño con productores y con estudiantes de la I.E.T.A. La Ceiba orientado a construir la línea estandarizada de producción de cacao (ver Figura 1) y a definir un plan de empalme generacional. Para ello, se integraron revisión de fuentes técnicas (fichas y matriz bibliográfica), mapeo de operaciones de la cadena (labores culturales, semillero, sistema de cultivo, cosecha, fermentación, secado, empaque y almacenamiento) y diseño de puntos de control y criterios de calidad para cada etapa. Las propuestas se validaron iterativamente mediante sesiones en aula y visitas a finca, ajustando procedimientos y materiales según la factibilidad local. El plan de empalme organizó mentorías intergeneracionales, talleres de creación de productos con valor agregado, uso de bitácoras de saberes y actividades de socialización comunitaria, definiendo responsables y mecanismos de seguimiento para la adopción de las prácticas. Esta fase articuló estándares técnicos con estrategias pedagógicas y culturales, habilitando la transferencia efectiva de conocimiento en el territorio.

**Figura 3**  
*Línea Estandarizada de Producción de Cacao*



Fuente: Elaboración propia (estudiantes del Semillero ARA).

Como se evidencia en la anterior imagen, la línea estandarizada ilustrada integra las etapas clave de la cadena cacaotera desde el precultivo hasta la transformación: inicia con labores culturales (estudios de clima y suelo), sigue con el semillero usando semillas certificadas y siembra directa o en vivero, define el sistema de cultivo en trazos de triángulo o cuadrado y prosigue a cosecha en el punto óptimo, acompañada de labores de siembra y poda. En poscosecha, la fermentación se realiza en cajones de marquesina, en capas de 20-25 cm y a 40-50 °C; luego el secado en marquesinas plásticas hasta 7-8 % de humedad, con mezclado homogéneo. El tostado se ejecuta entre 120-140 °C durante 20-30 min; el descascarillado por fricción/machacado en mortero; y la molienda genera licor, manteca y torta de cacao, insumos que finalmente se transforman en productos como chocolate de mesa o cacao en polvo.

La Tabla 3 sintetiza los principales indicadores técnicos y de mercado identificados en Rovira para establecer una línea base de la cadena cacaotera. Se observa una productividad promedio de 600 kg/ha, inferior al referente nacional considerado (900 kg/ha), un rendimiento de fermentación del 70% que evidencia oportunidades de mejora en poscosecha, un perfil de calidad medio-alto reconocido por su aroma fino, pero con heterogeneidad entre lotes, y una competitividad baja-media condicionada por la ausencia de certificaciones y sistemas de trazabilidad. Esta información orienta la priorización de acciones de estandarización y fortalecimiento organizativo.

**Tabla 3**

*Indicadores de productividad, calidad y competitividad en Rovira*

Indicador	Resultado promedio	Observación
Productividad por hectárea	600 kg/ha	Inferior al promedio nacional (900 kg/ha)
Rendimiento de fermentación	70%	Necesita mejoras en técnicas de poscosecha
Calidad (perfil de sabor)	Medio-alto	Reconocido por aroma fino, pero poco homogéneo
Competitividad en mercado	Baja-media	Limitada por falta de certificaciones y trazabilidad

Fuente: Elaboración propia

La lectura integrada de los indicadores sugiere centrar la intervención en tres frentes: primero, mejora de poscosecha (fermentación, secado y control de humedad) para elevar rendimiento y homogeneizar el perfil sensorial; segundo, adopción de buenas prácticas agrícolas y de calidad, con verificación sistemática de puntos críticos para cerrar la brecha de productividad; y tercero, gestión de certificaciones y trazabilidad que habiliten acceso a segmentos premium y mejoren la posición competitiva del cacao de Rovira. Estas medidas, articuladas con capacitación y asociatividad, constituyen la hoja de ruta para incrementar el desempeño técnico y comercial del territorio.

### **Plan de empalme generacional**

Como parte de la estrategia de transferencia e implementación, se desarrollaron dos espacios de socialización y validación en la Institución Educativa Técnica Agroindustrial La Ceiba (Rovira): la Feria Cacaotera (17 de abril, 08:00) y la Jornada de Experiencias Cacaoteras (9 de noviembre, 09:00). Ambos eventos, convocados por la Universidad de Ibagué, reunieron a estudiantes y productores para compartir aprendizajes, discutir oportunidades de fortalecimiento de la cadena y afianzar la identidad cacaotera del territorio. En este marco se presentó una muestra de productos innovadores a base de cacao, elaborados por estudiantes de la Técnica Agroindustrial, lo que permitió contrastar los contenidos técnicos del proyecto con la validación comunitaria.

**Figura 4**  
*Talleres para el Empalme Generacional*



Fuente: Elaboración propia

Como se evidencia en la Figura 4, “Plan de Empalme Generacional”, las actividades incluyeron mesas de diálogo y demostraciones/degustaciones de preparaciones a base de cacao, funcionando como mecanismos formales de transferencia de conocimiento y de validación comunitaria de los avances técnico-metodológicos del proyecto.

### **Elementos clave para el empalme generacional**

La siguiente figura sintetiza el dispositivo sociopedagógico del proyecto para asegurar relevo y apropiación de saberes en la cadena cacaotera de Rovira. El diagrama muestra el tránsito desde los saberes ancestrales de productores mayores hacia una mentoría intergeneracional, apoyada en transferencia de conocimiento mediante talleres, trabajo en aula y bitácora, que convoca y sostiene la participación juvenil. Este flujo se complementa con espacios de innovación y valor agregado, orientando los aprendizajes hacia resultados tangibles en productos y prácticas.

**Figura 5**

*Elementos clave para el empalme generacional*



Fuente: Elaboración de la IA GPT con base en datos del desarrollo del proyecto.

Los elementos de la estrategia, bajo esta estructura ilustran cómo el empalme generacional opera como puente entre tradición y técnica: la mentoría y la bitácora canalizan el conocimiento práctico hacia procedimientos estandarizados, mientras que la participación juvenil y los ejercicios de innovación traducen ese aprendizaje en mejoras productivas y culturales. En conjunto, el esquema explica el mecanismo por el cual se fortalece la cultura cacaotera del territorio y se generan condiciones para la continuidad de las prácticas aprendidas más allá del proyecto.

### **Bitácora para el diálogo generacional**

A continuación, en la Figura 5 muestra la Bitácora colectiva para el empalme generacional, instrumento pedagógico y de investigación-acción utilizado en PROCAPAZ para documentar el intercambio de saberes entre productores con trayectoria y jóvenes de la I.E.T.A. La Ceiba. La bitácora invitó a cada participante a registrar prácticas del cultivo y la poscosecha, relatos familiares, reflexiones sobre identidad cacaotera y evidencias de las actividades (talleres, mentorías y ejercicios en finca), funcionando como guía de trabajo durante el proceso formativo.

El uso de la bitácora fortaleció la apropiación cultural y la continuidad del aprendizaje más allá de las sesiones presenciales, al sistematizar conocimientos, acuerdos y compromisos intergeneracionales. Estos registros sirvieron como insumo para ajustar la línea estandarizada, orientar los talleres de innovación y retroalimentar la feria cacaotera, contribuyendo a la sostenibilidad social del proyecto y a la consolidación de una cultura cacaotera en Rovira.

**Figura 6**

*Bitácora colectiva para el empalme generacional*



Fuente: Elaboración propia.

**Figura 7**

*Bitácora Colectiva - Resultados Clave*



Fuente: Elaboración propia.

Como se evidencia en la anterior imagen “Bitácora colectiva - Resultados clave” la cual fue el resultado de un proceso de análisis de la Bitácora colectiva elaborada por los padres de familia de los estudiantes de la I.E.T.E. La Ceiba, presenta un diagrama radial cuyo núcleo es el empalme generacional. Desde ese centro se despliegan cinco ámbitos que se interrelacionan: la apropiación e identidad cultural cacaotera, que reúne relatos y categorías para fortalecer el sentido de pertenencia; la transferencia intergeneracional mediante mentorías y diálogo entre productores mayores y jóvenes; la innovación y el valor agregado, donde se registran los talleres de creación de productos a base de cacao con sello local; la validación y socialización comunitaria, que conduce los aprendizajes a la Feria Cacaotera como espacio de devolución y verificación social; y la sistematización para la mejora del proyecto, que utiliza la bitácora como repositorio de evidencias para orientar decisiones pedagógicas y técnicas. El conjunto muestra una ruta continua de aprendizaje y acción que conecta dimensiones culturales, formativas y productivas de la cadena cacaotera en Rovira.

### **Contribución de estudiantes del semillero Avancemos región activa - ARA**

El resultado más importante de este proyecto fue el poder generar una articulación desde la academia con la empresa y comunidad. Desde el “Semillero de Investigación Avancemos Región Activa” (ARA) participaron estudiantes del programa de Ingeniería Industrial de la Universidad de Ibagué. Su intervención fue decisiva: apoyaron el diagnóstico participativo en territorio, el levantamiento y la sistematización de información y la formulación de propuestas técnicas para la estandarización de procesos y el plan de empalme generacional. Este trabajo formativo-aplicado fortaleció el vínculo universidad-comunidad y, además, materializó el enfoque de la Triple Hélice, articulación universidad - empresa/sector productivo - Estado- propuesto por Etzkowitz y Leydesdorff (2000), en el que los estudiantes actúan como agentes de innovación que conectan la investigación académica con necesidades productivas y de política pública del territorio, contribuyendo de manera directa al logro de los objetivos del proyecto.

En este sentido, la tabla 4 sintetiza los aportes de dos Trabajos de Grado desarrollados en el marco del proyecto PROCAPAZ, organizando la evidencia en tres columnas; integrantes, conclusiones y recomendaciones, para facilitar la trazabilidad entre resultados y acciones propuestas. El primer estudio se centra en la estandarización técnico-operativa de la cadena de cacao; el segundo, en el empalme generacional y la apropiación cultural. Su lectura conjunta permite comprender el alcance de la intervención y orientar decisiones de mejora en el territorio.

**Tabla 4***Conclusiones y Recomendaciones de Trabajos de Grado (TG)*

Nombre del TG	Integrantes	Conclusiones (síntesis)	Recomendaciones
Construcción de una línea estandarizada del desarrollo en operaciones y actividades desde la siembra del cacao hasta la transformación como producto industrial con los estudiantes de la I.E.T.A. La Ceiba	Pablo Andrés Acosta Velásquez; Iván Camilo Alape Hernández	<ul style="list-style-type: none"> <li>- Se diseñó y validó una línea estandarizada de producción que mejora la calidad del grano y la eficiencia operativa.</li> <li>- La capacitación técnica y el trabajo con estudiantes fortalecen habilidades y promueven relevo generacional.</li> <li>- La colaboración productores-academia potencia la adopción de buenas prácticas y la sostenibilidad.</li> <li>- La estandarización abre opciones para acceder a mercados con mayores exigencias de calidad.</li> </ul>	<ul style="list-style-type: none"> <li>- Implementar capacitación continua en producción sostenible y estandarización.</li> <li>- Fortalecer la asociatividad para compras/ventas colectivas y acceso a recursos.</li> <li>- Promocionar el cacao fino de aroma en nichos premium.</li> <li>- Crear alianzas con entidades públicas/privadas para soporte técnico y de mercado.</li> </ul>
Construcción de una estrategia de empalme generacional a partir de un diagnóstico situacional para el fomento de una cultura cacaotera en el municipio de Rovira	Julián Mateo Saavedra; María Juliana Guerrero; Duván Leandro Gómez	<ul style="list-style-type: none"> <li>- El diagnóstico evidenció brechas en prácticas productivas, comercialización y organización; se requiere intervención integral.</li> <li>- Existen desafíos socioambientales (deforestación, clima, desigualdad) que afectan la sostenibilidad.</li> <li>- Productores ven potencial en la transformación con valor agregado, pero demandan apoyo técnico y financiero.</li> <li>- Actividades como bitácora colectiva, talleres y feria cacaotera fortalecen identidad y participación juvenil.</li> </ul>	<ul style="list-style-type: none"> <li>- Operacionalizar el programa de mentoría entre productores mayores y jóvenes de la I.E. La Ceiba.</li> <li>- Integrar la cultura local en la formación y en los productos derivados (identidad territorial).</li> <li>- Impulsar rutas de comercialización y trazabilidad con apoyo institucional.</li> <li>- Diseñar esquemas de financiación y asistencia técnica para adopción de tecnologías y buenas prácticas.</li> </ul>

Fuente: Elaboración propia

En conjunto, los hallazgos convergen en cuatro líneas estratégicas: primero, la estandarización y adopción de buenas prácticas agrícolas (BPA) en toda la cadena; por otra parte, la formación continua y mentoría intergeneracional para asegurar relevo juvenil; así mismo el fortalecimiento asociativo con mecanismos de financiación y asistencia técnica; y finalmente el acceso a mercados diferenciados mediante trazabilidad y certificaciones. La priorización e implementación coordinada de estas recomendaciones, bajo un esquema de Triple Hélice (universidad-sector productivo-Estado),

constituye la hoja de ruta para escalar PROCAPAZ y medir impactos en productividad, calidad y cohesión social del territorio.

Además de los hallazgos principales expuestos, el proyecto reporta otros resultados relevantes que refuerzan la articulación entre investigación, formación y acción en el territorio. Estos productos y logros operativos fortalecen la sostenibilidad técnica y social de la cadena cacaoera en Rovira y evidencian la traducción de la metodología en mejoras concretas para los actores locales, particularmente en el fortalecimiento organizativo. En esta línea, se promovieron mecanismos de asociatividad y colaboración territorial orientados a mejorar la comercialización, el acceso a fuentes de financiación y la adopción tecnológica. Este enfoque se refleja tanto en los análisis del proyecto como en las recomendaciones de los trabajos de grado, que subrayan la creación y consolidación de asociaciones y redes de innovación rural.

### ***Transferencia e Implementación: Curso Electivo en Bioeconomía***

Derivado de los hallazgos del diagnóstico y de la línea estandarizada del cacao, se diseñó el curso electivo “Bioeconomía: Industria, Medio Ambiente y Desarrollo Social Sostenible”, concebido como eje de transferencia curricular para consolidar capacidades en sostenibilidad, calidad y valorización de productos derivados del cacao. El programa fue presentado ante la Dirección del Departamento de Logística y Ciencias de la Computación de la Universidad de Ibagué y, posteriormente, aprobado por el Consejo de Facultad de Ingeniería para su oferta en el semestre A de 2026. Este resultado asegura la institucionalización de los aprendizajes del proyecto y su escalamiento: articula casos y prácticas construidas con la comunidad (estandarización de procesos, mentoría intergeneracional, innovación de productos y ferias cacaoeras), y establece una ruta formativa que da continuidad, desde la academia, a la adopción de buenas prácticas, trazabilidad y modelos de bioeconomía territorial en la cadena cacaoera de Rovira.

La siguiente figura resume los componentes curriculares del curso electivo para el Programa de Ingeniería Industrial (modalidad presencial, asignatura teórica, estrategias ABP-talleres- indagación, 3 créditos). En coherencia con PROCAPAZ, el cacao será una de las líneas de estudio, abordada mediante casos reales del territorio para articular los contenidos de bioeconomía con problemáticas y oportunidades locales.

**Figura 8**

*Programa Detallado de Asignatura (PDA) del curso electivo “Bioeconomía: Industria, Medio Ambiente y Desarrollo Social Sostenible”*



**Bioeconomía: Industria, Medio Ambiente y Desarrollo Social Sostenible.**

**1. Información general**

<b>Programa Académico</b>	Ingeniería Industrial		
<b>Núcleo de la formación</b>	Profesional		
<b>Tipo de asignatura</b>	Teórica		
<b>Modalidad</b>	<ul style="list-style-type: none"> <li>• Presencial</li> </ul>		
<b>Estrategia(s) pedagógica(s)</b>	ABP-Aprendizaje Basado en Proyectos, Talleres, Indagación		
<b>Tipología del salón</b>	<ul style="list-style-type: none"> <li>• Tipo I: Magistral y colaboración</li> </ul>		
<b>Dimensión</b>	<ul style="list-style-type: none"> <li>• Investigación</li> <li>• Espíritu emprendedor</li> </ul>		
<b>Código/grupo</b>			
<b>Número de Créditos</b>	3		
<b>Horas de acompañamiento directo</b>	48	<b>Horas de trabajo independiente</b>	96
		<b>Total de horas</b>	144

Fuente: Elaboración propia

Al incorporar estudios de caso sobre la cadena cacaoera del territorio, el curso garantiza la transferencia formativa de los resultados del proyecto hacia el currículo y promueve soluciones aplicadas en sostenibilidad, calidad y valorización, fortaleciendo la continuidad académica y la pertinencia regional de PROCAPAZ.

**Discusión**

Los resultados técnicos del proyecto -en particular la estandarización poscosecha- se alinean con lineamientos y evidencia internacional. Prácticas como la fermentación controlada seguida de un secado lento hasta ~7-8% de humedad son consistentes con las guías de la International Cocoa Organization y con revisiones recientes que relacionan parámetros de secado con la calidad sensorial y la inocuidad del grano (ICCO, s. f.; Dzelagha et al., 2020). Esta convergencia refuerza el foco metodológico adoptado por PROCAPAZ sobre puntos críticos de proceso (fermentación, secado y almacenamiento).

El déficit de productividad observado en Rovira se inscribe en tendencias nacionales que muestran rendimientos promedio históricamente bajos y heterogéneos en Colombia, lo que demanda extensión

técnica, estandarización y adopción tecnológica para cerrar brechas (FEDECACAO, s. f.; Ministerio de Agricultura y Desarrollo Rural, 2020). En paralelo, la apuesta por trazabilidad y certificaciones responde a señales claras del mercado: informes sectoriales evidencian la expansión de sellos (orgánico, Fairtrade, Rainforest Alliance) como vía de acceso a nichos premium, mientras que la Regulación de Deforestación de la UE (EUDR) incrementa los requisitos de debida diligencia y georreferenciación a nivel de finca para el cacao que ingrese al bloque (ITC, 2024; Rainforest Alliance, 2024; Ángel & Abenett, 2023; Ángel, 2024).

En el plano sociopedagógico, los dispositivos de mentorías intergeneracionales, bitácoras y talleres participativos están en línea con la literatura sobre extensión agrícola participativa. Experiencias como las “Farmer Field Schools” y los modelos de mentoría entre pares muestran aumentos en adopción de buenas prácticas, desarrollo de habilidades y empoderamiento juvenil en contextos rurales, tal como promueve el proyecto en Rovira (FAO, 2020, 2021).

El énfasis en asociatividad y redes colaborativas se sustenta en la teoría y evidencia de acción colectiva, que explica cómo las organizaciones de productores mejoran acceso a mercados, financiamiento y servicios, y se asocian con mejores resultados comerciales cuando existen reglas claras y apoyo institucional (Markelova et al., 2009). En términos de anclaje académico, el diseño del curso electivo de bioeconomía dialoga con marcos internacionales que posicionan la bioeconomía como vehículo para la circularidad y la descarbonización, y con iniciativas educativas que desarrollan competencias para su implementación, asegurando transferencia curricular y continuidad institucional de los aprendizajes del proyecto (European Commission, s. f.; EUBioNet, s. f.).

En conjunto, la triangulación entre estándares técnicos, extensión participativa y organización de productores, complementada por el anclaje curricular en bioeconomía, otorga validez externa a los resultados de PROCAPAZ y orienta su escalabilidad: priorizar poscosecha, trazabilidad/certificación y asociatividad, con soporte formativo y de políticas públicas, para capturar valor en mercados diferenciados y fortalecer la resiliencia territorial.

## **Conclusiones**

La implementación de enfoques de ingeniería industrial en el municipio de Rovira, Tolima, ha demostrado ser una herramienta eficaz para resolver problemáticas estructurales en la cadena de producción del cacao. A través de la estandarización de procesos, el diseño participativo de estrategias de producción y la introducción de tecnologías apropiadas, se han optimizado actividades desde la siembra hasta la transformación del cacao. Este enfoque no solo mejora la eficiencia operativa, sino que también contribuye al desarrollo regional, generando empleo digno, fortaleciendo el tejido social y promoviendo alternativas sostenibles frente a economías ilícitas.

La creación de una línea estandarizada de producción, acompañada por un programa de mentoría entre productores experimentados y jóvenes estudiantes de la Institución Técnica La Ceiba, ha

permitido mejorar la calidad del grano, reducir desperdicios y abrir oportunidades de acceso mercados premium, tanto nacionales como internacionales. Además, la certificación en prácticas sostenibles y el uso de tecnologías como sensores para la fermentación o el dron XAG para el monitoreo agrícola, han incrementado la competitividad del cacao producido en Rovira.

Frente a la migración rural y al envejecimiento de los productores, el proyecto “Cacao for Peace and Region” priorizó el relevo generacional mediante procesos participativos, mentoría y talleres con jóvenes de grados 10° y 11°. Estas acciones no solo transfieren conocimientos técnicos, sino también valores identitarios y saberes ancestrales, creando una cultura cacaotera sólida y resiliente. La apropiación cultural ha sido clave para conectar a las nuevas generaciones con su territorio y su historia productiva.

Dado el impacto positivo observado en Rovira, el modelo de mentoría, estandarización de procesos y participación comunitaria es replicable en otros municipios cacaoteros. No obstante, para sostener y escalar estos logros, es indispensable mejorar la infraestructura logística, reducir los tiempos de transporte, y avanzar en la trazabilidad del producto desde la finca hasta el mercado. Esto permitiría certificar la calidad y origen del cacao, agregando valor a la cadena y favoreciendo la entrada a mercados internacionales con mayores exigencias.

Como parte de las conclusiones, el proyecto identificó restricciones estructurales que condicionan la adopción de buenas prácticas y la sostenibilidad de los resultados: baja conectividad digital que limitó sesiones virtuales, debilidades de cultura organizacional en algunas asociaciones y carencias de infraestructura física para poscosecha y formación técnica.

Para mitigar estos obstáculos se implementaron modelos de mejora continua adaptados al contexto rural, estrategias lúdicas de capacitación orientadas a la apropiación práctica del conocimiento y acompañamiento desde Ingeniería para estructurar procesos productivos y logísticos, con avances verificables en estandarización y transferencia.

Persisten, no obstante, retos asociados al cierre de brechas tecnológicas, al fortalecimiento asociativo y a la inversión en infraestructura básica. Se recomienda consolidar alianzas territoriales y esquemas de financiación que garanticen continuidad a las acciones planteadas y posibiliten el escalamiento del modelo en la cadena cacaotera de Rovira.

## Referencias




- Afoakwa, E. O. (2014). *Cocoa processing and chocolate manufacture: From bean to bar*. CRC Press.
- Angel, M. & Abnett, K. (2023, June 13). *ICE aims to help cocoa, coffee firms meet EU deforestation rule*. Reuters. <https://www.reuters.com/sustainability/ice-aims-help-cocoa-coffee-firms-meet-eu-deforestation-rule-2023-06-13/>
- Angel, M. (2024, October 1). *Cocoa traceability rates fail to improve as EU deforestation law looms*. Reuters. <https://www.reuters.com/markets/commodities/cocoa-traceability-rates-fail-improve-eu-deforestation-law-looms-2024-10-01/>
- Chambers, R. (1994). The origins and practice of Participatory Rural Appraisal (PRA). *World Development*, 22(7), 953-969. [https://doi.org/10.1016/0305-750X\(94\)90141-4](https://doi.org/10.1016/0305-750X(94)90141-4)
- Creswell, J. W., & Plano, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). SAGE.
- Dillman, D. A., Smyth, J. D., & Christian, L. M. (2014). *Internet, phone, mail, and mixed-mode surveys: The tailored design method* (4th ed.). Wiley.
- Dzelagha, B. F., Nguemtchouin, M. M. G., Tamsa Arfao, A., Abdou Bouba, A., & Mbofung, C. M. F. (2020). A review of cocoa drying technologies and the effect on bean quality. *Foods*, 9(12), 1817. <https://doi.org/10.1155/2020/8830127>
- Emerson, K., & Nabatchi, T. (2015). *Collaborative governance regimes*. Georgetown University Press.
- Etzkowitz, H., & Leydesdorff, L. (2000). The dynamics of innovation: From national systems and “Mode 2” to a Triple Helix of university-industry-government relations. *Research Policy*, 29(2-3), 109-123. [https://doi.org/10.1016/S0048-7333\(99\)00055-4](https://doi.org/10.1016/S0048-7333(99)00055-4)
- EUBioNet. (s. f.). *Bioeconomy education initiatives*. <https://eubionet.eu/bioeconomy-education>
- European Commission. (s. f.). *Bioeconomy - Research and Innovation*. <https://research-and-innovation.ec.europa.eu>
- FAO. (2020). *Evaluation of FAO’s contributions to SDG2: Farmer Field Schools and youth empowerment*. Food and Agriculture Organization of the United Nations.
- FAO. (2021). *Evaluation of FAO’s contribution to Sustainable Development Goal 2 – “End hunger, achieve food security and improved nutrition and promote sustainable agriculture”*. Thematic Evaluation 10/2020. Rome: Food and Agriculture Organization of the United Nations. <https://sioc.minagricultura.gov.co/Acuicultura/Documentos/2020-09-30%20Cifras%20Sectoriales.pdf>

- Federación Nacional de Cacaoteros. (s. f.). *Investigación: estadísticas de rendimiento del cacao en Colombia*. Federación Nacional de Cacaoteros. <https://www.fedecacao.com.co/investigacion>
- International Cocoa Organization. (s. f.). *Harvesting & post-harvest; Processing cocoa*. International Cocoa Organization. <https://www.icco.org/harvesting-post-harvest-new>
- International Trade Centre. (2024). *The State of Sustainable Markets 2024: Statistics and Emerging Trends*. ITC/FiBL/ISEAL. <https://www.intracen.org/resources/publications/state-of-sustainable-markets-2024>
- Kania, J., & Kramer, M. (2011). Collective impact. *Stanford Social Innovation Review*, 9(1), 36-41. [https://ssir.org/articles/entry/collective\\_impact](https://ssir.org/articles/entry/collective_impact)
- Kemmis, S., McTaggart, R., & Nixon, R. (2014). *The action research planner: Doing critical participatory action research*. Springer. <https://doi.org/10.1007/978-981-4560-67-2>
- Markelova, H., Meinzen-Dick, R., Hellin, J., & Dohrn, S. (2009). Collective action for smallholder market access. *Food Policy*, 34(1), 1-7. <https://doi.org/10.1016/j.foodpol.2008.10.001>
- Ministerio de Agricultura y Desarrollo Rural. (2020). *Cifras sectoriales: Cadena de la acuicultura. Tercer trimestre 2020* [Presentación de PowerPoint]. Secretaría Técnica Nacional de la Cadena de la Acuicultura. <https://www.minagricultura.gov.co/paginas/default.aspx>
- Provan, K. G., & Kenis, P. (2008). Modes of network governance: Structure, management, and effectiveness. *Journal of Public Administration Research and Theory*, 18(2), 229-252. <https://doi.org/10.1093/jopart/mum015>
- Rainforest Alliance. (2024). *Cocoa Certification Data Report 2023*. Rainforest Alliance. <https://www.rainforest-alliance.org/business/certification/cocoa-certification-data-report-2023>
- Reason, P., & Bradbury, H. (Eds.). (2008). *The SAGE handbook of action research: Participative inquiry and practice* (2nd ed.). SAGE.
- Saavedra, J. M., Guerrero, M. J., & Gómez, D. L. (2024). *Construcción de una estrategia de empalme generacional para el fomento de una cultura cacaotera en Rovira*. Universidad de Ibagué.

# Integración de Redes y Protocolos de Comunicación para una Celda de Manufactura avanzada

## Integration of Communication Networks and Protocol Protocols in an Advanced Manufacturing Cell

---

Fernando Ricardez Rueda<sup>1</sup>, Hugo Enrique Noriega Pérez<sup>2</sup>, Carlos Adalberto Valenzuela Melchor<sup>3</sup>,  
Beatriz Adriana Esparza Ramírez<sup>4</sup>

---

Fecha de recepción: 01-08-2025 | Fecha de aprobación: 22-09-2025

### Resumen

La Manufactura Avanzada ha transformado los métodos de producción a través de la digitalización, la automatización y la interconexión de sistemas productivos. Este trabajo describe la integración del desarrollo de una celda de manufactura educativa basada en tecnologías como controladores lógicos programables PLC, sensores de visión, interfaces hombre-máquina HMI y para esta implementación el protocolo de comunicación PROFINET, destacándose que el sensor de visión artificial y PLC no son del mismo fabricante por lo que se usó una librería específica y un sistema de relevadores físicos para adaptar las señales digitales de cierre entre los equipos. El desarrollo e integración de la celda logró fortalecer las competencias tecnológicas de docentes y estudiantes participantes en la implementación del sistema automatizado y conectado en tiempo real.

---

1 Maestro en Enseñanza Universitaria del Tecnológico Nacional de México, Instituto Tecnológico José Mario Molina Pasquel y Henríquez, Unidad Académica de Lagos de Moreno. Cuerpo Académico: ITESLM-CA-2 Calidad, Productividad e Innovación Tecnológica en las Organizaciones. Contacto: [fernando.ricardez@lagos.tecmm.edu.mx](mailto:fernando.ricardez@lagos.tecmm.edu.mx)

2 Maestro en Manufactura Aditiva del Tecnológico Nacional de México, Instituto Tecnológico José Mario Molina Pasquel y Henríquez, Unidad Académica de Lagos de Moreno. Cuerpo Académico: ITESLM-CA-2 Calidad, Productividad e Innovación Tecnológica en las Organizaciones. Contacto: [hugo.noriega@lagos.tecmm.edu.mx](mailto:hugo.noriega@lagos.tecmm.edu.mx)

3 Ingeniero en Electrónica y Computación del Tecnológico Nacional de México, Instituto Tecnológico José Mario Molina Pasquel y Henríquez, Unidad Académica de Lagos de Moreno. Cuerpo Académico: C ITESLM-CA-2 Calidad, Productividad e Innovación Tecnológica en las Organizaciones. Contacto: [carlos.valenzuela@lagos.tecmm.edu.mx](mailto:carlos.valenzuela@lagos.tecmm.edu.mx)

4 Doctora en Manufactura Avanzada del Tecnológico Nacional de México, Instituto Tecnológico José Mario Molina Pasquel y Henríquez, Unidad Académica de Lagos de Moreno. Cuerpo Académico: C ITESLM-CA-2 Calidad, Productividad e Innovación Tecnológica en las Organizaciones. Contacto: [beatriz.esparza@lagos.tecmm.edu.mx](mailto:beatriz.esparza@lagos.tecmm.edu.mx)

## **Palabras Clave**

Automatización industrial, Manufactura avanzada, Redes de comunicación, Sistemas ciberfísicos, Tecnología educativa

## **Abstract**

Advanced Manufacturing has transformed production methods through digitization, automation and the interconnection of production systems. This work describes the integration of the development of an educational manufacturing cell based on technologies such as PLC programmable logic controllers, vision sensors, HMI Human-machine interfaces and the PROFINET communication protocol for this implementation, highlighting that the artificial vision sensor and PLC are not from the same manufacturer so a specific library and a system of physical relays were used to adapt the digital closing signals between the equipment. The development and integration of the cell succeeded in strengthening the technological competences of teachers and students involved in the implementation of an automated system connected in real time.

## **Keywords**

Advanced manufacturing, Cyber-physical systems, Communication network, Educational technology, Industrial automation.

## **Introducción**

La Manufactura Avanzada implica un cambio paradigmático en los métodos de producción, caracterizada por la convergencia de tecnologías digitales, automatización e interconectividad. En este contexto, una celda de manufactura es una unidad productiva automatizada que integra HMI, PLC, sensores, actuadores, todo interconectado mediante redes industriales, lo que implica que la formación tecnológica de los estudiantes universitarios deba alinearse a las nuevas competencias tecnológicas requeridas, lo que generó el problema de investigación porque la institución carecía de una celda experimental educativa de manufactura avanzada por restricción presupuestaria, el reto económico fue superado por la donación de equipo tecnológico actual y funcional de una empresa de alemana de autopartes con quien el cuerpo académico ha realizado proyectos de colaboración tecnológica desde 2016 los equipos donados se desensamblaron de maquinaria funcional que por cambios de línea la empresa desincorporo recientemente de sus líneas de manufactura lo que genero una adición al problema de investigación que consistió en la integración de la comunicación de componentes de diferente origen y marca en la celda de manufactura avanzada.

El objetivo educativo general: Que los estudiantes o investigadores puedan aprender y aplicar conceptos de manufactura, optimización de procesos y control de calidad en un entorno controlado. Para la generación de documentación didáctica de guías de prácticas de laboratorio de manufactura avanzada para estudiantes de ingeniería que les permita como ejemplo adquirir datos en tiempo real, en una celda de manufactura avanzada para la automatización de inspecciones mediante visión artificial y la interacción con el operador a través de una HMI para análisis de calidad y estadísticos.

El objetivo técnico general: Integrar una celda experimental de manufactura avanzada.

Objetivos específicos: a) Analizar las condiciones estructurales del desarrollo de la celda experimental para definir los requisitos de las redes. b) Analizar la configuración de tecnologías, equipos, hardware, softwares, periféricos e instalaciones del desarrollo de la celda experimental para definir los requisitos de los protocolos de comunicación. c) Definir las especificaciones de los protocolos de comunicación de las entidades de la celda por capas: física, de enlace de datos, capa de red, capa de transporte y la capa de aplicación.

Esta implementación se justifica por la promoción y la actuación en el aprendizaje práctico en entornos reales de manufactura inteligente de los estudiantes basado en la viabilidad económica a través de la serie de donaciones realizada por la empresa de autopartes y por la experiencia de los integrantes del cuerpo académico que tiene vinculación con la empresa de autopartes.

## **Marco Teórico**

La Manufactura Avanzada, conocida como Industria 4.0, ha cambiado de raíz los paradigmas tradicionales de la producción a través de la integración de tecnologías digitales, automatización inteligente y sistemas ciberfísicos. En el ámbito educativo, la implementación de una celda experimental de Manufactura Avanzada para prácticas de laboratorio representa un enfoque innovador y esencial para preparar a los estudiantes con las habilidades y competencias necesarias para enfrentar los desafíos de la industria moderna (Schwab, 2016; Moon, 2020).

### ***Fundamentos de la Manufactura Avanzada***

La Manufactura Avanzada se basa en la convergencia de diversas tecnologías innovadoras (Gilchrist, 2016; Mittal et al., 2017; Shabur, 2024; Wang et al., 2016):

- Internet de las Cosas (IoT): Permite la conectividad y comunicación entre objetos físicos empleando sensores y dispositivos electrónicos.
- Inteligencia Artificial (IA): Facilita la automatización inteligente, la toma de decisiones basada en datos en tiempo real y la optimización de procesos.
- Computación en la Nube: Proporciona almacenamiento de datos y capacidades de procesamiento que soportan aplicaciones en tiempo real y análisis de *big data*.
- Robótica Avanzada: Integra robots colaborativos y autónomos que trabajan junto a humanos en entornos de producción.

- **Fabricación Aditiva:** Utiliza técnicas como la impresión 3D para crear componentes y prototipos de manera personalizada y rápida.

### ***Importancia de la Celda Experimental en Educación***

La celda experimental de Manufactura Avanzada en el ámbito educativo tiene múltiples beneficios (Brecher & Schaefer, 2013; Edward, 2002):

- **Aprendizaje Práctico:** Impulsa a los estudiantes para aplicar teorías aprendidas en un entorno controlado, fortaleciendo su comprensión práctica de conceptos complejos.
- **Desarrollo de Habilidades Técnicas:** Capacita a los estudiantes en la programación de sistemas automatizados, el diseño de algoritmos de control y la integración de tecnologías emergentes permitiendo que su perfil de egreso este alineado a las necesidades actuales de la industria.
- **Preparación para el Mercado Laboral:** Los estudiantes adquieren habilidades altamente demandadas por la industria, como el manejo de datos en tiempo real, la resolución de problemas técnicos y la implementación de soluciones innovadoras.

### ***Diseño y Componentes de una Celda Experimental***

- **Selección de Tecnologías y Equipos.** Es crucial seleccionar tecnologías y equipos que cumplan con estándares reconocidos internacionalmente, como los definidos por el Instituto de Ingenieros Eléctricos y Electrónicos IEEE y otras normativas relevantes. Que incluye:
  - **Sensores Inteligentes y Actuadores:** Para la captura de datos y la ejecución de acciones físicas en respuesta a las instrucciones del sistema.
  - **Sistemas de Control y Automatización:** Que coordinen y optimicen los procesos de producción de manera eficiente y segura.
  - **Plataformas de Software:** Para la integración y gestión de datos, así como para el desarrollo de interfaces hombre-máquina (HMI) intuitivas.
- **Implementación de Sistemas Ciberfísicos.** Los sistemas ciberfísicos, que combinan elementos físicos con sistemas computacionales y redes de comunicación, son esenciales en la Manufactura Avanzada. En la celda experimental, estos sistemas permiten:
  - **Monitorización en Tiempo Real:** Para la supervisión continua de variables críticas de producción y condiciones operativas.
  - **Control Autónomo y Adaptativo:** Que ajusta automáticamente los parámetros de producción según las condiciones cambiantes del entorno.
- **Integración de Tecnologías Emergentes.** Además de los sistemas ciberfísicos, la celda experimental puede integrar tecnologías emergentes como (Tzimas et al., 2024):

- Realidad Aumentada (RA): Para la visualización de información contextual y guía en tiempo real durante las operaciones de manufactura.
- Inteligencia Artificial (IA) y Análisis de Datos: Para el análisis predictivo, la optimización de procesos y la detección de anomalías basadas en grandes volúmenes de datos.
- Integración horizontal y vertical. En el contexto del desarrollo de una celda experimental de manufactura avanzada, la integración horizontal y vertical son fundamentales para asegurar la eficiencia y la funcionalidad del sistema. La integración horizontal abarca la arquitectura de los sistemas de automatización, la especificación de instrumentación y equipos de control, la programación de PLCs, y la comunicación entre diferentes protocolos para asegurar la compatibilidad de equipos y periféricos, junto con una infraestructura de red robusta para garantizar la conectividad y la integración de dispositivos como robots. En telecomunicaciones, un protocolo se describe como un conjunto de reglas que permite la comunicación entre dos o más entidades para transmitir información, definiendo la sintaxis, la semántica y la sincronización que se utilizarán en dicha comunicación. Por otro lado, la integración vertical se enfoca en la conexión de redes y el flujo bidireccional de información a través de varios niveles jerárquicos: desde el nivel de campo, donde se interactúa con el proceso a través de sensores y actuadores, pasando por el control de sistemas automatizados, la supervisión y monitoreo a nivel local, hasta la planeación y gestión de la calidad del proceso.

### ***Beneficios Educativos y Profesionales***

- Formación Integral y Multidisciplinaria. La celda experimental no solo proporciona conocimientos técnicos específicos, sino que también promueve habilidades transversales como el trabajo en equipo, la comunicación efectiva y la capacidad de resolver problemas complejos.
- Innovación y Desarrollo Tecnológico. Los proyectos desarrollados en la celda experimental pueden generar avances significativos en áreas como la eficiencia energética, la sostenibilidad ambiental y la optimización de recursos en la producción (León-Bassantes et al., 2025; Hernández et al., 2020).

### ***Impacto en la Innovación y la Investigación***

La investigación realizada en la celda experimental puede contribuir a la asimilación tecnológica. (Mourtis et al., 2020):

- Nuevas Aplicaciones Tecnológicas: Explorar y desarrollar nuevas aplicaciones prácticas de tecnologías emergentes en la industria manufacturera.
- Optimización de Procesos: Mejorar la eficiencia y calidad de los procesos de producción mediante la implementación de soluciones innovadoras.

El desarrollo e integración de una celda experimental de Manufactura Avanzada para prácticas de laboratorio representa un paso estratégico hacia la preparación de una nueva generación de profesionales capacitados en tecnologías de vanguardia. Basado en un marco teórico robusto y en estándares internacionales, este enfoque no solo fortalece la educación técnica, sino que también impulsa la investigación y la innovación en el campo de la manufactura moderna, contribuyendo así al avance continuo de la industria global.

## Metodología

Se analizará la conformación de la celda experimental de los siguientes elementos:

- Sensores: Para adquirir datos.
- Elementos transmisores de información a las unidades de procesamiento.
- Unidades de procesamiento que toman decisiones de acuerdo con la información dada por los sensores y envían decisiones.
- Componentes que controlan las transmisiones y las instrucciones.
- Actuadores que ejecutan las órdenes.

Basados en estos requisitos se experimentarán y se seleccionarán protocolos de comunicación para la integración horizontal, y se desarrollará la integración vertical para que los datos se comuniquen con otras funciones como la evaluación de la calidad.

Los componentes de la integración horizontal a analizar son: Las arquitecturas integradas de los sistemas de automatización, la especificación de instrumentación y equipos de control, programación de PLC's la comunicación entre diferentes protocolos para lograr la compatibilidad de equipos y periféricos. La infraestructura de red para garantizar la conectividad. La integración de los dispositivos como robots.

La integración vertical se realiza por la conexión de redes y el flujo de la información es bidireccional en varios niveles jerárquicos. Campo: se interactúa con el proceso a través de sensores y actuadores. Control; Sistemas automatizados para máquinas y procesos. Supervisión: Monitoreo y operación a nivel local. Planeación: De la calidad de la gestión del proceso.

Los protocolos de comunicación entre las entidades analizarán: La Capa Física (*PHY Layer*) es el medio a través del cual viaja la información, ya sea aire, cable coaxial, cable de red, fibra óptica o cable eléctrico. Sobre esta capa se encuentra la Capa de Enlace de Datos (*Data Link Layer*), que proporciona direcciones necesarias para acceder al medio físico. La tercera capa es la Capa de Red (*Network Layer*), cuya función es asegurar que los datos lleguen de una dirección a otra, incluso si el origen y el destino no están conectados directamente. Sobre la Capa de Red se sitúa la Capa de Transporte (*Transport Layer*), que facilita el transporte de datos sin importar la red utilizada. En la parte superior se encuentra

la Capa de Aplicación (*Application/Data Layer*), que permite a las aplicaciones acceder a los servicios de las demás capas y define los protocolos para el intercambio de datos.

## **Resultados**

### ***Arquitectura del Sistema***

La celda se compone de los siguientes elementos:

- PLC SIEMENS como unidad de control central.
- Sensor de visión KEYENCE, encargado de realizar inspecciones de calidad.
- HMI KTP700 Basic, para visualización e interacción con el operador.
- Red PROFINET como protocolo de comunicación en tiempo real.

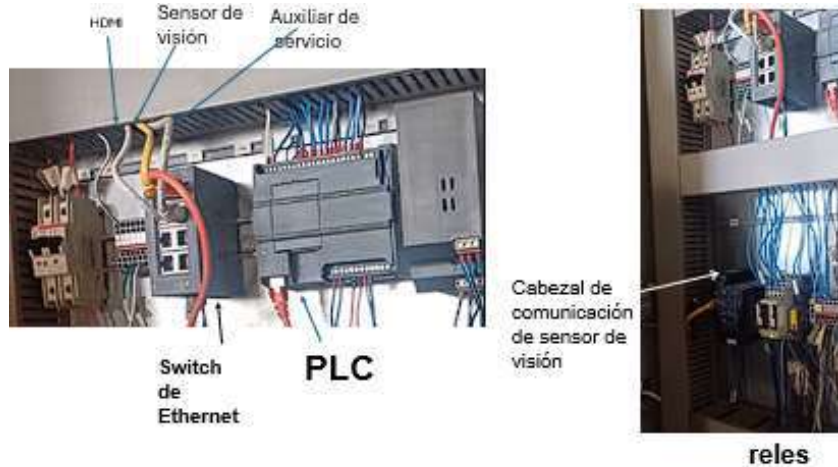
### ***Configuración y Comunicación***

Se utilizó el entorno de desarrollo TIA Portal (Portal de automatización totalmente integrado) para la programación del PLC y la configuración del sistema. La comunicación entre dispositivos se implementó mediante PROFINET RT, el cual garantiza una transmisión de datos determinista. Los pasos clave incluyeron:

- Asignación de áreas de memoria.
- Integración del archivo GSDML del sensor de visión. El archivo GSD es una forma estandarizada de descripción del de la información del dispositivo al describirse el archivo en formato XLM recibe el nombre GSDML.
- Configuración del direccionamiento de red y dispositivos.
- Dado que el sensor y el PLC son de marcas distintas (KEYENCE y SIEMENS, respectivamente), se utilizó una librería específica y un sistema de relevadores físicos para adaptar las señales digitales de cierre entre equipos.

La figura 1 muestra físicamente la arquitectura del sistema.

**Figura 1**  
*Arquitectura del sistema*



Fuente: propia

En la figura 2 se observa la pantalla HMI para una práctica desarrollada para un ensamble e inspección.

**Figura 2**  
*HMI de la estación de ensamble e inspección*

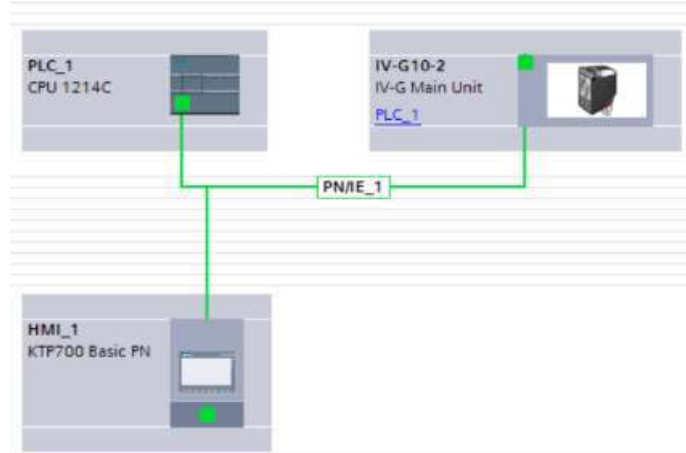


Fuente: propia.

La figura 3 despliega la configuración de la red ProfiNet del sensor de visión (IV-G10) de Keyence con el PLC CPU 1214C de Siemens y HMI ktp700 basic mediante el protocolo de ProfiNet.

**Figura 3**

Configuración del sensor de visión, PLC y HMI

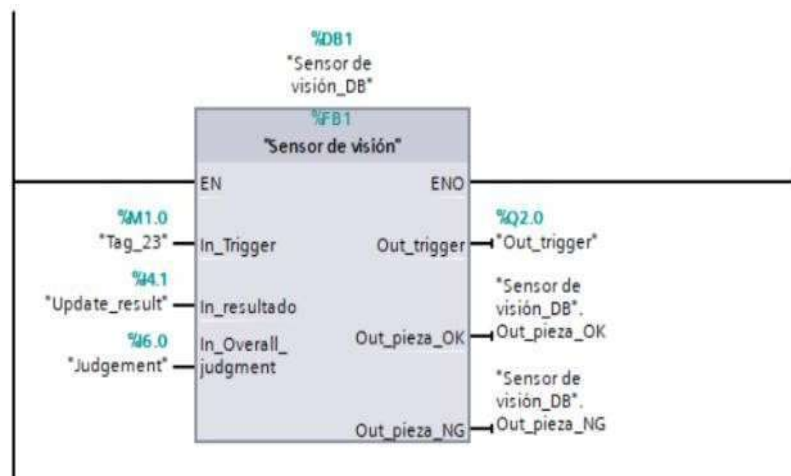


Fuente: propia.

La figura 4 muestra del programa principal, el primer segmento que hace el llamado de la función y en ella se definen cuáles serán las salidas y entradas tanto físicas como lógicas del programa, esta elección no es aleatoria, sino que responde a los datos que envía el sensor al conectarse en red con el PLC.

**Figura 4**

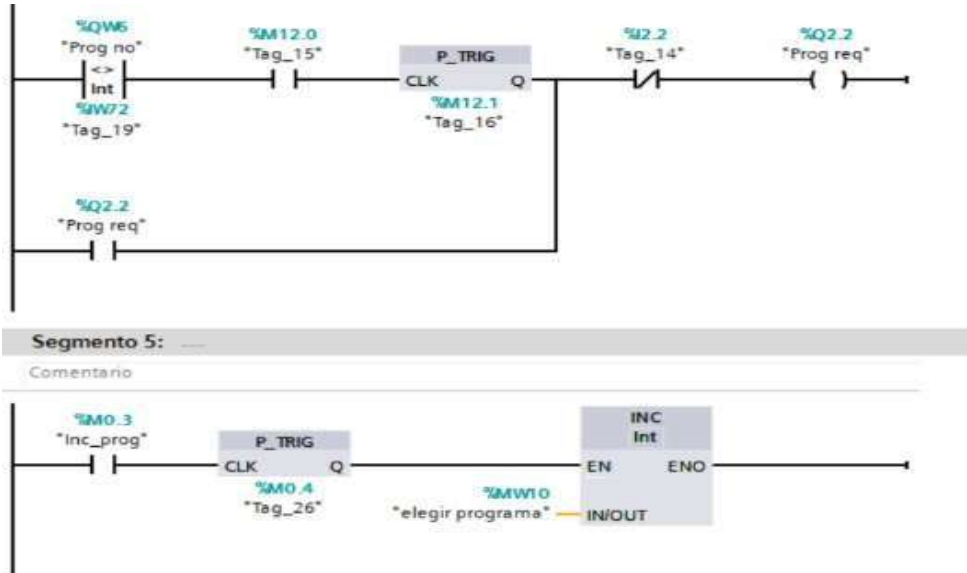
Segmento 1



Fuente propia

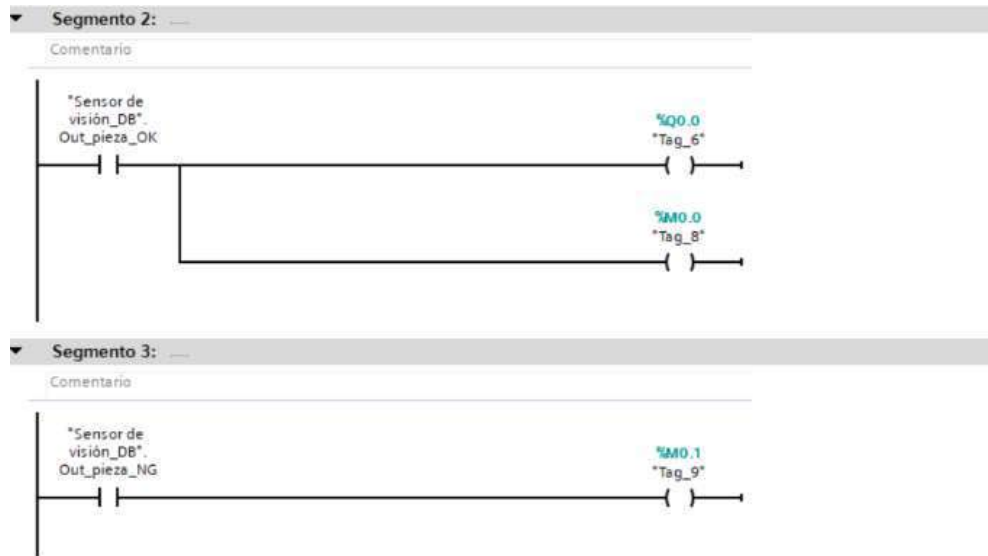
La figura 5 indica que los segmentos 2 y 3 son las salidas del juicio del sensor las cuales obedecen a las imágenes que se configuran en el sensor de visión.

**Figura 5**  
Salidas de juicio del sensor



Fuente Propia.

**Figura 6**  
Elección del programa a ejecutar



Fuente: Propia

La figura 6 muestra los segmentos involucrados en el programa a ejecutar.

## Conclusiones

La celda educativa funcionó exitosamente en tareas de inspección y ensamblaje automatizado. El sensor de visión realizó evaluaciones de calidad, mientras que el PLC coordinó las operaciones y la HMI permitió la interacción humana.

Las principales ventajas observadas fueron:

- Comunicación determinista y confiable gracias a PROFINET RT.
- Facilidad de configuración y diagnóstico en tiempo real mediante TIA Portal.
- Integración efectiva entre diferentes marcas mediante adaptación de señales.

El sistema permitió observar de forma didáctica la funcionalidad de una celda 4.0, desde la adquisición de datos hasta la toma de decisiones automatizadas.

La celda de manufactura implementada evidencia cómo la inteligencia y la autonomía de la Industria 4.0 emergen de la interacción entre sensores, redes de comunicación, unidades de procesamiento y actuadores. Se destacan dos niveles de integración:

- Integración horizontal, que comprende:
  - Arquitecturas de automatización
  - Programación de PLC
  - Compatibilidad entre protocolos
  - Infraestructura de red
- Integración vertical, orientada a:
  - Comunicación bidireccional entre niveles jerárquicos
  - Flujo de datos desde sensores hasta niveles de planificación y gestión

Finalmente, la participación de estudiantes y docentes fortaleció las competencias tecnológicas en automatización, programación, visión artificial y redes industriales, elementos clave en la formación de profesionales para la industria del futuro.

## Referencias

Brecher, C., & Schaefer, D. (Eds.). (2013). *Innovative Education and Training Concepts for Advanced Manufacturing*. Springer.


Edward, N. S. (2002). The Role of Laboratory Work in Engineering Education: Student and Staff Perceptions. *International Journal of Electrical Engineering Education*, 39(1), 11–19. <https://doi.org/10.7227/IJEEE.39.1.2>

- Gilchrist, A. (2016). *Industry 4.0: The Industrial Internet of Things*. Apress. <https://doi.org/10.1007/978-1-4842-2047-4>
- Hernandez-de-Menendez, M., Escobar, C.A., & Morales-Menendez, R. (2020). Engineering Education for Smart 4.0 Technology: A Review. *International Journal on Interactive Design and Manufacturing*, 14, 789–803. <https://doi.org/10.1007/s12008-020-00672-x>
- León-Bassantes, L. S., González-Sánchez, E., Sánchez-Granja, A. E., & Mosquera-Viejo, J. L. (2025). Industria 4.0 y sostenibilidad: Sinergias entre digitalización y responsabilidad ambiental. *MQRInvestigar*, 9(2), e761. <https://doi.org/10.56048/MQR20225.9.2.2025.e761>
- Mittal, S., Khan, M. A., Romero, D., & Wuest, T. (2017). Smart Manufacturing: Characteristics Technologies and Enabling Factors. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, 233 (5), 1342-1361. doi:[10.1177/0954405417736547](https://doi.org/10.1177/0954405417736547)
- Moon, I. (Ed.). (2020). *Smart Manufacturing: Applications and Case Studies*. Springer.
- Mourtzis, D., Angelopoulos, J., & Dimitrakopoulos, G. (2020). Design and development of a flexible manufacturing cell in the concept of learning factory paradigm for the education of generation 4.0 engineers. *Procedia Manufacturing*, 45, 361–366. <https://doi.org/10.1016/j.promfg.2020.04.035>
- Schwab, K. (2016). *The Fourth Industrial Revolution*. World Economic Forum. <https://www.weforum.org/about/the-fourth-industrial-revolution-by-klaus-schwab>
- Shabur, A. (2024). A comprehensive review on the impact of Industry 4.0 on the development of a sustainable environment. *Discov Sustain* 5, 97. <https://doi.org/10.1007/s43621-024-00290-7>
- Tzimas, E., Papazetis, G., Benardos, P., & Vosniakos, G.-C. (2024). Design and Development of a Flexible Manufacturing Cell Controller Using an Open-Source Communication Protocol for Interoperability. *Machines*, 12(8), 519. <https://doi.org/10.3390/machines12080519>
- Wang S, Wan J, Li D, Zhang C. (2016) Implementing Smart Factory of Industrie 4.0: An Outlook. *International Journal of Distributed Sensor Networks*, 12(1). doi:[10.1155/2016/3159805](https://doi.org/10.1155/2016/3159805)

# Análisis y Predicción de la Siniestralidad Vial en Ibagué Mediante Inteligencia Artificial: Un Enfoque Desde la Ingeniería Industrial

## Predicting Road Traffic Accidents in Ibagué with Artificial Intelligence: An Industrial Engineering Perspective

---

David Trujillo Durán<sup>1</sup> 

Fecha de recepción: 04-07-2025 | Fecha de aprobación: 19-09-2025

---

### Resumen

Este estudio aborda la siniestralidad vial en Ibagué, Colombia, mediante el uso de inteligencia artificial para identificar patrones y predecir accidentes. Analizar datos históricos (2020–2023) y desarrollar un modelo predictivo basado en el algoritmo Random Forest, complementado con un dashboard interactivo para visualizar dimensiones temporal, espacial y demográfica. Se procesaron bases de datos de siniestros viales mediante técnicas de análisis exploratorio (EDA) y balanceo de clases Synthetic Minority Over-Sampling Technique (SMOTE) para mitigar sesgos. El modelo alcanzó un F1-score de 0.71, demostrando capacidad moderada para clasificar causas de accidentes. Se identificó que los jóvenes hombres de 18 a 24 años son el grupo más vulnerable, con alta incidencia en motocicletas durante noches de fin de semana, especialmente en la comuna 1. La integración de IA y análisis multidimensional permite priorizar intervenciones en zonas críticas y horarios de riesgo. Estos hallazgos refuerzan la necesidad de políticas de seguridad vial focalizadas, como campañas educativas y optimización de recursos en áreas de alta accidentalidad.

### Palabras Clave

Siniestro Vial, Inteligencia Artificial, Predicción, Análisis de Datos, Ingeniería Industrial

### Abstract

This study addresses road safety in Ibagué, Colombia, using artificial intelligence to analyze and predict traffic accidents. To identify patterns and develop a predictive model based on the Random Forest algorithm, supported by an interactive dashboard for temporal, spatial, and demographic visualization. Historical accident data (2020–2023) was processed through exploratory data analysis (EDA) and class balancing Synthetic Minority Over-Sampling Technique (SMOTE) to reduce bias. The model achieved

---

1 Maestro Universidad de Ibagué, Ingeniero Industrial. Contacto: [dtrujillo1003@gmail.com](mailto:dtrujillo1003@gmail.com)

an F1-score of 0.71, showing moderate effectiveness in classifying accident causes. Young men aged 18–24 were identified as the most vulnerable group, with high motorcycle accident rates during weekend nights, particularly in Comuna 1. Integrating AI and multidimensional analysis enables prioritization of high-risk zones and timeframes. These insights highlight the need for targeted road safety policies, such as educational campaigns and resource optimization in critical areas.

### **Keywords**

Accident, Artificial Intelligence, Prediction, Data analysis, Industrial Engineering

La siniestralidad vial representa un desafío crítico para la salud pública, la seguridad y el desarrollo urbano sostenible en Ibagué, Colombia. En 2023, la ciudad reportó aproximadamente 17 muertes por cada 100.000 habitantes debido a incidentes viales, lo que la posiciona entre las más afectadas del país según El Colombiano (2023). La Agencia Nacional de Seguridad Vial (ANSV) (s.f.) también evidenció un índice de fatalidad municipal significativamente alto durante este periodo. En el departamento del Tolima, se registraron 69 muertes de mujeres por siniestros viales en 2022, ubicándolo en el sexto lugar nacional. Además, se identificó que estos hechos constituyen la principal causa de muerte violenta entre las mujeres en la región.

Como plantea Infocop (2024), los siniestros viales afectan no solo la integridad física, mental, social y profesional de las víctimas, sino también la estabilidad económica de sus familias y de las organizaciones en las que participan. Las consecuencias abarcan desde incapacidades y pérdida de movilidad hasta ansiedad, estrés postraumático y fallecimiento. Las empresas también enfrentan costos significativos derivados de estas situaciones, desde aseguradoras hasta entidades empleadoras.

En este contexto, resulta pertinente diferenciar entre los términos *accidente* y *siniestro vial*. Según Educación Bogotá (s. f.):

es el que permite vincular causas, consecuencias y responsabilidades de la persona en un evento de tránsito. Incluso, la palabra "siniestro" tiene un significado de catástrofe y se asocia con circunstancias dolorosas, como las lesiones o la pérdida de una vida, las cuales se pudieron haber prevenido en el marco de la responsabilidad y la autorregulación. En este sentido, en seguridad vial se opta por siniestro vial y no accidente vial, ya que éste es un suceso imprevisible e inevitable asociado al azar donde se exonera a la persona de toda responsabilidad.

Esto sugiere que los siniestros viales deben ser entendidos como eventos prevenibles, con responsabilidades identificables, y no como simples hechos fortuitos lo cual a título personal cambia completamente la perspectiva desde la que se aborda estos eventos ya que abordarlos como un evento fortuito exime responsabilidades e invisibiliza a las víctimas.

Frente a este contexto social, conceptual y estadístico, esta investigación parte de la hipótesis de que el uso integrado de inteligencia artificial (IA) y análisis multidimensional de datos históricos permite predecir causas comunes de siniestros y priorizar zonas de intervención. Su relevancia radica en ofrecer herramientas aplicables a la gestión pública en movilidad, promoviendo decisiones basadas en evidencia. Además, se alinea con los Objetivos de Desarrollo Sostenible (ODS) 8, 9, 10 y 11, y se inserta en el enfoque sistémico de la ingeniería industrial, orientado a la optimización de procesos urbanos.

El objetivo general es realizar un análisis multidimensional de los siniestros viales ocurridos en Ibagué entre 2020 y 2023, con énfasis en la identificación de zonas críticas, perfiles de actores viales y condiciones contextuales asociadas. Asimismo, se desarrollará un modelo predictivo basado en el algoritmo Random Forest para clasificar causas probables de siniestros, proporcionando insumos para el diseño de políticas públicas orientadas a la reducción de la accidentalidad vial.

En cuanto a la revisión documental el enfoque fue utilizar estudios previos para el análisis de la siniestralidad vial en Colombia e implementaciones de algoritmos de *machine learning* a nivel internacional, identificando avances metodológicos, vacíos operativos y limitaciones técnicas. Por consiguiente, esta exploración resulta necesaria para fundamentar el diseño metodológico adoptado en la presente investigación, que combina análisis exploratorio de datos con modelos de aprendizaje automático aplicados al contexto de una ciudad intermedia como Ibagué. Se examinan estudios con enfoques espaciales, multivariados y computacionales, con especial atención a su aplicabilidad en entornos con restricciones de datos y capacidades institucionales limitadas.

En el ámbito colombiano, diversos estudios han abordado la siniestralidad vial desde enfoques descriptivos, espaciales y temporales. Sin embargo, la mayoría presenta limitaciones en el uso de tecnologías emergentes como la inteligencia artificial. Uno de los trabajos más robustos es el de González (2024), quien desarrolló un análisis espacial y multiescalar de los siniestros viales ocurridos en Bogotá entre 2007 y 2022. El estudio se apoyó en el procesamiento de información proveniente de bases de datos oficiales (como el Observatorio de Movilidad y la Policía Metropolitana) y aplicó técnicas como el análisis de puntos calientes (Getis-Ord  $G_i^*$ ), análisis espacio-temporal, y regresión geográficamente ponderada multiescala (MGWR). Estas herramientas permitieron identificar zonas de concentración significativa de siniestros, así como modelar la relación entre características del entorno urbano (densidad poblacional, infraestructura vial, iluminación, uso del suelo) y la probabilidad de ocurrencia de incidentes.

Entre los hallazgos más relevantes se encuentra la identificación sistemática de clústeres persistentes de alta siniestralidad en corredores viales primarios y áreas con alta densidad de usos mixtos, así como la concentración temporal de los eventos en franjas horarias específicas. La desagregación por UPZ permitió generar mapas de riesgo que sustentan una priorización territorial de intervenciones públicas, como el mejoramiento de señalización, control de velocidad y rediseño urbano.

No obstante, este estudio fue desarrollado en un contexto metropolitano con un sistema robusto de información georreferenciada y una estructura institucional avanzada para el monitoreo del tránsito, lo cual favorece la aplicación de modelos estadísticos multivariados de alta resolución. En contraste, ciudades intermedias como Ibagué enfrentan limitaciones en la calidad, estandarización y cobertura de los registros, lo que restringe la aplicabilidad directa de metodologías como MGWR. Por esta razón, el presente estudio adopta un enfoque complementario, basado en algoritmos de aprendizaje automático como Random Forest y técnicas de balanceo de clases (SMOTE), que resultan más tolerantes a datos incompletos o desbalanceados y ofrecen capacidades predictivas útiles para priorizar zonas de intervención.

Además, mientras que el estudio de González (2022) enfatiza el análisis espacial estático con herramientas de SIG, esta investigación añade un componente interactivo mediante la implementación de *dashboards* dinámicos que permiten a los tomadores de decisiones explorar visualmente los datos, filtrar por dimensiones clave (actor, vehículo, tiempo, comuna) y utilizar los resultados como insumo directo para el diseño de estrategias preventivas. En conjunto, ambos enfoques son complementarios y refuerzan la necesidad de adaptar las metodologías al contexto de disponibilidad de datos, escala urbana y capacidades institucionales.

De manera complementaria, Ardila (2017) desarrolló un análisis espacial y temporal de la accidentalidad y mortalidad asociada al uso de motocicletas en el Área Metropolitana de Bucaramanga. Su estudio adoptó un enfoque mixto que combinó análisis estadístico, revisión documental y georreferenciación de casos con el fin de identificar los factores que explican la alta incidencia de siniestros entre motociclistas. La investigación utilizó datos de fuentes institucionales como Medicina Legal, el Instituto Nacional de Salud (INS) y la Dirección de Tránsito, abarcando variables sociodemográficas, técnicas, ambientales y comportamentales. El modelo teórico incorporado se basó en el enfoque de causalidad estructural, que considera tanto causas directas (conducción imprudente, exceso de velocidad, desobediencia de normas) como causas indirectas (déficit en infraestructura vial, débil control institucional, falta de formación vial).

En términos de propuestas, Ardila (2017) planteó una batería de acciones orientadas a mitigar el riesgo: campañas educativas para motociclistas, incremento de la vigilancia en puntos críticos, mejora de la infraestructura vial en zonas priorizadas, fortalecimiento de la capacidad institucional para el monitoreo de la movilidad, e integración de la gestión del riesgo vial en los planes de desarrollo municipal y metropolitano. No obstante, aunque el estudio generó insumos relevantes para la formulación de políticas públicas, no incorporó herramientas de análisis predictivo ni automatizado, lo que reduce su aplicabilidad en contextos donde se requiere proyección de escenarios futuros, estimación de riesgo en tiempo real o priorización dinámica de recursos. Tampoco se consideraron mecanismos de visualización interactiva que facilitaran la consulta de resultados por parte de actores no técnicos, lo cual puede limitar su impacto en procesos de toma de decisiones distribuidos o multisectoriales.

El contraste con el presente estudio es evidente en términos metodológicos y operativos. Mientras Ardila (2017) se centró en un análisis retrospectivo con fuerte anclaje institucional y énfasis cualitativo, esta investigación implementa un enfoque computacional automatizado basado en técnicas de aprendizaje automático (Random Forest), balanceo de clases (SMOTE), validación cruzada y visualización interactiva de datos con herramientas como Dash, Plotly y Folium. La segmentación por hipótesis causales, la modelación de perfiles de riesgo y la implementación de un *dashboard* web abierto constituyen avances sustantivos que permiten transformar los datos en decisiones operativas, especialmente en contextos de ciudades intermedias como Ibagué.

Finalmente, resulta pertinente resaltar que la investigación de Ardila (2017) proporciona un marco conceptual y estratégico valioso para entender los factores estructurales de la siniestralidad motociclista, pero requiere ser complementada con enfoques algorítmicos que permitan generar alertas tempranas, predicciones territorializadas y herramientas visuales adaptables al usuario. La combinación de ambos enfoques estructural y computacional representa una oportunidad concreta para avanzar hacia sistemas inteligentes de gestión del riesgo vial.

A nivel internacional, la aplicación de técnicas de *machine learning* ha mostrado resultados prometedores en el ámbito de la predicción de siniestros. Arakelyan (2023) desarrolló un modelo predictivo para estimar la probabilidad de ocurrencia de accidentes automovilísticos en Armenia, utilizando una base de datos del sector asegurador. El estudio empleó algoritmos como Regresión Logística, Random Forest, XGBoost y Redes Neuronales Artificiales (ANN), y evaluó su desempeño sobre una variable binaria (accidente/no accidente), en un conjunto de datos con un desbalance significativo (aproximadamente 6% de registros positivos frente a 94% negativos). En términos de rendimiento, los modelos basados en XGBoost y ANN lograron los mejores resultados, alcanzando un *F1-score* cercano a 0,80 y una métrica *Cohen's Kappa* por encima de 0,74, lo que indica una alta capacidad del modelo para discriminar eventos en escenarios con clases desproporcionadas.

El conjunto de variables utilizadas en el estudio incluyó factores como la edad del conductor, historial de reclamos, monto asegurado, tiempo de exposición y tipo de vehículo, todos ellos comparables conceptualmente con las dimensiones exploradas en el presente trabajo, como edad, género, localización y contexto del siniestro. Esta coincidencia confirma la relevancia operativa de integrar atributos sociodemográficos y de comportamiento histórico en modelos predictivos de riesgo vial. Asimismo, el uso de técnicas como *undersampling* y *oversampling* en Arakelyan (2023) se alinea con la aplicación de SMOTE en esta investigación, reafirmando la necesidad de estrategias de balanceo para evitar sesgos hacia las clases mayoritarias en contextos de siniestralidad vial.

No obstante, el enfoque de Arakelyan (2023) se limita a contextos privados del sector asegurador, centrado en la optimización del riesgo financiero y la segmentación de pólizas. No considera la dimensión territorial ni variables espaciales como latitud, longitud o comunas, lo que restringe su aplicabilidad directa a políticas públicas urbanas o sistemas de alerta georreferenciados. Tampoco se incorpora una interfaz de visualización para la interpretación institucional o ciudadana de los

resultados, lo que representa una barrera para su adopción por parte de gestores públicos sin formación técnica especializada. En contraste, el presente estudio amplía el alcance metodológico al incorporar variables espaciales y temporales, desarrollar una plataforma visual interactiva y vincular los resultados a escenarios reales de intervención territorial.

En síntesis, el trabajo de Arakelyan (2023) valida técnicamente el uso de algoritmos avanzados como Random Forest, XGBoost y ANN en la predicción de eventos infrecuentes con alto costo social, y sirve como referencia metodológica para el diseño de modelos similares en dominios públicos. Su principal aporte radica en la demostración empírica de que, incluso con variables estructuradas de carácter administrativo, es posible alcanzar una precisión aceptable en contextos de alta incertidumbre. Sin embargo, su alcance se ve limitado por la falta de articulación con políticas públicas, visualización geográfica o análisis multidimensional del territorio, aspectos que esta investigación busca integrar en el caso de Ibagué como ciudad intermedia.

A pesar de estos avances, la literatura evidencia vacíos significativos. En primer lugar, existen pocos estudios en Colombia que apliquen inteligencia artificial en contextos locales con datos incompletos o de baja calidad. En segundo lugar, la integración de modelos predictivos con dashboards interactivos sigue siendo escasa, lo que limita la exploración dinámica de los datos por parte de actores institucionales. Por último, persiste una débil conexión entre los hallazgos técnicos y la formulación efectiva de políticas públicas. Esta investigación busca contribuir a cerrar estas brechas mediante el desarrollo de una herramienta predictiva e interactiva adaptada a la ciudad de Ibagué, que no solo permita identificar zonas críticas y causas frecuentes de siniestros, sino que también funcione como insumo estratégico para la gestión pública de la movilidad urbana.

En conjunto, los estudios revisados muestran avances significativos en el análisis de la siniestralidad vial desde perspectivas geoespaciales, estructurales y predictivas. Sin embargo, persisten limitaciones comunes en la articulación entre análisis técnico y herramientas operativas para la gestión pública. La escasa integración de inteligencia artificial en entornos urbanos intermedios, la baja disponibilidad de dashboards interactivos institucionales, y la ausencia de sistemas de predicción adaptados a variables contextuales, justifican la necesidad de un enfoque metodológico híbrido como el desarrollado en este estudio. La combinación de modelos de *machine learning*, balanceo de clases y visualización interactiva representa un aporte concreto para avanzar en sistemas de análisis de siniestralidad más accesibles, aplicables y orientados a la toma de decisiones en ciudades con capacidades limitadas de infraestructura de datos.

En esta investigación se empleó inteligencia artificial (IA) como herramienta complementaria para la estructuración metodológica y el desarrollo computacional. El modelo seleccionado fue Qwen 2.5 Max, accesible a través de la plataforma Qwen Chat, elegido por su capacidad para procesar entradas extensas (hasta 10.000 palabras), realizar razonamiento contextual, mantener coherencia temática entre turnos conversacionales y acceder a información en línea de forma integrada. A la fecha de elaboración de este estudio, se consideraron alternativas como ChatGPT, Claude, Gemini y Copilot.

No obstante, muchas de estas plataformas restringen sus funciones avanzadas a planes de suscripción, lo cual limita su accesibilidad para investigaciones con restricciones presupuestarias. En este sentido, la elección de Qwen se basó en su equilibrio entre capacidad técnica, apertura de funciones, velocidad de respuesta y compatibilidad con flujos de trabajo iterativos en entornos de desarrollo.

Por lo anterior, desde un criterio práctico, se valoró especialmente su desempeño en tareas que requerían generación de código estructurado, depuración de errores, segmentación de funciones, reformulación de scripts y adaptación al entorno Python, sin necesidad de recurrir constantemente a documentación externa. El punto de partida fue la formulación del siguiente *prompt* de arranque:

Quiero realizar una investigación, el proyecto sería en el municipio de Ibagué, como recursos cuento con una base de datos recopilatoria de los siniestros entre 2020–2022 y 2023, cuento con Python y HTML, dame ideas de qué podríamos hacer.

Este *prompt* inicial no buscaba únicamente generar ideas temáticas, sino también evaluar el alcance funcional de la IA en términos de interpretación de contexto, diseño de flujo metodológico, y sugerencias de implementación computacional viables.

A lo largo del desarrollo, la IA fue utilizada como soporte continuo para acelerar la producción de código en Python y facilitar tareas de análisis estructurado, como la limpieza de datos, construcción de visualizaciones y elaboración del modelo predictivo. El flujo de trabajo se articuló en torno a tres scripts principales, cada uno asociado a una fase distinta del proceso:

1. Preprocesamiento y depuración de las bases de datos (detección y eliminación de registros nulos, tipificación de variables, y estandarización de formatos).
2. Visualización exploratoria con generación de gráficos y publicación de resultados en un entorno web.
3. Modelado predictivo mediante Random Forest con balanceo de clases y validación cruzada.

Cada bloque funcional se diseñó mediante *prompts* específicos que indicaban de manera explícita las entradas esperadas (input), las salidas deseadas (output), y los formatos requeridos para integración en el flujo de trabajo. Un ejemplo recurrente fue: “Necesito el código de una función con las siguientes entradas (edad, género, día) y quiero que me retorne un gráfico de barras con la frecuencia de siniestros.”

Este tipo de estructuras facilitó una interacción precisa y reproducible con el modelo, reduciendo errores de interpretación y favoreciendo la consistencia del código.

Durante la investigación, cuando surgieron errores durante la ejecución, se reutilizó el entorno conversacional de la IA como sistema de depuración, copiando directamente los mensajes arrojados

por la consola para obtener diagnósticos rápidos y soluciones específicas. En casos donde el modelo empezaba a generar respuestas erráticas o inconsistentes, se optó por reiniciar la conversación desde cero, aplicando una estrategia de iteración controlada para evitar acumulación de errores lógicos o alucinaciones. Esta práctica, aunque rudimentaria, permitió mantener trazabilidad sobre los cambios implementados y preservar el control en cada etapa crítica.

En conjunto, esta metodología asistida por IA no solo aceleró el tiempo de desarrollo, sino que permitió mantener una estructura lógica en la producción de código, facilitando la modularidad, reutilización y documentación interna del proceso. Si bien no se delegaron decisiones sustantivas ni análisis estadísticos a la IA, su papel como soporte computacional resultó fundamental para alcanzar un producto replicable, escalable y con valor operativo concreto. Este enfoque evidencia el potencial de los modelos de lenguaje como asistentes técnicos en investigaciones aplicadas, especialmente en contextos de recursos limitados donde se requiere maximizar eficiencia sin comprometer la calidad técnica.

Una vez estructurado el flujo de trabajo inicial con IA, se procedió al análisis exploratorio de datos (EDA), desarrollado a partir de las bases de datos suministradas por los líderes del Observatorio de Movilidad y Transporte de la Universidad de Ibagué (OMTU), Juliana Rojas y Juan Zuluaga quienes amablemente suministraron la información y brindaron su apoyo en el proceso de análisis e interpretación de la información allí contenida. Las bases de datos comprenden registros de siniestralidad vial ocurridos entre los años 2020 y 2023. Estas bases consolidan información procedente de reportes oficiales, integrando variables como tipo de siniestro, actores involucrados, ubicación aproximada que el equipo del observatorio se dio a la tarea de georreferenciar ya que la base original no cuenta con esa información, fecha, hora y características de los vehículos. Exceptuando la base correspondiente a 2021, debido a que, presentaba limitaciones de cobertura y ausencia de coordenadas geográficas, el resto de los años contaban con referenciación espacial aproximada (latitud y longitud), lo que permitió su incorporación en los análisis territoriales. Sin embargo, desde este punto ya se evidencia una oportunidad de mejora a nivel institucional en la medición de datos ya que se debería contar con bases georreferenciadas de siniestralidad al entender el valor que aporta conocer donde ocurren los siniestros viales.

El uso de información geográfica producto de la minuciosa labor del equipo del OMTU, aportó para la agregación por comunas y la generación de mapas de calor, clústeres y visualizaciones comparativas por zonas de la ciudad. Esta condición permitió realizar segmentaciones espaciales de la siniestralidad, identificar áreas con alta recurrencia de incidentes y establecer correlaciones entre la ocurrencia de siniestros y su localización urbana. Adicionalmente, la integración de la dimensión temporal (año, mes, día de la semana y hora) posibilitó el análisis multivariable de patrones espacio-temporales, facilitando la caracterización de franjas críticas según tipo de vehículo, actor vial o severidad del evento.

En esta etapa se aplicaron procesos sistemáticos de depuración, transformación y normalización de los datos. Esto incluyó la eliminación de registros incompletos, homogenización de formatos, recodificación de variables categóricas y ajuste de escalas temporales. A partir de esta base depurada, se organizaron los datos en estructuras específicas para cada una de las dimensiones analíticas definidas en el estudio: actores viales, demografía, características del siniestro, variables temporales, ubicación espacial y tipo de vehículo.

El proceso de limpieza incluyó la eliminación de registros incompletos, columnas sin información significativa y datos inconsistentes que comprometían la calidad analítica. Esta etapa fue fundamental para garantizar que los modelos y visualizaciones posteriores se basaran en información confiable y estructurada. La base consolidada original contaba con 110 columnas, muchas de las cuales eran redundantes, presentaban nomenclaturas irregulares o contenían vacíos sistemáticos. Tras la depuración, se conservaron 4.544 registros válidos, representativos del fenómeno de siniestralidad vial en Ibagué entre 2020 y 2023.

Una vez estabilizada la estructura del conjunto de datos, se definieron cinco dimensiones clave de análisis, orientadas a descomponer el fenómeno en componentes interpretables:

- Actores viales: permitiendo diferenciar entre conductores, peatones, acompañantes u otros tipos de participantes, considerando variables como edad y género.
- Demografía: enfocada exclusivamente en la caracterización sociodemográfica de los involucrados, facilitando análisis por grupos etarios, género y su distribución relativa.
- Características del siniestro: abarcando el tipo de evento (choque, volcamiento, atropello), su gravedad (con heridos o fallecidos) y la participación relativa de cada tipo de actor.
- Dimensión espacial y temporal: con base en la ubicación geográfica (latitud, longitud, comuna) y aspectos cronológicos (día de la semana, hora, mes, año) que permiten construir perfiles espacio-temporales del riesgo vial.
- Tipo de vehículo: clasificando las unidades involucradas en motocicletas, vehículos particulares, transporte público, bicicletas, entre otros.

Para cada dimensión, se diseñaron scripts personalizados en Python que automatizaron el procesamiento, transformación y representación gráfica de los datos. Estas herramientas facilitaron el desarrollo modular de los análisis y permitieron replicar la lógica exploratoria para distintas combinaciones de variables.

Inicialmente se consideró la exportación de resultados en formato estático (imágenes, PDF), pero dicha opción limitaba la capacidad de exploración cruzada de los datos. En respuesta, y aprovechando la infraestructura digital del OMTU, se optó por implementar un dashboard interactivo en línea que permite a los usuarios explorar los resultados.

Para este entorno interactivo se integraron herramientas de desarrollo específicas:

- Pandas y NumPy: para el tratamiento, filtrado y agregación eficiente de grandes volúmenes de datos.
- Plotly: para la generación de gráficos dinámicos (barras, líneas, histogramas) con capacidad de interacción directa.
- Folium: para la representación de los datos georreferenciados en mapas por comunas, permitiendo identificar zonas críticas.
- Dash: como marco principal para construir la arquitectura del dashboard, integrando gráficos, menús desplegados, selectores de tiempo y filtros dinámicos.
- Render: utilizado para el despliegue y publicación web del producto final, asegurando su disponibilidad pública desde el portal del OMTU.

La visualización interactiva, disponible en línea, refuerza el valor operativo del análisis exploratorio al facilitar una interpretación directa y segmentada de los datos. Este componente no solo convierte las bases de datos en información legible, sino que habilita la toma de decisiones basada en evidencia, sin requerir formación técnica avanzada por parte del usuario final. La arquitectura modular del sistema permite, además, su futura ampliación con nuevos años, variables o capas de análisis, manteniendo la lógica de navegación ya construida.

Este enfoque modular permitió una lectura más precisa de los datos y facilitó la generación de visualizaciones dirigidas, adaptadas a cada eje de análisis. En paralelo, se diseñaron scripts reutilizables en Python para automatizar la producción de gráficos y facilitar su posterior integración en un entorno web interactivo. Esta combinación de depuración estructurada, segmentación temática y visualización multicanal constituyó la base operativa sobre la cual se construyó el dashboard exploratorio final.

Para el desarrollo del modelo predictivo, se definió como variable objetivo la columna denominada Hipótesis, correspondiente al juicio preliminar emitido por el agente de tránsito sobre la causa probable del siniestro vial. Esta variable representaba una entrada textual no estandarizada, por lo que su uso requería un tratamiento previo de limpieza, normalización y clasificación.

Durante la revisión inicial se identificaron inconsistencias ortográficas, uso de sinónimos no uniformes, abreviaciones irregulares y presencia de múltiples criterios superpuestos en una misma entrada. Además, algunos registros estaban vacíos o eran irrelevantes para el modelo (por ejemplo: “se desconoce” o “en investigación”). Ante esto, se procedió a una consolidación semántica, que consistió en agrupar las hipótesis dispersas en un conjunto de 10 categorías unificadas, construidas con base en criterios conceptuales, frecuencia de aparición y claridad operacional:

1. Incumplimiento de prelación
2. Falta de atención a la vía y actores viales
3. No mantener distancia de seguridad

4. Desobediencia de señales y normas de tránsito
5. Uso incorrecto de carriles
6. Falta de precaución
7. Exceso de velocidad
8. Impericia en el manejo
9. Conductas de riesgo en maniobras específicas
10. Embriaguez

Este proceso de normalización categórica fue crítico para garantizar que el modelo pudiera operar sobre clases discretas, homogéneas y con suficiente volumen de datos por grupo. Cabe aclarar que la construcción de estas categorías fue de tipo inductivo, basada tanto en la recurrencia empírica como en el criterio técnico.

Por otra parte, se consideró la estructura relacional de los datos, donde un mismo siniestro podía registrar múltiples involucrados (actores viales) asociados a una misma hipótesis. Esta característica generaba el riesgo de duplicidad de registros al momento de entrenar el modelo, lo cual podía distorsionar el balance de clases o sobre representar ciertos perfiles. Para evitar este sesgo, se decidió reducir la unidad de análisis a un actor vial por siniestro, conservando únicamente el primer registro según el orden cronológico de aparición en la base.

En esta selección se priorizó la conservación de atributos como edad y género, que resultaban esenciales para la construcción de perfiles sociodemográficos y además evitaban un número muy grande de datos vacíos. Esta depuración permitió obtener un subconjunto final de 1.621 registros válidos, distribuidos de forma suficientemente representativa entre las categorías de hipótesis.

Este conjunto final sirvió como base estructurada para el entrenamiento del modelo predictivo, garantizando tanto la calidad del etiquetado como la unicidad de los casos, condiciones fundamentales para evitar sobreajuste y asegurar interpretabilidad en los resultados del modelo.

El algoritmo de clasificación seleccionado para el modelo predictivo fue Random Forest, una técnica ampliamente utilizada en problemas de clasificación multiclase por su capacidad para manejar relaciones no lineales, interacción entre variables y conjuntos de datos con ruido o valores atípicos. Este modelo opera mediante la construcción de múltiples árboles de decisión, entrenados sobre subconjuntos aleatorios del conjunto de datos original, lo cual permite reducir el sobreajuste y aumentar la generalización. Cada árbol genera una predicción independiente, y la decisión final se obtiene mediante un proceso de votación mayoritaria entre todos los árboles del bosque. En el presente caso, el objetivo fue predecir la hipótesis más probable asociada a un siniestro vial, a partir de atributos geoespaciales (latitud, longitud), temporales (día de la semana) y sociodemográficos (edad, género, comuna).

Cada árbol de decisión sigue una estructura jerárquica donde, en cada nodo, se selecciona la variable y el umbral de división que maximizan la ganancia de información o reducen la impureza. A través de este proceso recursivo, los árboles agrupan observaciones similares en función de los valores de sus atributos hasta llegar a una clasificación final, en este caso, una de las 10 hipótesis causales previamente normalizadas.

Para evaluar el rendimiento del modelo, se utilizaron métricas estándar de clasificación: precisión, *recall* y F1-score. La precisión indica el porcentaje de predicciones positivas correctas frente al total de predicciones positivas realizadas. El *recall*, en cambio, mide la capacidad del modelo para identificar correctamente los casos positivos reales. Finalmente, el F1-score representa la media armónica entre precisión y *recall* (Kundu, 2022), siendo especialmente útil en contextos donde las clases están desbalanceadas. Esta métrica, al ser menos sensible al desequilibrio que la precisión simple, permite una evaluación más robusta de modelos en escenarios reales de siniestralidad.

Dada la alta disparidad en la frecuencia de aparición de las hipótesis, se implementó la técnica de balanceo SMOTE (*Synthetic Minority Over-sampling Technique*). Esta técnica genera artificialmente nuevos registros para la clase minoritaria mediante interpolación entre observaciones cercanas, en lugar de simplemente duplicar registros existentes. El objetivo es reducir el sesgo del modelo hacia las clases mayoritarias, facilitando un entrenamiento más equilibrado y justo. SMOTE es particularmente efectivo cuando se desea evitar la pérdida de información asociada a técnicas de *undersampling*.

El algoritmo se apoya en el uso del algoritmo k-vecinos más cercanos (k-NN), el cual permite identificar las observaciones más próximas (según una métrica de distancia) para la generación de ejemplos sintéticos. El valor de k especifica cuántos vecinos se consideran en este proceso, lo cual influye directamente en la diversidad y representatividad de los nuevos registros generados (IBM, 2025). Este enfoque mejora la representación de las clases minoritarias sin alterar la distribución estadística global del conjunto de datos.

En cuanto al conjunto de variables utilizadas para el entrenamiento, se incluyeron: Latitud, Longitud, Edad, Género, Día y Comuna, seleccionadas por su relevancia tanto estadística como operacional, y su disponibilidad en los registros procesados. Estas variables aportan información suficiente para capturar patrones de riesgo asociados al contexto físico, temporal y social del siniestro.

Para evitar sobreajuste y asegurar una evaluación más precisa, se empleó la técnica de validación cruzada K-Folds con 10 particiones, mediante la cual el conjunto de datos se divide aleatoriamente en 10 subconjuntos del mismo tamaño. En cada iteración, 9 subconjuntos se utilizan para el entrenamiento y 1 para la validación, rotando sucesivamente hasta que cada subconjunto haya sido utilizado como conjunto de prueba. Este proceso, implementado con la librería Scikit-Learn de Python, maximiza el uso del conjunto de datos disponible, que por su volumen moderado no permitía una separación clásica de entrenamiento/prueba sin pérdida significativa de información.

Finalmente, los resultados se visualizaron mediante una matriz de confusión, que permite identificar cuántas veces el modelo acertó o erró en la predicción de cada clase. Se definió además un umbral de riesgo de 0.9, utilizado para filtrar solo aquellas predicciones cuya probabilidad asignada por el modelo superara dicho valor, garantizando que las salidas finales fueran altamente confiables y aptas para uso operativo.

En conjunto, estos procedimientos metodológicos permitieron construir una arquitectura robusta de análisis predictivo, capaz de operar con volúmenes intermedios de datos, lidiar con desequilibrios estructurales entre clases, y ofrecer resultados interpretables y que son coherentes con la realidad de la ciudad.

Desde una perspectiva personal, esta metodología responde a la necesidad de compatibilizar capacidades técnicas limitadas con una alta exigencia de resultados estructurados, reproducibles y útiles. No se partió de un entorno ideal, sino de un contexto realista en términos de disponibilidad de recursos, tiempo y acceso a herramientas avanzadas. Por ello, la elección de componentes como Python, Dash y la IA conversacional Qwen no obedece únicamente a su condición de libre acceso, sino a su capacidad comprobada de articular un flujo de trabajo controlado, flexible y escalable, sin renunciar a estándares de calidad técnica.

Frente a otras alternativas cerradas, opacas o excesivamente dependientes de licencias, esta arquitectura ofreció autonomía operativa, trazabilidad del código y posibilidad de adaptación modular. Cada componente del sistema fue elegido no por moda o disponibilidad, sino por su alineación funcional con los objetivos analíticos y su compatibilidad con entornos institucionales de bajo presupuesto. Esto permitió una integración racional de tecnologías, donde cada herramienta cumple una función específica en el ciclo de procesamiento, visualización o predicción.

El uso de IA como asistente funcional se justificó no como una delegación de criterio, sino como un acelerador técnico supervisado. En este proyecto, la IA se utilizó con la premisa de ser puntual, dirigida y bajo control constante. Las solicitudes se diseñaron con *prompts* estructurados, las respuestas se validaron línea por línea y cualquier desviación se corregía mediante reinicio o reformulación. Nunca se delegó el diseño analítico ni se aceptaron respuestas automáticas sin revisión crítica. Esta elección metodológica refleja un principio rector: la herramienta debe estar subordinada al diseño lógico del sistema, no al revés.

En cuanto al enfoque general, se priorizó un modelo reproducible, documentado y adaptable, antes que uno hipercomplejo, opaco o dependiente de estructuras externas no replicables. Este enfoque no pretende explotar al máximo la sofisticación algorítmica, sino garantizar que el modelo pueda ser comprendido, ajustado y aplicado a nivel general sin requerir capacidades técnicas avanzadas. Esto tiene implicaciones directas en la escalabilidad, la transferencia metodológica y la sostenibilidad operativa en el tiempo.

Adicionalmente, se tomó la decisión consciente de evitar la dependencia de librerías experimentales o entornos de alto mantenimiento, lo que refuerza la viabilidad de actualizar o extender el modelo mediante cambios incrementales, sin necesidad de rediseños completos. Este criterio de diseño es especialmente relevante en contextos públicos, donde la rotación de personal, la obsolescencia tecnológica o los cambios en prioridades institucionales son frecuentes.

En resumen, se construyó un modelo técnicamente sólido, operacionalmente viable y estratégicamente útil para el contexto local, con un diseño orientado a decisiones prácticas, adaptable a futuras extensiones y compatible con procesos institucionales reales.

Como producto del análisis exploratorio, se diseñó un *dashboard* interactivo disponible públicamente en <https://resumen-siniestros-ibague.onrender.com>. Allí se visualizan los principales patrones espacio-temporales, demográficos y modales identificados. Adicionalmente, en el repositorio <https://github.com/Dtrujillo-d/Dashboard.git> se encuentra el código para acceder al *dashboard* así como la base de datos implementada y demás elementos para replicar o mejorar el cuadro de control construido. A continuación, se sintetizan los hallazgos más relevantes según cada dimensión analizada.

#### Actores Viales

- 2.578 de los registros corresponde a conductores.
- Del total de conductores involucrados en siniestros, el 85,5% son hombres.
- La edad promedio de los actores viales varía según su rol, pero se concentra entre los 30 y 34 años.
- Los conductores con mayor participación en siniestros se ubican entre los 21 y 24 años.

#### Demografía

- El grupo etario con mayor frecuencia de siniestros corresponde a personas entre 21 y 23 años (842 registros).
- En ese rango, el 78% de los involucrados son hombres y el 22% mujeres
- La relación general entre hombres y mujeres involucrados en siniestros es de 4 a 1.

#### Temporalidad

- El mes con mayor número de siniestros fue marzo.
- Los viernes y sábados concentran la mayoría de los siniestros (707 y 711 casos, respectivamente).
- La hora con mayor número de siniestros fue las 12:00 AM.
- Entre semana se identificaron picos adicionales a las 16:00 y 18:00 horas.
- Los fines de semana, la siniestralidad incrementa significativamente entre las 23:00 y las 5:00 horas.

### Características del siniestro

- El tipo de siniestro más común es el choque, que representa el 82,3% de los registros.
- Marzo fue el mes con mayor número de heridos (299 casos).
- Junio presentó el mayor número de fallecimientos (18 casos).
- El 61,1% de los casos, los implicados fueron conductores.
- La proporción entre heridos y fallecidos es de 32 a 1.

### Vehículos Involucrados

- Las motocicletas fueron los vehículos más involucrados en siniestros.
- Febrero fue el mes con mayor siniestralidad en motocicletas (326 casos).
- Marzo registró el mayor número total de siniestros vehiculares (353).
- Los vehículos particulares lideran la participación en siniestros.
- Se reportaron 420 buses de servicio público implicados en siniestros viales.

### Modelo Predictivo

La variable objetivo que fue seleccionada fue la columna Hipótesis, categorizada en diez causas probables tras un proceso de depuración y normalización de texto. La distribución inicial de clases evidenció un fuerte desbalance entre las categorías, con diferencias de hasta 1.300 registros entre la clase más y menos frecuente. Por esta razón, se implementó la técnica SMOTE para equilibrar las clases. La categoría con menor representación (“Impericia en el manejo”) fue eliminada para evitar distorsión en la predicción.

El modelo de clasificación empleado fue Random Forest, con un esquema de validación cruzada mediante K-Folds ( $k=10$ ). Las variables predictoras utilizadas fueron: Latitud, Longitud, Edad, Género, Día y Comuna. El rendimiento promedio del modelo fue el siguiente:

- F1-score promedio: 0,71
- Precisión (precisión): 0,70
- Recall (exhaustividad): 0,72

Estos valores reflejan un desempeño equilibrado del modelo, con una capacidad aceptable para predecir correctamente las causas probables de siniestros viales a partir de condiciones espacio-temporales y sociodemográficas.

La matriz de confusión resultante evidenció una adecuada correspondencia entre las clases reales y las predichas, con un margen de error aceptable en la mayoría de las categorías. Adicionalmente, se construyó un gráfico de importancia de variables, donde se identificó que la Edad, seguida por la Latitud y Longitud, fueron los factores más influyentes en la predicción de las hipótesis.

Por último, se estableció un umbral de riesgo de 0,9, lo que permitió priorizar únicamente aquellas predicciones con alta certeza, mejorando así la confiabilidad del modelo para su uso como herramienta de apoyo en la toma de decisiones públicas.

A partir del análisis exploratorio y predictivo realizado en esta investigación, se derivan las siguientes conclusiones clave para la comprensión y gestión de la siniestralidad vial en la ciudad de Ibagué:

- **Riesgo elevado en jóvenes conductores:** El grupo etario entre 18 y 24 años concentra el mayor número de siniestros viales registrados en la base de datos. Esta concentración no solo se refleja en volumen absoluto, sino también en la proporción relativa dentro de los segmentos poblacionales analizados. La reiterada implicación de este grupo sugiere una exposición significativa al riesgo vial, probablemente asociada a factores como baja experiencia al volante, sobreconfianza, conductas impulsivas y limitado reconocimiento de situaciones de peligro. Este hallazgo refuerza la necesidad de establecer programas de intervención dirigidos específicamente a esta población, incluyendo formación obligatoria en gestión del riesgo vial, simuladores de conducción, campañas de sensibilización con enfoque conductual, y restricciones progresivas de licencia según nivel de experiencia. La estrategia debe ir más allá del enfoque informativo y priorizar acciones que modifiquen la percepción del riesgo y los hábitos concretos de conducción.
- **Mayor siniestralidad en horarios nocturnos de fin de semana:** Se identificó un aumento marcado de incidentes entre viernes y domingo en las franjas comprendidas entre las 23:00 y las 05:00 horas. Este patrón se desvía significativamente de la distribución esperada bajo condiciones de tráfico promedio y coincide con periodos de alta movilidad asociada al ocio nocturno. La coincidencia temporal sugiere la influencia de factores como consumo de alcohol, fatiga, condiciones de baja visibilidad o tránsito reducido, lo que incrementa las velocidades medias y reduce la percepción de riesgo. Estos hallazgos justifican la implementación de estrategias específicas para estas horas críticas, como refuerzo de los operativos de control de alcoholemia, mejoramiento del alumbrado público, campañas específicas de prevención nocturna y diseño de esquemas de transporte alternativo para el retorno seguro durante la madrugada.
- **Brecha de género en la participación vial:** El estudio evidenció una participación masculina desproporcionada en los siniestros viales, con una relación aproximada de 4 a 1 respecto a las mujeres. Esta brecha puede estar relacionada con mayores niveles de exposición, mayor frecuencia de conducción, diferencias en la elección del tipo de vehículo (por ejemplo, motocicleta) y patrones diferenciales de comportamiento vial. Aunque la estadística no permite inferencias causales directas, estos resultados abren la puerta a la formulación de hipótesis operativas y a la necesidad de incorporar un enfoque de género en las políticas de seguridad vial. Esto incluye la revisión de los programas de formación, la identificación de factores de riesgo específicos según género y la promoción de investigaciones complementarias que profundicen en los determinantes sociales y conductuales de esta disparidad.

- **Alta incidencia de motocicletas:** Las motocicletas representaron el tipo de vehículo con mayor número de siniestros reportados. Esta sobre-representación es consistente con estudios nacionales y evidencia un problema estructural de seguridad en este modo de transporte. La facilidad de acceso a motocicletas, su bajo costo relativo, la alta densidad de uso en ciudades intermedias y la exposición directa del conductor explican en parte esta incidencia. Frente a esto, resulta urgente diseñar e implementar estrategias específicas dirigidas a motociclistas: capacitaciones obligatorias antes de la expedición del pase, fortalecimiento de los controles sobre el uso de casco y elementos de protección, rediseño de la infraestructura para reducir los puntos críticos de fricción con otros vehículos, y políticas de regulación diferenciada por cilindrada o perfil del conductor.
- **Concentración espacial de los siniestros:** El análisis geográfico reveló una mayor concentración de siniestros en la comuna 1, lo que sugiere una combinación de factores de riesgo urbano que requieren intervención prioritaria. Esta concentración puede estar relacionada con características del entorno como alta densidad poblacional, mezcla de usos del suelo, congestión vial, falta de infraestructura peatonal segura o deterioro en las condiciones de señalización y visibilidad. La identificación de este patrón permite delimitar una zona crítica que debería ser considerada como polígono piloto para la implementación de acciones preventivas integrales. Esto incluye auditorías de seguridad vial, intervenciones de urbanismo táctico y control permanente con monitoreo institucional.
- **Cobertura institucional y posibles vacíos en el registro:** Finalmente, como se comentó en la metodología, se detectaron indicios de posibles inconsistencias o vacíos en la recolección de información georreferenciada y demográfica en ciertas zonas de la ciudad. Esta situación puede estar vinculada a una distribución desigual del recurso humano (agentes de tránsito), limitaciones logísticas en el levantamiento de información o ausencia de mecanismos estandarizados de reporte. Estos vacíos no solo afectan la representatividad del análisis, sino que pueden conducir a un subregistro sistemático que distorsiona la planeación pública. Se recomienda revisar la cobertura operativa de los agentes de tránsito, evaluar la calidad del proceso de captura de datos en campo, e impulsar el diseño de herramientas móviles unificadas para el reporte de siniestros, con validación automática de datos mínimos obligatorios y georreferenciación en tiempo real.

El modelo predictivo basado en el algoritmo Random Forest arrojó un F1-score promedio de 0,71, lo cual puede considerarse un desempeño aceptable y funcional, especialmente si se contextualiza en un entorno de datos limitados, con calidad heterogénea y un fuerte desbalance de clases. Esta métrica sugiere que el modelo logra una razonable capacidad de generalización y puede ser utilizado como herramienta de apoyo en procesos de clasificación preliminar. El desempeño no solo valida la aplicabilidad del enfoque, sino que también demuestra la pertinencia de utilizar algoritmos de ensamble robustos cuando se trabaja con fenómenos multivariados y distribuciones asimétricas, como ocurre en los registros de siniestralidad vial.

En términos de importancia relativa de las variables predictoras, se destaca que edad, latitud y longitud fueron los atributos con mayor peso dentro de la estructura del modelo. Esto confirma que los factores demográficos y geoespaciales son determinantes para predecir las causas probables de los siniestros, y sugiere que cualquier estrategia de intervención debe tener una base territorial clara, segmentada por grupo etario. El hecho de que el componente espacial tenga un peso comparable al sociodemográfico refuerza la necesidad de trabajar con mapas de riesgo dinámicos y con herramientas de análisis territorial continuo, lo cual abre la posibilidad de integrar capas adicionales como infraestructura, densidad vehicular o presencia de equipamientos urbanos.

La aplicación del método SMOTE (*Synthetic Minority Over-sampling Technique*) fue fundamental para corregir el sesgo del modelo hacia las clases mayoritarias. Este tipo de balanceo es especialmente útil cuando se busca conservar la mayor cantidad de registros posibles, evitando la pérdida de información inherente al undersampling. No obstante, como limitación metodológica, debe señalarse que el uso de muestras sintéticas puede alterar parcialmente la estructura interna de los datos y generar sobreajuste si no se controla adecuadamente. Por este motivo, se sugiere que futuras versiones del modelo incluyan una comparación sistemática con otras técnicas de balanceo.

Adicionalmente, se identificó que la variable objetivo, la columna *Hipótesis*, presenta un desafío metodológico no menor: se trata de una clasificación cualitativa de carácter subjetivo, emitida por agentes de tránsito sin un protocolo estandarizado de codificación. Esto introduce un nivel de variabilidad interobservador que puede afectar la consistencia del modelo. Para reducir este ruido, se sugiere avanzar en un proceso de normalización institucional de esta variable mediante una matriz de codificación técnica, orientada por criterios periciales, que garantice homogeneidad semántica y operativa. Alternativamente, podrían explorarse modelos jerárquicos de predicción que primero clasifiquen el tipo de siniestro y luego infieran la hipótesis probable, disminuyendo el impacto de la ambigüedad inicial.

Por otra parte, la incorporación de un dashboard interactivo construido con Dash, Plotly y Folium representa una mejora sustantiva frente a modelos tradicionales que limitan sus resultados a reportes estáticos o análisis descriptivos. Esta interfaz no solo democratiza el acceso a la información, sino que también habilita la exploración multidimensional de los resultados, filtrando según comuna, franja horaria, tipo de actor vial y tipo de vehículo. Su publicación dentro del portal del OMTU facilita su adopción institucional y convierte el modelo en un instrumento operativo, accesible incluso para usuarios sin conocimientos técnicos en programación o análisis de datos.

En conjunto, el sistema de análisis implementado permite estructurar una lógica operativa para la acción pública en movilidad segura, basada en cinco dimensiones clave:

1. Edad (identificación de poblaciones de riesgo);
2. Tiempo (segmentación por horarios críticos);

3. Geografía (zonificación territorial de alta siniestralidad)
4. Actor (tipificación por tipo de vehículo y perfil del involucrado);
5. Causa probable (clasificación predictiva basada en condiciones observadas).

A partir de los hallazgos y limitaciones del presente estudio, se delinean varias líneas de trabajo que no solo permiten consolidar y perfeccionar el modelo propuesto, sino también escalar su aplicabilidad hacia un sistema integral de apoyo a la toma de decisiones en seguridad vial. Estas líneas se estructuran en cinco dimensiones funcionales: mejora del modelo, expansión de datos, integración tecnológica, evaluación de impacto y replicabilidad territorial.

Con base en los hallazgos y limitaciones del presente estudio, se identifican varias líneas estratégicas que pueden fortalecer la continuidad y escalabilidad del análisis, tanto en términos metodológicos como operativos:

- Ampliación del modelo mediante variables contextuales críticas  
Actualmente, el modelo se apoya en variables geoespaciales y demográficas básicas. Sin embargo, existe un margen considerable para enriquecer su capacidad explicativa mediante la incorporación de nuevas variables que capturen mejor el contexto operativo de los siniestros. Entre ellas se destacan:
  1. Condiciones climáticas (lluvia, niebla, temperatura, humedad relativa).
  2. Estado y tipo de vía (pavimento, pendiente, número de carriles, presencia de señalización horizontal o vertical).
  3. Iluminación pública (tipo y cobertura del alumbrado en la zona).
  4. Flujo vehicular estimado por hora, extraído de sensores o estimaciones satelitales.
  5. Eventos extraordinarios (festividades, marchas, conciertos, accidentes previos).

La inclusión de estas variables no solo incrementaría la precisión del modelo, sino que permitiría generar perfiles de riesgo altamente localizados, útiles para la planificación táctica de intervenciones preventivas.

- Evolución hacia un sistema de predicción en tiempo real  
Una proyección natural de este trabajo es su integración con fuentes de datos en línea y sensores urbanos que operen en tiempo real. Esto habilitaría:
  - Modelos adaptativos capaces de actualizar sus predicciones con base en condiciones dinámicas del entorno.
  - Alertas tempranas para autoridades de tránsito, basadas en umbrales de riesgo derivados del modelo.
  - Integración con plataformas ciudadanas o aplicaciones móviles, permitiendo retroalimentación bidireccional (reportes de eventos, condiciones, incidentes).

Este componente requiere tanto desarrollo tecnológico como protocolos de gobernanza de datos, así como acuerdos interinstitucionales para la interoperabilidad de sistemas.

- **Profundización metodológica en modelos geoespaciales y de inteligencia artificial**  
El uso de Random Forest demostró ser efectivo para una primera aproximación. No obstante, se recomienda explorar modelos avanzados con mayor sensibilidad espacial y capacidad de modelado no lineal. Además, es pertinente comparar el rendimiento de técnicas de balanceo como SMOTE con métodos alternativos para evitar distorsiones por síntesis artificial.
- **Diseño y validación de intervenciones piloto orientadas por evidencia**  
La utilidad del modelo debe ser validada operativamente mediante su traducción en acciones piloto territorializadas, especialmente en las zonas y franjas horarias identificadas como críticas. Esto implica:
  - Diseñar intervenciones diferenciadas por comuna, tipo de vehículo y edad de los involucrados.
  - Establecer indicadores de impacto (reducción de siniestros, cambios en tipos de hipótesis, modificación de patrones temporales).
  - Aplicar diseño experimental para medir el efecto real de las políticas basadas en datos.Esta línea permitiría cerrar el ciclo de retroalimentación entre análisis predictivo, toma de decisiones y evaluación de resultados, aportando evidencia concreta sobre la eficacia de las acciones.
- **Replicabilidad nacional e institucionalización de la herramienta**  
El modelo desarrollado para Ibagué es escalable a otras ciudades intermedias de Colombia, bajo un enfoque modular. Para ello se sugiere:
  - Establecer criterios técnicos para la selección de municipios replicables (población, estructura vial, disponibilidad de datos).
  - Elaborar una guía metodológica para la implementación local del modelo.
  - Establecer alianzas entre observatorios de movilidad, gobiernos locales y universidades regionales para garantizar la adopción y sostenibilidad del sistema.En paralelo, se recomienda avanzar hacia la institucionalización de un sistema nacional de análisis predictivo vial, que estandarice metodologías, fomente la interoperabilidad de bases de datos, y ofrezca soporte técnico continuo a los entes territoriales.
- **Reforzamiento de capacidades institucionales y cultura de datos**  
La sostenibilidad del modelo requiere no solo tecnología, sino capacidades institucionales consolidadas en:
  - Recolección, depuración y estandarización de datos
  - Formación técnica en ciencia de datos aplicada a la movilidad.
  - Adopción de protocolos éticos para el uso de inteligencia artificial en decisiones públicas.

Se recomienda promover la creación de células de análisis predictivo en secretarías de movilidad, con perfiles mixtos (ingeniería, estadística, geografía, ciencia de datos), responsables de mantener y escalar los modelos en operación.

En síntesis, el trabajo futuro no debe centrarse únicamente en perfeccionar el modelo desde un punto de vista algorítmico, sino en consolidar un ecosistema completo de datos, capacidades institucionales, herramientas interoperables y cultura organizacional orientada a la acción basada en evidencia. Esta visión permitiría convertir modelos como el desarrollado en este estudio en pilares operativos de una política pública moderna de seguridad vial.

## Referencias

- Agencia Nacional de Seguridad Vial [ANSV]. (s. f.). Índice de Fatalidad Municipal. Agencia Nacional de Seguridad Vial. Recuperado 23 de abril de 2025, de <https://ansv.gov.co/es/observatorio/estad%C3%ADsticas/indice-de-fatalidad-municipal>
- González, F. H. (2022). *Análisis espacial de siniestros viales en Bogotá, Colombia durante el periodo 2007 – 2022* [Tesis de Maestría en Salud Pública]. Pontificia Universidad Javeriana. <https://repositorio.javeriana.edu.co/items/3bfe2949-59f9-479e-8ce2-460c1ed1d401>
- Ardila, J. A. (2018). *Análisis espacial y temporal de la accidentalidad y muerte generado por el uso de la motocicleta en el área Metropolitana de Bucaramanga* [Tesis de maestría]. Universidad de Santander, Bucaramanga. <https://repositorio.udes.edu.co/entities/publication/ee0eb9e8-cc8f-4906-994a-660732b7291e>
- Arakelyan, D. (2023). *Machine Learning-Powered Accident Prediction for the Automotive Insurance Industry*. American University of Armenia. [https://cse.aua.am/files/2023/12/Dawid\\_Arakelyan\\_Capstone.pdf](https://cse.aua.am/files/2023/12/Dawid_Arakelyan_Capstone.pdf)
- Educación Bogotá. (s. f.). Protocolo de atención de siniestros viales para establecimientos educativos del Distrito Capital. [https://www.educacionbogota.edu.co/portal\\_institucional/sites/default/files/inline-files/Anexo%2011%20protocolo\\_atencion\\_siniestros\\_viales\\_Establecimientos\\_edu.pdf](https://www.educacionbogota.edu.co/portal_institucional/sites/default/files/inline-files/Anexo%2011%20protocolo_atencion_siniestros_viales_Establecimientos_edu.pdf)
- El Colombiano. (2023, 21 junio). *Las 10 ciudades con más accidentes de tránsito en Colombia, ¿está la suya?* El Colombiano. <https://www.elcolombiano.com/colombia/cuales-son-las-diez-ciudades-de-colombia-con-mas-accidentes-de-transito-GM21789609>
- IBM. (2025, abril 1). *¿Qué es el algoritmo de k vecinos más cercanos (KNN)?* IBM. <https://www.ibm.com/mx-es/think/topics/knn>
- Infocop. (2024, 6 febrero). *Consecuencias psicológicas, físicas y socioeconómicas de los accidentes de tráfico*. Infocop. <https://www.infocop.es/consecuencias-psicologicas-fisicas-y-socioeconomicas-de-los-accidentes-de->

[trafico/#:~:text=Los%20accidentes%20de%20tr%C3%A1fico%20pueden,depresi%C3%B3n%20y%20trastorno%20de%20ansiedad.](#)

Kundu, R. (2022). *F1 Score in Machine Learning: Intro & Calculation*. V7. <https://www.v7labs.com/blog/f1-score-guide>