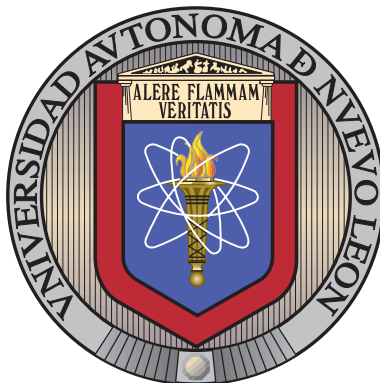


UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN

FACULTAD DE INGENIERÍA MECÁNICA Y ELÉCTRICA

SECRETARIA DE ESTUDIOS DE POSGRADO



DECISIÓN INTELIGENTE EN COMPRAS
INDIRECTAS: INTEGRACIÓN DE
CATEGORIZACIÓN Y ANÁLISIS PREDICTIVO

POR

JAVIER ALFONSO LÓPEZ PARRA

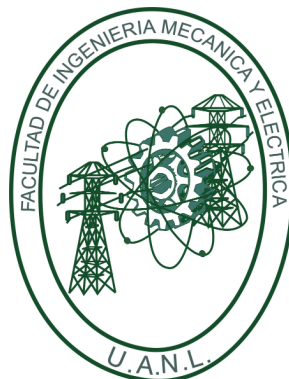
COMO REQUISITO PARA OBTENER EL GRADO DE
MAESTRÍA EN LOGÍSTICA Y CADENA DE SUMINISTRO

JUNIO 2026

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN

FACULTAD DE INGENIERÍA MECÁNICA Y ELÉCTRICA

SECRETARIA DE ESTUDIOS DE POSGRADO



DECISIÓN INTELIGENTE EN COMPRAS
INDIRECTAS: INTEGRACIÓN DE
CATEGORIZACIÓN Y ANÁLISIS PREDICTIVO

POR

JAVIER ALFONSO LÓPEZ PARRA

COMO REQUISITO PARA OBTENER EL GRADO DE
MAESTRÍA EN LOGÍSTICA Y CADENA DE SUMINISTRO

JUNIO 2026

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN
Facultad de Ingeniería Mecánica y Eléctrica
Posgrado

Los miembros del Comité de Evaluación de Tesis recomendamos que la Tesis Decisión Inteligente en compras indirectas: Integración de categorización y análisis predictivo, realizada por el estudiante Javier Alfonso López Parra, con número de matrícula 2225288, sea aceptada para su defensa como requisito parcial para obtener el grado de Maestría en Logística y Cadena de Suministro.

El Comité de Evaluación de Tesis

Dra. Jania Astrid Saucedo Martínez
Director

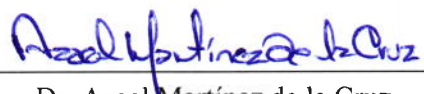
Dr. José Daniel Mosquera Artamonov
Co-director

Dr. Eduardo Valdés García
Revisor

Dr. Tomás Eloy Salas Fierro
Revisor

MLyCS Fanny Palma Vázquez
Revisor

Vo.Bo.



Dr. Azael Martínez de la Cruz
Secretario de Estudios de Posgrado



SUBDIRECCIÓN
ACADÉMICA

Institución 190001

Programa 642597

Acta Núm. 4701

Ciudad Universitaria, a 22/06/2026

*Dedico este trabajo, en primer lugar, a **Dios**, por darme la fortaleza necesaria para continuar cuando el camino se volvió incierto. A mi **padre**, que ya no está físicamente conmigo, pero cuya memoria sigue acompañándome. Recordarlo me dio fuerza durante este proceso y me ayudó a no olvidar de dónde vengo. A mi **madre**, por estar presente, apoyarme y compartir conmigo cada etapa de este recorrido. A mi hermano, **Andrés Parra**, por su respaldo y por motivarme a seguir adelante cuando más lo necesitaba. A **Fátima Medina**, por su amor, su paciencia y la confianza que depositó en mí. Gracias por acompañarme en los momentos difíciles y por recordarme que no tenía que enfrentar este proceso solo. A **José Daniel Mosquera**, gran amigo y compañero de camino desde mi llegada a México. Su orientación, lealtad y apoyo hicieron más llevaderos muchos momentos de adaptación y aprendizaje. A la Dra. **Janía Astrid Saucedo Martínez**, por brindarme la oportunidad de continuar mis estudios cuando las circunstancias económicas ponían en duda mi permanencia en la maestría. Su confianza fue decisiva para que pudiera seguir trabajando hasta llegar a este momento. A **Eduardo Valdez García** y a su familia, por recibirme con generosidad y apoyarme desde mi llegada a México. Su cercanía y sus enseñanzas han formado parte importante de esta etapa de mi vida. Este logro también pertenece a quienes estuvieron conmigo, me escucharon y me ayudaron a no detenerme. A todos ellos, gracias.*

ÍNDICE GENERAL

Agradecimientos	XI
Resumen	XII
1. Planteamiento del problema y alcance de la investigación	1
1.1. Contexto del estudio	1
1.2. Planteamiento del problema	3
1.3. Objetivo general	4
1.4. Objetivos específicos	5
1.5. Hipótesis	5
1.6. Justificación	5
1.7. Alcances y limitaciones del estudio	8
1.8. Estructura de la tesis	9
2. Marco teórico y antecedentes	11
2.1. Compras indirectas y decisiones basadas en información	11
2.2. Proceso de revisión de literatura	12

2.3. Abastecimiento estratégico y gestión por categorías	21
2.4. Gobernanza de datos y datos maestros en compras	22
2.5. Calidad del dato: dimensiones, estándares y evaluación	23
2.6. Metodologías para proyectos de datos: CRISP-DM y OSEMN	23
2.7. Abastecimiento estratégico y contribución al cumplimiento de las 7R de la logística	24
2.8. Analítica predictiva por categoría: utilidad, alcances y limitaciones . .	26
2.9. Fundamentos teóricos de los modelos de pronóstico considerados . . .	27
2.9.1. Modelo Prophet	27
2.9.2. Modelo ARIMA/SARIMA	28
2.9.3. Pronóstico basado en series de tiempo difusas	29
2.9.4. Modelo TimesFM	30
2.9.5. Métodos especializados para demanda intermitente	30
2.10. Métricas de evaluación y criterios de comparación de modelos	31
2.11. Casos organizacionales y aprendizajes para OC_QA	32
2.11.1. Unilever: simplificación operativa y gobierno del dato	33
2.11.2. Grupo Bimbo: datos accesibles como soporte para decisiones .	33
2.11.3. Lecciones transferibles para el caso OC_QA	34
2.12. Implicaciones para la investigación	34

3. Metodología

3.1. Enfoque del estudio	36
3.2. Diseño y unidad de análisis	37
3.3. Fuente de datos y preparación de OC_QA	38
3.4. Adaptación de CRISP-DM al caso	39
3.5. Selección de la categoría piloto	43
3.6. Variables y registro de las transformaciones	44
3.7. Protocolo de comparación	45
3.8. Aplicación de los modelos	46
3.8.1. Prophet	46
3.8.2. ARIMA/SARIMA	47
3.8.3. Series de tiempo difusas	47
3.8.4. TimesFM	48
3.8.5. SBA y SBJ para demanda intermitente	49
3.9. Herramientas de trabajo	50
3.10. Alcances metodológicos y posibilidad de reutilización	51
4. Resultados	53
4.1. Resultados del análisis exploratorio	54
4.1.1. Cobertura de la categorización	54
4.1.2. Estandarización de las unidades de medida	55
4.1.3. Comportamiento mensual por categoría	56

4.1.4. Concentración y consumo acumulado	57
4.1.5. Descripciones de artículos y servicios	59
4.1.6. Valores atípicos en la cantidad	60
4.2. Preparación de la serie LP_RE02	61
4.3. Resultados de Prophet	62
4.4. Resultados de ARIMA/SARIMA	64
4.5. Resultados de FTS-Chen	66
4.6. Resultados de TimesFM	68
4.7. Resultados de SBA y SBJ	69
4.8. Comparación integral de los modelos	71
5. Conclusiones	74
5.1. Conclusiones generales	74
5.2. Cumplimiento de los objetivos y revisión de la hipótesis	76
5.3. Aportaciones del trabajo	77
5.4. Implicaciones prácticas y limitaciones	78
5.5. Trabajo futuro	79
5.6. Reflexión final	80

ÍNDICE DE FIGURAS

3.1. Fases principales de la metodología CRISP-DM	41
3.2. Flujo metodológico aplicado al caso OC_QA	42
4.1. Cobertura de categorización antes y después de la preparación de OC_QA	54
4.2. Reducción de unidades de medida diferentes	55
4.3. Consumo mensual de cinco categorías sin valores superiores al per- centil 99	56
4.4. Distribución del consumo dentro de las cinco categorías seleccionadas	57
4.5. Consumo acumulado de cinco categorías	58
4.6. Diez productos o servicios con mayor cantidad acumulada	59
4.7. Distribución y valores atípicos de la variable Cantidad	60
4.8. Serie semanal y partición temporal de LP_RE02	61
4.9. Pronóstico de Prophet base para LP_RE02	62
4.10. Pronóstico de Prophet logístico para LP_RE02	63
4.11. Entrenamiento y prueba utilizados en ARIMA/SARIMA	64

4.12. Pronóstico de SARIMA para LP_RE02	65
4.13. Pronóstico FTS-Chen para LP_RE02	67
4.14. Pronóstico de TimesFM para LP_RE02	68
4.15. Serie semanal utilizada para caracterizar la demanda intermitente . .	70
4.16. Calibración de SBA y SBJ en las últimas seis semanas	70
4.17. Comparación de modelos en las 21 semanas de prueba	72

ÍNDICE DE TABLAS

2.1. Literatura relevante para la investigación	13
2.2. Criterios de selección de literatura	18
2.3. Fuentes consultadas	20
4.1. Resultados de las variantes de Prophet	63
4.2. Resultados de SARIMA para LP_RE02	65
4.3. Resultados de FTS-Chen para LP_RE02	67
4.4. Resultados de TimesFM para LP_RE02	69
4.5. Métricas utilizadas para calibrar SBA y SBJ	71
4.6. Comparación final de modelos para LP_RE02	72

AGRADECIMIENTOS

Quiero agradecer de corazón a mi comité de tesis por su apoyo y orientación a lo largo de este proceso. Sus consejos y paciencia han sido invaluable.

A mis profesores, gracias por compartir su conocimiento y por motivarme siempre a dar lo mejor de mí.

A la Facultad de Ingeniería Mecánica y Eléctrica (FIME) y a la Universidad Autónoma de Nuevo León (UANL), les agradezco por ofrecerme un entorno tan enriquecedor y lleno de oportunidades.

Asimismo, expreso mi profundo agradecimiento a la Secretaría de Coordinación Institucional de Humanidades, Tecnologías e Innovación (SECIHTI) por el invaluable apoyo brindado a través de la beca de manutención, la cual fue fundamental para continuar con mis estudios de posgrado.

Gracias a todos por su ayuda y por creer en mí.

RESUMEN

Javier Alfonso López Parra.

Candidato para obtener el grado de Maestría en Logística y Cadena de Suministro.

Universidad Autónoma de Nuevo León.

Facultad de Ingeniería Mecánica y Eléctrica.

Título del estudio: DECISIÓN INTELIGENTE EN COMPRAS INDIRECTAS: INTEGRACIÓN DE CATEGORIZACIÓN Y ANÁLISIS PREDICTIVO.

Número de páginas: 86.

OBJETIVOS Y MÉTODO DE ESTUDIO La investigación comenzó con el análisis de la base OC_QA, que registra las compras indirectas de una empresa alimentaria en México. La empresa contaba con información histórica, pero esta no se podía utilizar directamente. Las descripciones eran muy generales, había campos incompletos y las unidades de medida se registraban de formas distintas. La categorización existente tampoco permitía entender claramente qué se compraba ni cómo evolucionaba el consumo entre rubros.

El objetivo principal fue diseñar y validar una metodología que ayudara a ordenar esta información y convertirla en un apoyo real para la planeación del abastecimiento. Se utilizó CRISP DM como guía y sus etapas se adaptaron a las condiciones reales de la base. El proceso empezó con el diagnóstico de los registros y la identificación de sus limitaciones. Después se homologaron las variables clave, sobre todo las unidades de medida, y se construyó una categorización jerárquica.

Con la base analítica preparada, se seleccionó LP_RE02 como categoría piloto. La serie temporal se dividió en entrenamiento y prueba, con un horizonte común de 21 semanas. Esto permitió comparar bajo las mismas condiciones los modelos Prophet, ARIMA/SARIMA, series de tiempo difusas, TimesFM y los métodos SBA

y SBJ para demanda intermitente.

CONTRIBUCIONES Y CONCLUSIONES El estudio de OC_QA dejó claro que tener registros históricos no significa que los datos estén listos para analizarse. Fue necesario corregir descripciones, homologar unidades de medida y organizar los registros en categorías útiles para el área de compras. Sin esta preparación inicial, cualquier análisis posterior habría tenido poca confiabilidad.

El análisis exploratorio reveló comportamientos muy diferentes entre categorías. Algunas concentraban gran parte del consumo, mientras que otras mostraban semanas sin registros, picos aislados y valores atípicos. LP_RE02 presentó claramente este patrón intermitente. Por eso, la elección del modelo no se basó solo en las métricas de error. También se verificó si el pronóstico reflejaba de forma razonable el comportamiento real de la serie. En esta evaluación, ARIMA/SARIMA logró el mejor desempeño general para LP_RE02.

Este resultado corresponde únicamente a la categoría piloto y no se puede extender automáticamente al resto de las compras indirectas. Cada rubro puede presentar un comportamiento diferente. El principal valor de la investigación está en el procedimiento aplicado. Primero se revisa la calidad de los datos y se organizan los registros por categoría. Después se prepara la serie temporal y se comparan los modelos con el mismo horizonte de prueba.

Esta ruta puede repetirse en otras categorías de compras indirectas. En cada caso será necesario considerar la calidad de la información y el comportamiento del consumo antes de utilizar el pronóstico como apoyo para la planeación del abastecimiento.

Director

Codirector

Dra. Jania Astrid Saucedo Martínez Dr. José Daniel Mosquera Artamonov

CAPÍTULO 1

PLANTEAMIENTO DEL PROBLEMA Y ALCANCE DE LA INVESTIGACIÓN

1.1 CONTEXTO DEL ESTUDIO

Las compras indirectas reúnen los bienes y servicios que una empresa necesita para mantener su operación, aunque no formen parte directa del producto que vende. En este grupo pueden encontrarse materiales de mantenimiento, uniformes, productos de limpieza, papelería, servicios y otros apoyos utilizados por distintas áreas. La variedad de artículos y requerimientos hace que su seguimiento sea más complejo que el de las compras relacionadas directamente con la producción.

Delke *et al.* (2023) explican que las compras indirectas tienen características propias y no deben tratarse simplemente como una extensión de las compras directas. En ellas puede intervenir una gran cantidad de artículos, proveedores y solicitudes, cada una con necesidades y formas de consumo distintas. Cuando los registros no siguen criterios comunes, resulta difícil obtener una visión clara de lo que se compra.

El trabajo del área de compras también ha cambiado. Ya no consiste únicamente en recibir solicitudes, cotizar y emitir órdenes. Di Mauro *et al.* (2024) señalan que las compras pueden participar en decisiones de mayor alcance cuando conocen mejor sus

categorías, identifica patrones y cuenta con información para anticipar necesidades.

Para lograrlo, los datos deben estar organizados. Karttunen *et al.* (2023) encontraron que la transformación digital de compras necesita procesos claros e información que pueda utilizarse sin depender de correcciones constantes. Una nueva plataforma o una herramienta de análisis aporta poco cuando los registros siguen teniendo descripciones ambiguas, unidades distintas o categorías demasiado generales.

Bernardo *et al.* (2024) relacionan la gobernanza del dato con la definición de reglas, responsabilidades y criterios que ayuden a conservar información consistente. En compras, esto se refleja en acciones concretas: utilizar descripciones entendibles, mantener unidades de medida uniformes, completar los campos necesarios y asignar cada registro a una categoría que represente lo adquirido.

Estas dificultades aparecieron en la base OC_QA, utilizada por una empresa alimentaria en México para registrar sus compras indirectas. La empresa contaba con información histórica, pero no podía analizarla directamente. Había descripciones generales, campos incompletos y distintas formas de registrar las unidades de medida. La categorización existente tampoco ayudaba a reconocer con claridad qué se compraba ni cómo cambiaba el consumo entre rubros.

El problema inicial no era la ausencia de registros. La dificultad estaba en aprovechar los que ya existían. Antes de estudiar una categoría era necesario revisar descripciones, corregir diferencias y reorganizar la información.

Por esta razón, el trabajo comenzó con el diagnóstico de OC_QA. Después se homologaron las variables necesarias y se construyó una categorización jerárquica. Con la base preparada, se analizó el consumo y se eligió una categoría piloto para comparar distintos modelos de pronóstico.

1.2 PLANTEAMIENTO DEL PROBLEMA

La base OC_QA reúne registros históricos de compras indirectas. En su estado original, una parte de la información presentaba problemas que dificultaban su consulta. Algunos artículos tenían descripciones poco específicas, ciertos campos estaban incompletos y las unidades de medida se capturaban con criterios diferentes.

La clasificación disponible tampoco ofrecía suficiente detalle. Registros con características distintas podían aparecer dentro de una misma categoría, mientras que otros no contenían información suficiente para reconocer el tipo de compra. Esto dificultaba consolidar el consumo de manera consistente.

En la práctica, cada nuevo análisis exigía una revisión previa. Era necesario interpretar las descripciones, homologar unidades y decidir cómo agrupar los registros antes de comparar cantidades o periodos. El tiempo dedicado a este trabajo reducía la posibilidad de concentrarse en preguntas de mayor utilidad para el área, como identificar consumos recurrentes, cambios importantes o categorías que necesitaban mayor seguimiento.

Fontes *et al.* (2025) explican que la gestión por categorías ayuda a tomar decisiones de acuerdo con las características de cada grupo de bienes o servicios. Una categoría no debe funcionar solamente como una etiqueta administrativa. También debe reunir registros que puedan analizarse bajo criterios similares.

Guida *et al.* (2025) señalan que la clasificación del gasto ayuda a convertir transacciones dispersas en información útil para compras. En OC_QA, los campos originales no eran suficientes para lograr esta organización, por lo que fue necesario construir una clasificación adicional.

La falta de una estructura clara también afectaba la preparación de las series temporales. Para estudiar cómo cambia el consumo con el paso del tiempo, primero se necesita identificar de forma consistente los registros que pertenecen a una misma

categoría. De lo contrario, el comportamiento observado puede quedar alterado por errores de captura o agrupaciones poco claras.

La planeación de compras indirectas también se relaciona con decisiones logísticas. Swamidass (2000, p. 684) presenta las siete condiciones de la logística: producto, cantidad, condición, lugar, tiempo, cliente y costo correctos. En esta investigación se retoman principalmente la cantidad, el tiempo y el costo, ya que estas decisiones necesitan información previa sobre el consumo.

Cuando no se conoce con claridad qué se compró, cuánto se consumió o con qué frecuencia apareció una necesidad, anticipar requerimientos se vuelve más difícil. El pronóstico puede servir como referencia, pero primero debe construirse una serie que represente de forma razonable el comportamiento de la categoría.

El problema que guía esta investigación es que la base OC_QA no contaba con descripciones, unidades de medida y categorías suficientemente consistentes para dar seguimiento al consumo por rubro y preparar pronósticos que apoyaran la planeación de las compras indirectas.

1.3 OBJETIVO GENERAL

Diseñar y validar una metodología para compras indirectas que organice la información de la base OC_QA mediante la estandarización de sus variables y una categorización jerárquica, y que compare modelos de pronóstico en una categoría piloto para mejorar el seguimiento por categoría y apoyar la planeación del abastecimiento.

1.4 OBJETIVOS ESPECÍFICOS

1. Diagnosticar y estandarizar la base OC_QA, identificando los problemas de descripción, unidades de medida, campos incompletos y clasificación que afectan su uso para el análisis.
2. Diseñar e implementar una categorización jerárquica que organice las compras indirectas por rubro y facilite el análisis del consumo.
3. Comparar el desempeño de Prophet, ARIMA/SARIMA, series de tiempo difusas, TimesFM, SBA y SBJ en la categoría piloto LP_RE02, utilizando un horizonte común de 21 semanas y los mismos criterios de evaluación.

1.5 HIPÓTESIS

Si la base OC_QA se estandariza y sus registros se organizan mediante una categorización jerárquica, entonces mejorará la trazabilidad por categoría y será posible comparar modelos de pronóstico en una categoría piloto como apoyo para la planeación de compras indirectas.

1.6 JUSTIFICACIÓN

La investigación surgió de un problema observado directamente en OC_QA. La empresa tenía información histórica, pero utilizarla requería un trabajo previo de revisión y limpieza. Las diferencias entre descripciones, unidades de medida y categorías impedían consultar los registros de forma inmediata.

Organizar esta información tiene una utilidad práctica. Una base más clara ayuda a reconocer qué categorías concentran mayor consumo, cuáles presentan

cambios importantes y en qué casos existe suficiente historial para estudiar su comportamiento. También reduce el trabajo repetido antes de preparar un reporte.

Di Mauro *et al.* (2024) sostienen que el área de compras puede participar en decisiones de mayor alcance cuando deja de concentrarse únicamente en la ejecución de solicitudes. Para asumir esa función necesita información que ayude a entender qué ocurre en cada categoría y no solo una lista de órdenes de compra.

En el caso estudiado, esa información no podía obtenerse directamente de la base original. Antes de comparar periodos o artículos era necesario corregir diferencias. Por eso, la primera parte del trabajo se concentró en ordenar los registros.

Karttunen *et al.* (2023) explican que las herramientas digitales producen mejores resultados cuando los procesos y los datos ya cuentan con cierta organización. Aplicar modelos sobre OC_QA sin revisar primero su estructura habría trasladado los mismos problemas a la etapa de análisis.

La categorización se incorporó para dar un orden común a los registros. Fontes *et al.* (2025) describen la gestión por categorías como una forma de adaptar las decisiones de compra a grupos con características propias. Esta idea ayudó a separar registros que, aunque pertenecían a compras indirectas, no debían analizarse de la misma manera.

Guida *et al.* (2025) explican que la clasificación del gasto facilita la identificación de patrones dentro de una gran cantidad de transacciones. En esta tesis, la clasificación se utilizó para organizar el consumo y preparar la información antes de construir las series temporales.

La etapa de pronóstico comenzó cuando los registros ya estaban organizados y fue posible construir una serie temporal para LP_RE02. Los modelos no se eligieron bajo la idea de que uno de ellos sería adecuado para cualquier categoría. Se buscó reunir alternativas con formas distintas de representar el comportamiento de la demanda y evaluarlas con las mismas 21 semanas de prueba.

Prophet se incluyó porque trabaja con tendencias y cambios temporales mediante una estructura que puede interpretarse con relativa facilidad, como explican Taylor y Letham (2018). ARIMA/SARIMA ofrecía una lógica diferente. Hyndman y Athanasopoulos (2021) señalan que esta familia de modelos utiliza la relación entre las observaciones anteriores para estimar el comportamiento futuro de una serie.

También se incorporaron series de tiempo difusas. Song y Chissom (1993) propusieron este enfoque para representar información temporal mediante relaciones difusas. Chen (1996) presentó después un procedimiento más sencillo para construir y utilizar esas relaciones en el pronóstico.

TimesFM se agregó como una alternativa más reciente. Das *et al.* (2024) presentan este modelo como una propuesta entrenada con grandes colecciones de series temporales y diseñada para generar pronósticos sin requerir un ajuste completo para cada nueva serie.

Durante la preparación de LP_RE02 se encontraron semanas sin consumo y picos aislados. Este comportamiento llevó a incluir métodos diseñados para demandas que no aparecen de forma continua. Croston (1972) propuso uno de los primeros procedimientos para separar el tamaño de la demanda y el tiempo entre ocurrencias. A partir de esta lógica se evaluaron los ajustes SBA y SBJ.

La presencia de periodos sin movimiento no es exclusiva del caso estudiado. Ghobbar y Friend (2002) identificaron este comportamiento en la demanda de refacciones aeronáuticas. Gutierrez *et al.* (2008) explican que las series con alta irregularidad necesitan un tratamiento diferente al utilizado para demandas más estables. Estos antecedentes ayudaron a justificar una comparación que incluyera modelos generales y métodos para demanda intermitente.

El trabajo siguió un orden definido. Primero se revisaron los registros y se corrigieron las diferencias que afectaban su análisis. Después se organizaron las compras por categoría y se preparó la serie temporal. La comparación de modelos comenzó cuando estas etapas estuvieron terminadas.

Para la empresa, este procedimiento ofrece una forma más clara de trabajar con su información histórica y evita comenzar cada análisis desde cero. En el ámbito académico, el caso reúne la preparación de datos, la categorización de compras indirectas y el pronóstico dentro de un mismo proceso aplicado.

1.7 ALCANCES Y LIMITACIONES DEL ESTUDIO

El estudio se concentra en las compras indirectas registradas en OC_QA. La base pertenece a una empresa alimentaria ubicada en México y contiene información histórica de artículos, servicios, proveedores, unidades de medida y categorías.

El trabajo incluye el diagnóstico de la base original, la homologación de variables, la construcción de una categorización jerárquica y el análisis exploratorio del consumo. También comprende la preparación de una serie temporal para una categoría piloto.

LP_RE02 fue seleccionada para la etapa de pronóstico. La serie se dividió en datos de entrenamiento y prueba. Se utilizó un horizonte de 21 semanas para evaluar Prophet, ARIMA/SARIMA, series de tiempo difusas, TimesFM, SBA y SBJ bajo las mismas condiciones.

La tesis no plantea un rediseño completo del proceso de compras. Tampoco contempla la instalación de la metodología dentro del sistema utilizado por la empresa. El trabajo se concentra en preparar la información, analizarla y probar el procedimiento con una categoría.

Una de las limitaciones es que el análisis predictivo se realizó únicamente con LP_RE02. Los resultados obtenidos no deben trasladarse automáticamente a otras categorías. Algunas pueden tener consumo continuo, mientras que otras pueden presentar movimientos aislados o largos periodos sin registros.

Otra limitación proviene de la calidad de la información original. Aunque se

corrigieron diferencias y se homologaron variables, no fue posible recuperar datos que nunca se capturaron. Una descripción demasiado general, por ejemplo, puede seguir limitando la identificación exacta del artículo.

La comparación incluyó modelos clásicos, difusos, especializados para demanda intermitente y una propuesta reciente basada en grandes colecciones de series temporales. Aun así, existen otras alternativas que no formaron parte del estudio.

El desempeño de un modelo tampoco depende únicamente de su estructura matemática. La cantidad de datos disponibles, los cambios en el consumo y la presencia de semanas en cero pueden modificar los resultados. El modelo que funcionó mejor para LP_RE02 no debe considerarse como una solución general para todas las compras indirectas.

La aplicación continua del procedimiento requerirá trabajo adicional dentro de la empresa. Será necesario definir quién actualizará la base, cómo se clasificarán los nuevos artículos y con qué frecuencia se revisarán los resultados. Esa etapa no fue evaluada en esta tesis.

1.8 ESTRUCTURA DE LA TESIS

La tesis está organizada en cinco capítulos. El primer capítulo presenta el contexto del estudio, el problema encontrado en OC_QA, los objetivos, la hipótesis, la justificación y los límites de la investigación.

El segundo capítulo reúne los conceptos y antecedentes utilizados para estudiar el problema. Incluye compras indirectas, abastecimiento estratégico, gestión por categorías, calidad de los datos, gobernanza, CRISP-DM y los modelos de pronóstico considerados.

El tercer capítulo explica cómo se realizó el trabajo. Se describe la fuente de datos, la preparación de la base, la categorización jerárquica, la selección de LP_RE02

y el procedimiento utilizado para evaluar los modelos.

El cuarto capítulo presenta los hallazgos del análisis exploratorio y los resultados de los pronósticos. Cada modelo se revisa por separado y después se comparan sus métricas y su comportamiento frente a los valores reales.

El quinto capítulo contiene las conclusiones, la revisión de los objetivos y la hipótesis, las aportaciones del estudio, las implicaciones para compras indirectas y las posibles líneas de trabajo posteriores.

CAPÍTULO 2

MARCO TEÓRICO Y ANTECEDENTES

2.1 COMPRAS INDIRECTAS Y DECISIONES BASADAS EN INFORMACIÓN

Compras no siempre ha tenido un lugar claro dentro de las decisiones estratégicas de una empresa. Durante mucho tiempo su trabajo consistió en recibir solicitudes, buscar proveedores y emitir órdenes. Di Mauro *et al.* (2024) documentan que esta situación está cambiando: la gestión de compras y abastecimiento, referida en la literatura como *purchasing and supply management*, puede participar en decisiones de mayor alcance cuando cuenta con información para hacerlo.

Esta tesis parte de esa idea. Las compras indirectas no solo implican ejecutar solicitudes; también generan registros que, bien organizados, pueden usarse para planear, priorizar y anticipar necesidades. El problema es que esa organización no siempre existe.

Karttunen *et al.* (2023) encontraron que la transformación digital de compras no depende únicamente de adoptar nuevas herramientas. Requiere procesos claros y datos que puedan utilizarse. Delke *et al.* (2023) agregan que las compras indirectas tienen características propias y no deben tratarse como una versión simplificada de

las compras directas.

Cuando la información está organizada por categoría y los registros son trazables, es posible hacer algo más que consultar el historial. Fontes *et al.* (2025) muestran que la gestión por categorías da mayor coherencia a las decisiones de compra. Guida *et al.* (2025) explican que la clasificación del gasto, o *spend classification*, ayuda a convertir transacciones dispersas en información útil para el análisis.

Esta tesis estudia esas tareas como partes de un mismo proceso. Primero se revisan los datos, después se organizan por categoría y, cuando la información lo permite, se construye una serie para comparar pronósticos. El orden importa porque cada etapa depende de la anterior.

2.2 PROCESO DE REVISIÓN DE LITERATURA

La revisión se organizó alrededor de los temas que estructuran el problema: compras indirectas, abastecimiento estratégico, gobernanza de datos, calidad del dato, clasificación del gasto, metodologías para proyectos de datos y modelos de pronóstico aplicables a series temporales por categoría.

Los documentos se seleccionaron con criterios explícitos: pertinencia temática, rigor metodológico, calidad editorial, actualidad y aplicabilidad al caso OC_QA. Cada fuente se registró en una matriz elaborada en Excel con tema central, método, contexto, hallazgos, limitaciones y relevancia para la investigación. El objetivo fue construir una síntesis crítica, no solo una lista de referencias.

La matriz también permitió registrar por qué se conservó cada documento y en qué parte de la tesis sería utilizado. De esta manera, la revisión mantuvo un vínculo directo con las decisiones tomadas durante el estudio.

La Tabla 2.1 resume la literatura seleccionada y su aporte específico para esta tesis.

TABLA 2.1: Literatura relevante para la investigación

Referencia	Tema central	Uso dentro de la tesis
Di Mauro <i>et al.</i>		
(2024)	Papel de compras en la organización	Explican por qué compras puede participar en decisiones que van más allá de emitir órdenes.
Delke <i>et al.</i>		
(2023)	Compras directas e indirectas	Muestran que las compras indirectas tienen condiciones y habilidades propias.
Fontes <i>et al.</i>		
(2025)	Gestión por categorías	Ayudan a entender la categoría como una unidad para organizar decisiones de abastecimiento.
Guida <i>et al.</i>		
(2025)	Clasificación del gasto e inteligencia artificial	Describen cómo ordenar transacciones dispersas y por qué la revisión humana sigue siendo necesaria.
Karttunen <i>et al.</i>		
(2023)	Transformación digital en compras	Relacionan el uso de tecnología con procesos claros y datos que puedan utilizarse.
Bernardo <i>et al.</i>		

Continúa en la siguiente página

Cuadro 2.1 – Continuación

Referencia	Tema central	Uso dentro de la tesis
(2024)	Gobernanza y calidad de los datos	Explican cómo las reglas y responsabilidades influyen en el uso de la información.
OECD		
(2022)	Gobernanza de datos	Presenta los datos como un recurso que necesita criterios claros de administración.
International Organization for Standardization y International Electrotechnical Commission		
(2008)	Calidad de datos estructurados	Aporta dimensiones para revisar la exactitud, completitud, consistencia y comprensión de los registros.
International Organization for Standardization		
(2019)	Medición de la calidad de los datos	Orienta el uso de objetivos, indicadores y métricas para evaluar la calidad.
Chapman <i>et al.</i>		
(2000)	CRISP-DM	Presenta las fases originales del método utilizado como guía.

Continúa en la siguiente página

Cuadro 2.1 – Continuación

Referencia	Tema central	Uso dentro de la tesis
Wirth y Hipp		
(2000)	Desarrollo inicial de CRISP-DM	Documenta uno de los antecedentes que ayudaron a consolidar el modelo.
Schröer <i>et al.</i>		
(2021)	Aplicaciones de CRISP-DM	Revisa su uso en distintos proyectos analíticos.
Martínez-Plumed <i>et al.</i>		
(2021)	Evolución de CRISP-DM	Explica que sus fases pueden adaptarse sin perder la lógica del proceso.
Shameti <i>et al.</i>		
(2026)	Comparación entre CRISP-DM y OSEMN	Aporta elementos para justificar la elección de CRISP-DM.
Taylor y Letham		
(2018)	Prophet	Presenta el modelo y su forma de trabajar con tendencia, estacionalidad y eventos.
Hyndman y Athanopoulos		
(2021)	ARIMA y SARI-MA	Aporta la base estadística para explicar y evaluar estos modelos.
Croston		

Continúa en la siguiente página

Cuadro 2.1 – Continuación

Referencia	Tema central	Uso dentro de la tesis
(1972)	Demanda intermitente	Presenta un procedimiento pensado para series con periodos sin demanda.
Syntetos y Boylan		
(2005)	Corrección SBA	Propone un ajuste para reducir el sesgo del método de Croston.
Shale <i>et al.</i>		
(2006)	Variante SBJ	Estudia la estimación de un promedio no sesgado para demanda intermitente.
Ghobbar y Friend		
(2002)	Intermitencia en refacciones	Muestra una aplicación con consumos separados por periodos en cero.
Gutierrez <i>et al.</i>		
(2008)	Demanda <i>lumpy</i>	Explica por qué las series muy irregulares necesitan un tratamiento distinto.
Song y Chissom		
(1993)	Series de tiempo difusas	Aporta la base conceptual del enfoque difuso.
Chen		
(1996)	Pronóstico mediante relaciones difusas	Presenta un procedimiento más sencillo para construir pronósticos.
Nishad <i>et al.</i>		

Continúa en la siguiente página

Cuadro 2.1 – Continuación

Referencia	Tema central	Uso dentro de la tesis
(2023)	Series difusas de orden superior	Muestra un uso reciente del enfoque ante comportamientos irregulares.
Samanta y Al-Araimi		
(2001)	Lógica difusa e inventarios	Relaciona la incertidumbre con decisiones de inventario.
Glock <i>et al.</i>		
(2012)	EOQ con demanda difusa	Conecta la lógica difusa con decisiones de abastecimiento.
Kumar y Evers		
(2015)	Registros imprecisos e inventario de seguridad	Muestra cómo los errores en los datos pueden afectar decisiones posteriores.
Das <i>et al.</i>		
(2024)	TimesFM	Presenta un modelo preentrenado para pronóstico de series temporales.
Choi <i>et al.</i>		
(2018)	Analítica de datos en operaciones	Relaciona los datos históricos con decisiones operativas.
Spreitzenbarth <i>et al.</i>		
(2024)	Inteligencia artificial en compras	Advierte que estas herramientas dependen de los datos y de la preparación de la organización.

Continúa en la siguiente página

Cuadro 2.1 – Continuación

Referencia	Tema central	Uso dentro de la tesis
Hanke y Wichern		
(2006)	Evaluación de pronósticos	Aporta criterios para interpretar medidas de error.
Swamidass		
(2000)	Siete condiciones de la logística	Ayuda a relacionar el estudio con cantidad, tiempo y costo correctos.

Lo que muestra esta tabla es que la base teórica del estudio no proviene de un solo campo. Combina literatura de compras, gobernanza del dato, metodologías de ciencia de datos y pronóstico. Esa combinación no es casual: el problema OC_QA no se resuelve solo con un modelo predictivo. Primero hay que estructurar la información, mejorar su trazabilidad y definir una lógica de análisis por categoría.

Los criterios de selección se presentan en la Tabla 2.2 y las fuentes consultadas en la Tabla 2.3.

TABLA 2.2: Criterios de selección de literatura

Criterio	Aplicación en la revisión	Propósito en la tesis
Pertinencia temática	Se buscaron trabajos sobre compras indirectas, categorías, calidad del dato, CRISP-DM, demanda intermitente y pronóstico.	Mantener la revisión vinculada con el problema estudiado.

Continúa en la siguiente página

Cuadro 2.2 – Continuación

Criterio	Aplicación en la revisión	Propósito en la tesis
Claridad metodológica	Se eligieron documentos que explicaran cómo obtuvieron sus resultados.	Distinguir entre opiniones generales y evidencia con un procedimiento identificable.
Calidad de la fuente	Se priorizaron artículos indexados, normas y documentos institucionales.	Utilizar referencias que pudieran verificarse.
Actualidad	Se incluyeron publicaciones recientes y referencias clásicas necesarias para explicar los modelos.	Combinar desarrollos actuales con sus bases originales.
Relación con OC_QA	Se valoró si la fuente ayudaba a interpretar problemas de clasificación, calidad, consumo o pronóstico.	Evitar una revisión alejada del caso.
Comparación entre trabajos	La matriz registró coincidencias, diferencias y limitaciones.	Evitar que cada fuente apareciera como un resumen aislado.

TABLA 2.3: Fuentes consultadas

Fuente	Tipo de evidencia	Uso dentro de la revisión
Scopus	Base de datos bibliográfica	Localizar artículos recientes sobre compras, logística y análisis de datos.
Web of Science y JCR	Índice y base bibliográfica	Revisar la procedencia y la calidad editorial de las revistas.
ScienceDirect	Plataforma académica	Consultar artículos sobre compras, inventarios y series de tiempo.
Emerald Insight	Plataforma académica	Buscar trabajos relacionados con abastecimiento y transformación digital.
IEEE Xplore	Base técnica	Localizar publicaciones sobre ciencia de datos y modelos computacionales.
ISO	Organismo de normalización	Consultar modelos y procesos relacionados con calidad de datos.
OECD iLibrary	Organismo internacional	Revisar el tratamiento de los datos como recurso organizacional.
Springer	Editorial académica	Consultar conceptos logísticos y referencias especializadas.
Google Scholar	Buscador académico	Rastrear citas y localizar versiones disponibles de los documentos.

Una vez sistematizada la evidencia, la revisión buscó identificar coincidencias, tensiones y vacíos. Karttunen *et al.* (2023) y Di Mauro *et al.* (2024) coinciden en que

compras necesita datos organizados para asumir un rol más estratégico. Spreitzenbarth *et al.* (2024) encontraron, en una revisión mixta sobre inteligencia artificial en compras, que la adopción de estas herramientas depende de la disponibilidad y preparación del dato, además de factores organizacionales. Esto coincide con una idea central del estudio: el análisis predictivo solo tiene valor cuando la base transaccional está suficientemente estructurada.

Algo que también apareció en la revisión fue una diferencia con el enfoque de esta tesis. Buena parte de la literatura trabaja por separado la gobernanza del dato, la clasificación del gasto, la gestión por categorías o los modelos predictivos. Aquí se articulan dentro de una misma ruta: primero se diagnostica y estructura la información, después se categoriza y finalmente se evalúa la posibilidad de pronosticar una categoría piloto. Ese vacío práctico orienta el estudio.

2.3 ABASTECIMIENTO ESTRATÉGICO Y GESTIÓN POR CATEGORÍAS

Las compras indirectas reúnen bienes y servicios que no forman parte del producto final pero sostienen la operación diaria. Delke *et al.* (2023) señalan que este tipo de compras demanda habilidades específicas, con mayor peso en competencias como flexibilidad, comunicación y adaptación. No pueden gestionarse bajo los mismos supuestos que las compras directas.

El abastecimiento estratégico intenta superar esa fragmentación organizando las decisiones por categorías, patrones de consumo y estrategias de aprovisionamiento. Di Mauro *et al.* (2024) sostienen que compras debe ganar un lugar más activo en la toma de decisiones del negocio. Fontes *et al.* (2025) muestran que la gestión por categorías se ha consolidado como un enfoque ampliamente utilizado, aunque advierten que todavía existen diferencias entre su desarrollo teórico y su aplicación práctica.

Antes de analizar una categoría, también es necesario ordenar las transacciones que la forman. Guida *et al.* (2025) muestran que el *spend classification* se ha convertido en un frente relevante para la aplicación de inteligencia artificial en compras, aunque subrayan que el juicio humano sigue siendo necesario para que la clasificación tenga sentido dentro de la organización. En esta tesis, la clasificación del gasto es la base que conecta transacciones dispersas con decisiones de abastecimiento por categoría.

2.4 GOBERNANZA DE DATOS Y DATOS MAESTROS EN COMPRAS

La gobernanza de datos agrupa las estructuras, roles, reglas y mecanismos que permiten tratar los datos como un activo organizacional. La OECD (OECD, 2022) sostiene que los datos tienen valor estratégico y que su gobernanza requiere una visión coherente. Bernardo *et al.* (2024) explican que la gobernanza y la calidad del dato se han vuelto condiciones que influyen en el uso real de la información en distintos contextos organizacionales.

En compras, esto se vuelve muy concreto. Implica definir cómo se describen los artículos, qué unidades de medida se usan, cómo se nombran los proveedores y qué estructura de categorías se aplica. Cuando esos criterios no están alineados, el dato transaccional pierde consistencia. Karttunen *et al.* (2023) muestran que la digitalización de compras requiere procesos estructurados y datos utilizables para evitar que las correcciones se repitan en cada análisis.

En el caso OC-QA, estas condiciones no se cumplían del todo. Por eso el trabajo empezó por ahí: antes de analizar el consumo, fue necesario revisar cómo estaba capturada la información y corregir lo que se podía corregir.

2.5 CALIDAD DEL DATO: DIMENSIONES, ESTÁNDARES Y EVALUACIÓN

Tener datos no es lo mismo que tener datos útiles. La norma ISO/IEC 25012 (International Organization for Standardization y International Electrotechnical Commission, 2008) define un modelo general de calidad para información estructurada y establece criterios para especificar requisitos, medidas y evaluaciones. ISO 8000-63 (International Organization for Standardization, 2019) propone una estructura para medir los procesos de gestión de calidad mediante objetivos, indicadores y métricas.

Aplicado a OC_QA, este marco ayuda a entender que campos incompletos, descripciones ambiguas o unidades heterogéneas no son solo errores de captura. Son señales de una calidad de datos insuficiente para soportar decisiones analíticas. Bernardo *et al.* (2024) explican que la calidad del dato es inseparable de su aprovechamiento organizacional. Kumar y Evers (2015) muestran que incluso decisiones de inventario de seguridad se ven afectadas cuando los registros son imprecisos.

Limpiar la base no fue una tarea auxiliar en esta investigación. Fue el primer paso metodológico, sin el cual no era posible categorizar, construir series temporales ni comparar modelos de forma confiable.

2.6 METODOLOGÍAS PARA PROYECTOS DE DATOS: CRISP-DM Y OSEM N

Cuando un proyecto combina problema de negocio, preparación de datos, modelado y evaluación, necesita un orden. CRISP-DM es uno de los marcos más usados para ese propósito. Chapman *et al.* (2000) lo presentan como una guía paso a paso para proyectos de minería de datos. Wirth y Hipp (2000) son un antecedente

temprano en su consolidación como modelo de referencia. Schröer *et al.* (2021) confirman que sigue siendo un estándar de facto en proyectos analíticos.

Martínez-Plumed *et al.* (2021) proponen releer CRISP-DM veinte años después. Su argumento es que el valor del modelo no está en ser un esquema rígido, sino en ofrecer trayectorias comprensibles para conectar el problema con las decisiones analíticas. Eso es exactamente lo que se necesitaba aquí: no solo correr modelos, sino documentar con claridad la lógica que une el diagnóstico de datos, la categorización y la comparación de pronósticos.

Como contraste, OSEMN resume el flujo de ciencia de datos en cinco momentos: obtener, depurar, explorar, modelar e interpretar. Shameti *et al.* (2026) comparan ambos marcos y concluyen que CRISP-DM ofrece una estructura más completa para organizar proyectos analíticos de forma trazable. Por eso se eligió como estructura principal en esta tesis.

2.7 ABASTECIMIENTO ESTRATÉGICO Y CONTRIBUCIÓN AL CUMPLIMIENTO DE LAS 7R DE LA LOGÍSTICA

Las 7R de la logística son un marco conocido para interpretar el desempeño del abastecimiento. Swamidass (2000, p. 684) las sintetiza como la capacidad de entregar el producto correcto, en la cantidad correcta, en la condición correcta, en el lugar correcto, en el tiempo correcto, al cliente correcto y al costo correcto.

En esta tesis no se usan las 7R como definición general de logística. Se usan para mostrar dónde impacta concretamente el trabajo realizado. Cuando la información de compras indirectas es más confiable y trazable, mejoran especialmente tres dimensiones: cantidad correcta, tiempo correcto y costo correcto. Sin saber qué se compra, cuánto se consume y con qué frecuencia aparece una necesidad, es difícil

anticiparse.

En términos operativos, las 7R pueden describirse así:

- **Producto correcto:** que lo adquirido corresponda realmente a lo que la operación necesita, evitando compras ambiguas o sustituciones inadecuadas.
- **Cantidad correcta:** comprar el volumen adecuado para responder al consumo esperado, sin generar sobreabasto ni desabasto.
- **Tiempo correcto:** tener disponible el recurso cuando la operación lo requiere, reduciendo urgencias y retrasos.
- **Lugar correcto:** que el producto llegue al punto donde se usará o almacenará, con la trazabilidad necesaria.
- **Condición correcta:** que el bien o servicio llegue en condiciones utilizables, sin deterioro ni incumplimiento de especificaciones.
- **Cliente correcto:** que el recurso se dirija al área o proceso que realmente lo necesita.
- **Costo correcto:** que el abastecimiento ocurra con una relación adecuada entre precio, calidad y costo total.

La conexión entre las 7R y esta tesis no es decorativa. Mejores datos y mejores pronósticos pueden apoyar decisiones más oportunas sobre cuánto comprar, cuándo anticiparse y cómo analizar el costo del abastecimiento.

2.8 ANALÍTICA PREDICTIVA POR CATEGORÍA: UTILIDAD, ALCANCES Y LIMITACIONES

No todas las categorías de compra se comportan igual. Algunas tienen consumo estable y predecible. Otras tienen semanas sin movimiento, picos aislados y una variabilidad que complica cualquier estimación. Analizar el abastecimiento de forma agregada puede servir para una vista general, pero no alcanza cuando se necesita tomar decisiones específicas por rubro.

El pronóstico por categoría puede disminuir la dependencia de decisiones tomadas únicamente cuando aparece una urgencia. En lugar de esperar a que aparezca una necesidad para reaccionar, permite estimar patrones futuros, identificar recurrencias y anticipar periodos de mayor presión. Esto también ayuda a priorizar inventarios, revisar frecuencias de reabastecimiento y detectar categorías que merecen más atención analítica.

Pero no todas las categorías necesitan el mismo enfoque. Algunas funcionan bien con modelos tradicionales. Otras requieren métodos que reconozcan la intermitencia, los ceros frecuentes o los cambios abruptos de nivel. El pronóstico por categoría no sigue una receta uniforme; es una forma de adaptar el análisis a la naturaleza real del consumo.

En esta tesis, el pronóstico no aparece como punto de partida. Se incorpora después de preparar la base, construir la categorización y verificar que la serie piloto tiene suficiente consistencia para ser modelada. Bernardo *et al.* (2024) vinculan la calidad del dato con su aprovechamiento organizacional, y eso es exactamente lo que ocurre aquí: sin una base gobernable, el modelo hereda los problemas de los datos en lugar de resolverlos.

También hay límites que deben reconocerse. La utilidad de la predicción depende de la consistencia temporal de los datos, la calidad del registro y la naturaleza

de la serie. Cuando el dato es insuficiente, el modelo no corrige el problema; puede amplificarlo. Por eso en esta investigación el pronóstico se usa como herramienta de apoyo a la planeación, no como sustituto del criterio operativo.

2.9 FUNDAMENTOS TEÓRICOS DE LOS MODELOS DE PRONÓSTICO CONSIDERADOS

Esta investigación compara enfoques con naturalezas distintas: Prophet, ARIMA/SARIMA, series de tiempo difusas, TimesFM y métodos especializados para demanda intermitente. La selección buscó reunir enfoques con distintos niveles de interpretabilidad, sensibilidad a la intermitencia y capacidad de adaptación a patrones irregulares de consumo.

2.9.1 MODELO PROPHET

Prophet es un modelo de pronóstico para series temporales que representa el comportamiento de la variable como una combinación aditiva de tendencia, estacionalidad, efectos especiales y error. Taylor y Letham (2018) explican que este enfoque fue desarrollado para facilitar pronósticos prácticos, interpretables y escalables en series con componentes de tendencia y estacionalidad. Su formulación general puede expresarse como:

$$y(t) = g(t) + s(t) + h(t) + \varepsilon_t \quad (2.1)$$

donde $g(t)$ representa la tendencia, $s(t)$ la estacionalidad, $h(t)$ los efectos asociados a eventos o calendarios y ε_t el término de error no explicado por el modelo.

En su variante base, Prophet modela la tendencia mediante una función lineal

por tramos:

$$g(t) = (k + a(t)^\top \delta)t + (m + a(t)^\top \gamma) \quad (2.2)$$

En la variante logística, la tendencia se representa mediante una forma de crecimiento acotado:

$$g(t) = \frac{C}{1 + \exp(-k(t - m))} \quad (2.3)$$

donde C actúa como capacidad máxima, k controla la tasa de crecimiento y m representa el punto medio de la curva.

La ventaja principal de Prophet está en su interpretabilidad y en la posibilidad de modelar tendencia y estacionalidad de forma flexible. Su limitación aparece cuando la serie tiene intermitencia marcada, valores frecuentes en cero o cambios abruptos que dificultan una representación suavizada.

2.9.2 MODELO ARIMA/SARIMA

ARIMA combina componentes autorregresivos, diferenciación y medias móviles. Hyndman y Athanasopoulos (2021) lo presentan como uno de los enfoques clásicos más utilizados para analizar y pronosticar series temporales con estructura serial. Su formulación general es:

$$\phi(B)(1 - B)^d y_t = \theta(B)\varepsilon_t \quad (2.4)$$

Cuando la serie tiene patrón estacional, la extensión SARIMA incorpora componentes específicos para esa recurrencia:

$$\Phi(B^s)\phi(B)(1-B)^d(1-B^s)^D y_t = \Theta(B^s)\theta(B)\varepsilon_t \quad (2.5)$$

donde s representa la longitud del ciclo estacional.

Su fortaleza está en la base estadística consolidada para modelar series con estructura. Su limitación aparece cuando la serie es demasiado irregular o cuando la relación señal-ruido dificulta identificar patrones robustos.

2.9.3 PRONÓSTICO BASADO EN SERIES DE TIEMPO DIFUSAS

Este enfoque representa los valores observados mediante conjuntos difusos y relaciones lógicas entre estados. Song y Chissom (1993) establecen la base conceptual de este tipo de series. Chen (1996) propone un enfoque práctico de pronóstico a partir de relaciones difusas. La función de pertenencia puede expresarse como:

$$\mu_{A_i}(x) : U \rightarrow [0, 1] \quad (2.6)$$

Las relaciones difusas se representan como transiciones entre estados:

$$A_i \rightarrow A_j \quad (2.7)$$

y el valor pronosticado se obtiene a partir de un proceso de desfusificación:

$$\hat{y}_{t+1} = \text{Defuzz}(A_j) \quad (2.8)$$

Nishad *et al.* (2023) muestran que este tipo de modelos sigue siendo pertinente para representar incertidumbre y comportamiento irregular. Su fortaleza está en la flexibilidad conceptual. Su limitación es que puede tender a soluciones demasiado rígidas cuando la serie exhibe intermitencia intensa y cambios bruscos de nivel.

2.9.4 MODELO TIMESFM

TimesFM es un modelo fundacional de series de tiempo basado en aprendizaje profundo. Das *et al.* (2024) lo describen como un modelo diseñado para trabajar a partir de contexto histórico y producir horizontes futuros de predicción. Su lógica puede resumirse como:

$$\hat{y}_{t+1:t+h} = f_{\theta}(y_{t-c+1:t}) \quad (2.9)$$

donde $y_{t-c+1:t}$ es la secuencia de contexto, $\hat{y}_{t+1:t+h}$ el vector de pronóstico y f_{θ} la función aprendida por el modelo.

En esta tesis sirve como alternativa reciente para comparar un enfoque pre-entrenado con modelos más clásicos. Su interés está en aplicar patrones aprendidos previamente a una nueva serie. Su limitación principal es que la sofisticación del modelo no garantiza mejor desempeño en series cortas, intermitentes y con muchos ceros.

2.9.5 MÉTODOS ESPECIALIZADOS PARA DEMANDA INTERMITENTE

No todas las series de compras tienen un comportamiento continuo. Algunas categorías registran periodos prolongados sin consumo, intercalados con ocurrencias positivas de magnitud variable. Ese patrón se conoce como demanda intermitente y requiere un tratamiento distinto al de las series más regulares.

Croston (1972) desarrolló uno de los enfoques clásicos para este tipo de demanda, separando el modelado del tamaño de la demanda positiva y el intervalo entre ocurrencias. Ghobbar y Friend (2002) muestran que la intermitencia puede caracterizarse mediante indicadores como el promedio del intervalo entre demandas y

la variabilidad relativa del tamaño, lo que permite distinguir patrones intermitentes, erráticos o *lumpy* antes de elegir el método.

Además del método original de Croston, en esta tesis se incorporan dos variantes. La aproximación SBA corrige parte del sesgo del método original (Syntetos y Boylan, 2005). SBJ parte de la misma familia y busca una estimación no sesgada del promedio (Shale *et al.*, 2006). Ambas se incluyeron para revisar cómo responden ante una serie irregular. Gutierrez *et al.* (2008) señalan que las series con alta irregularidad necesitan tratamiento diferenciado respecto a una demanda estable. Choi *et al.* (2018) agregan que el uso de datos históricos y analítica avanzada mejora la calidad de las decisiones cuando la selección metodológica es consistente con la estructura real de los datos.

La inclusión de Croston, SBA y SBJ responde a una decisión metodológica concreta: cuando la categoría analizada tiene múltiples semanas en cero, alta separación entre ocurrencias y variabilidad considerable en los valores positivos, tiene sentido complementar los modelos generales con enfoques diseñados para ese comportamiento específico.

2.10 MÉTRICAS DE EVALUACIÓN Y CRITERIOS DE COMPARACIÓN DE MODELOS

Los modelos se evaluaron con las mismas métricas y el mismo horizonte de prueba. Además del error, se revisó si los valores pronosticados tenían sentido para la operación. Hanke y Wichern (2006) presentan una base útil para la evaluación cuantitativa del desempeño predictivo. Se utilizaron el error absoluto medio, la raíz del error cuadrático medio y una versión robusta del error porcentual absoluto medio:

$$MAE = \frac{1}{n} \sum_{t=1}^n |y_t - \hat{y}_t| \quad (2.10)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2} \quad (2.11)$$

$$MAPE_{robusto} = \frac{100}{n} \sum_{t=1}^n \left| \frac{y_t - \hat{y}_t}{\max(|y_t|, \varepsilon)} \right| \quad (2.12)$$

Además de estas métricas, la comparación considera si el modelo genera predicciones negativas y cuál es el valor mínimo pronosticado:

$$N_{\text{neg}} = \sum_{t=1}^n I(\hat{y}_t < 0) \quad (2.13)$$

$$\hat{y}_{\min} = \min(\hat{y}_1, \hat{y}_2, \dots, \hat{y}_n) \quad (2.14)$$

En un problema de consumo o abastecimiento, un pronóstico no solo debe minimizar el error promedio. También debe mantener sentido para la operación. Por eso la selección final del modelo no depende de una sola métrica, sino de una lectura integrada entre precisión y utilidad práctica.

2.11 CASOS ORGANIZACIONALES Y APRENDIZAJES PARA OC_QA

La transformación digital no produce valor por sí sola. Su impacto aparece cuando se apoya en procesos más ordenados, datos consistentes y criterios claros para usarlos en la toma de decisiones. Los casos revisados no son equivalentes al estudio OC_QA, pero ofrecen lecciones útiles sobre cómo la gobernanza del dato y la visibilidad de la información pueden mejorar la planeación.

2.11.1 UNILEVER: SIMPLIFICACIÓN OPERATIVA Y GOBIERNO DEL DATO

Unilever (Unilever, 2025) ha documentado cómo la digitalización se utilizó para mejorar visibilidad, simplificar procesos y favorecer decisiones más ágiles dentro de su operación global. Pega (Pega, n.d.) registra un caso en el que la empresa redujo significativamente los tiempos de incorporación de proveedores al transformar sus procesos de datos maestros y simplificar flujos de trabajo internos.

Lo que se puede traer de este caso es directo: estandarizar y gobernar el dato no son tareas administrativas menores. Son condiciones para que las decisiones puedan repetirse, compararse y trazarse. En OC_QA, donde había descripciones genéricas, unidades inconsistentes y categorías poco claras, esto significó que ordenar los registros era el primer paso real hacia una función de abastecimiento con más capacidad de anticipación.

2.11.2 GRUPO BIMBO: DATOS ACCESIBLES COMO SOPORTE PARA DECISIONES

Microsoft (Microsoft, 2022) documenta cómo Grupo Bimbo utilizó soluciones de Azure IoT para mejorar la visibilidad y la eficiencia en sus procesos productivos. Microsoft News (Microsoft News Center Europe, 2022) reporta la transformación del análisis de datos del área comercial mediante Power BI y Azure Synapse Analytics.

El caso de Grupo Bimbo no se sitúa en compras indirectas, pero el principio es transferible: cuando los datos están accesibles, comparables y estructurados, su valor para decidir crece. En esta tesis eso se tradujo en la necesidad de convertir OC_QA en un insumo analítico gobernable, capaz de sostener una categorización jerárquica y una evaluación comparable de modelos de pronóstico.

2.11.3 LECCIONES TRANSFERIBLES PARA EL CASO OC_QA

De los dos casos anteriores se pueden extraer tres ideas aplicables. Primera: la tecnología genera más valor cuando va acompañada de reglas de consistencia y simplificación de procesos, no solo de herramientas nuevas. Segunda: que los datos sean accesibles y estén estructurados es una condición para pasar de un uso reactivo de la información a uno más estratégico. Tercera: la trazabilidad del dato y su organización por criterios compartidos mejoran la capacidad de planeación, algo especialmente relevante en compras indirectas donde la dispersión del gasto y la ambigüedad descriptiva complican las decisiones.

Estos casos también complementan la discusión sobre demanda intermitente: la utilidad de la analítica predictiva depende no solo de la calidad del dato, sino de que el método elegido corresponda al patrón real de la serie.

2.12 IMPLICACIONES PARA LA INVESTIGACIÓN

La revisión de literatura deja varias ideas que sostienen esta tesis. Delke *et al.* (2023) muestran que las compras indirectas necesitan una lógica propia de gestión. Bernardo *et al.* (2024) explican que la gobernanza y la calidad del dato son condiciones para aprovechar la información para el análisis. Fontes *et al.* (2025) permiten entender la gestión por categorías como base para decisiones diferenciadas en compras. Guida *et al.* (2025) precisan que la clasificación del gasto conecta registros dispersos con decisiones por categoría.

La revisión también muestra que la analítica avanzada en compras no puede aplicarse sin condiciones previas. Spreitzenbarth *et al.* (2024) advierten que adoptar inteligencia artificial y aprendizaje automático en compras requiere datos disponibles, preparación organizacional y condiciones de implementación. Sin eso, el pronóstico no resuelve el problema; lo traslada a otra etapa.

Hay además una implicación técnica relevante: no todas las categorías de consumo se pueden analizar con los mismos modelos. Cuando una serie tiene muchos periodos en cero, alta separación entre ocurrencias y variabilidad considerable en los valores positivos, tiene sentido incorporar métodos especializados como Croston, SBA y SBJ. Eso amplía el comparativo principal y evita forzar un modelo diseñado para demanda continua sobre una serie que no lo es.

En conjunto, estos elementos justifican la tesis como una propuesta aplicada: no solo comparar modelos predictivos, sino construir primero la base que los hace posibles. El orden seguido conecta gobernanza del dato, categorización jerárquica y análisis predictivo dentro de un mismo flujo metodológico aplicado a compras indirectas.

CAPÍTULO 3

METODOLOGÍA

3.1 ENFOQUE DEL ESTUDIO

La investigación se planteó como un estudio aplicado. El punto de partida fue un problema encontrado en la base OC_QA. La empresa tenía registros históricos de compras indirectas, pero estos no podían utilizarse directamente para analizar el consumo ni preparar pronósticos por categoría.

El trabajo se organizó alrededor de esa dificultad. Primero se revisó el estado de la información. Después se corrigieron diferencias en las variables principales y se construyó una categorización jerárquica. La comparación de modelos comenzó cuando la base ya tenía una estructura adecuada para formar una serie temporal.

CRISP-DM dio el orden general del proceso. Sus seis fases ayudaron a separar la comprensión del problema, la revisión de OC_QA, la preparación de los registros, el modelado y la evaluación. La estructura original fue presentada por Chapman *et al.* (2000) y Wirth y Hipp (2000), pero las fases se adaptaron a lo que fue apareciendo durante el análisis.

No siempre fue posible avanzar en línea recta. En algunos momentos hubo que regresar a los datos, corregir una regla o revisar nuevamente la serie antes de

continuar. Este tipo de ajustes también aparece en otras aplicaciones de CRISP-DM revisadas por Schröer *et al.* (2021).

El método ayudó a mantener la relación entre la necesidad del área de compras y las decisiones tomadas durante el análisis. El punto de partida no fue elegir un modelo. Antes era necesario entender qué contenía la base, qué problemas presentaba y hasta dónde podía utilizarse.

3.2 DISEÑO Y UNIDAD DE ANÁLISIS

El estudio se desarrolló como un caso aplicado sobre OC_QA. Se trabajó con los registros disponibles durante el periodo analizado, sin extraer una muestra aleatoria.

Cada fila de OC_QA corresponde a una transacción de compra. Esta fue la unidad de observación durante el diagnóstico y la preparación de los datos. Para estudiar el consumo, las transacciones se agruparon por categoría y periodo. De esta transformación surgieron las series semanales utilizadas en el pronóstico.

El problema se revisó dentro de su contexto original. Las descripciones, unidades de medida, proveedores, almacenes y categorías no se trataron como campos independientes. Se analizaron como partes de un registro que debía conservar sentido para el área de compras.

Cada transformación respondió a una dificultad concreta. La homologación de unidades surgió de las diferencias en la captura. Las descripciones se limpiaron porque varios registros no permitían reconocer con claridad lo comprado. La nueva jerarquía se construyó debido a que la clasificación original no ofrecía suficiente detalle.

Esta forma de trabajar mantiene la relación entre la necesidad inicial, las decisiones tomadas y los resultados obtenidos, un vínculo que Martínez-Plumed *et al.* (2021) consideran necesario en los proyectos de datos.

La investigación no busca afirmar que los resultados de una categoría sean válidos para todas las compras indirectas. La categoría piloto se utilizó para probar el procedimiento completo y observar cómo respondían distintos modelos frente a una misma serie.

3.3 FUENTE DE DATOS Y PREPARACIÓN DE OC_QA

La fuente principal fue una base de órdenes de compra indirectas de una empresa alimentaria en México. El nombre de la organización y la denominación original del archivo se mantuvieron reservados. OC_QA es el identificador utilizado en la tesis para referirse al conjunto de datos.

La base contenía información histórica de artículos, servicios, proveedores, almacenes, cantidades, unidades de medida y categorías. También presentaba problemas que dificultaban su uso. Había campos incompletos, descripciones generales, unidades registradas de distintas maneras y categorías con poco detalle.

La preparación comenzó con una revisión de la estructura del archivo. Se identificaron los campos disponibles, el tipo de dato almacenado en cada columna y la cantidad de registros con valores faltantes. Después se revisaron las descripciones, las unidades de medida y la clasificación existente.

La norma ISO/IEC 25012 sirvió como referencia para interpretar problemas de completitud, consistencia y comprensión de los registros (International Organization for Standardization y International Electrotechnical Commission, 2008). ISO 8000-63 también se consideró porque propone evaluar la calidad mediante objetivos, indicadores y medidas definidas (International Organization for Standardization, 2019).

Las principales actividades fueron las siguientes:

1. Revisión de campos vacíos y valores inconsistentes.
2. Limpieza de las descripciones utilizadas para identificar artículos y servicios.
3. Homologación de unidades de medida registradas con nombres o abreviaturas diferentes.
4. Creación de variables para identificar problemas de calidad.
5. Construcción de una jerarquía de categorías.
6. Preparación de agregaciones semanales para el análisis temporal.

La base original no fue reemplazada. Se conservó como referencia y las modificaciones se realizaron en una versión de trabajo. Esto permitió comparar los campos iniciales con las variables creadas durante el proceso.

3.4 ADAPTACIÓN DE CRISP-DM AL CASO

Las seis fases de CRISP-DM se adaptaron al trabajo realizado con OC_QA.

La comprensión del negocio comenzó con el problema del área de compras. La empresa tenía información histórica, pero no contaba con una estructura que facilitara el seguimiento del consumo por categoría. También se definió que el pronóstico funcionaría como apoyo y no como una decisión automática de compra.

La comprensión de los datos se concentró en el diagnóstico de OC_QA. Se revisaron los registros, la cobertura de las variables, las unidades de medida, las descripciones y la clasificación disponible. Esto ayudó a reconocer qué campos podían utilizarse y cuáles necesitaban ajustes.

Durante la preparación se aplicaron las reglas de limpieza, se homologaron las unidades y se crearon las variables derivadas. En esta misma fase se construyó la jerarquía formada por macrocategoría, nivel 1 y nivel 2. Después se agruparon las transacciones para preparar las series semanales.

La fase de modelado incluyó Prophet, ARIMA/SARIMA, series de tiempo difusas mediante FTS-Chen, TimesFM y dos variantes para demanda intermitente derivadas del método de Croston: SBA y SBJ. Todos los modelos recibieron la misma serie y se evaluaron con un horizonte común de 21 semanas.

Durante la evaluación se compararon los pronósticos con las observaciones del conjunto de prueba. También se revisó si los resultados tenían sentido para el problema estudiado. Un error bajo no era suficiente cuando el modelo producía cantidades negativas o una trayectoria poco relacionada con la serie real.

El procedimiento no llegó a instalarse de forma permanente dentro de la empresa, ya que esa etapa quedó fuera del alcance de la investigación. La fase de despliegue se concentró en organizar los archivos, gráficas, parámetros y resultados necesarios para continuar el trabajo o repetirlo con otra categoría.

CRISP-DM se eligió sobre OSEMN porque el trabajo no comenzaba con la obtención o limpieza de datos. Primero era necesario entender el problema del área de compras y definir para qué se utilizaría la información. OSEMN resultaba útil para describir las tareas técnicas, pero CRISP-DM daba mayor espacio a la comprensión del negocio, la evaluación y la continuidad del proyecto. La comparación presentada por Shameti *et al.* (2026) reconoce esta diferencia entre ambos enfoques.

Las figuras siguientes presentan dos niveles del método. La primera muestra la estructura general de CRISP-DM. La segunda explica cómo se adaptó al caso de compras indirectas. En ambas, el modelado aparece después de entender y preparar la información.

La Figura 3.1 resume las fases generales de CRISP-DM.

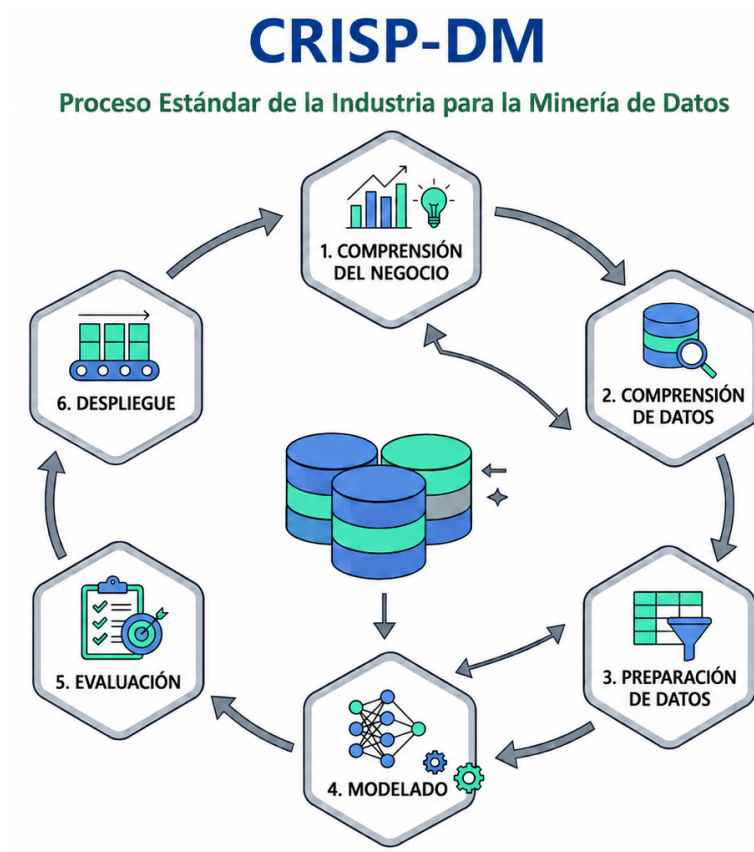


FIGURA 3.1: Fases principales de la metodología CRISP-DM

Fuente: elaboración propia con base en Chapman *et al.* (2000).

Flujo metodológico aplicado al caso de compras indirectas OC_QA

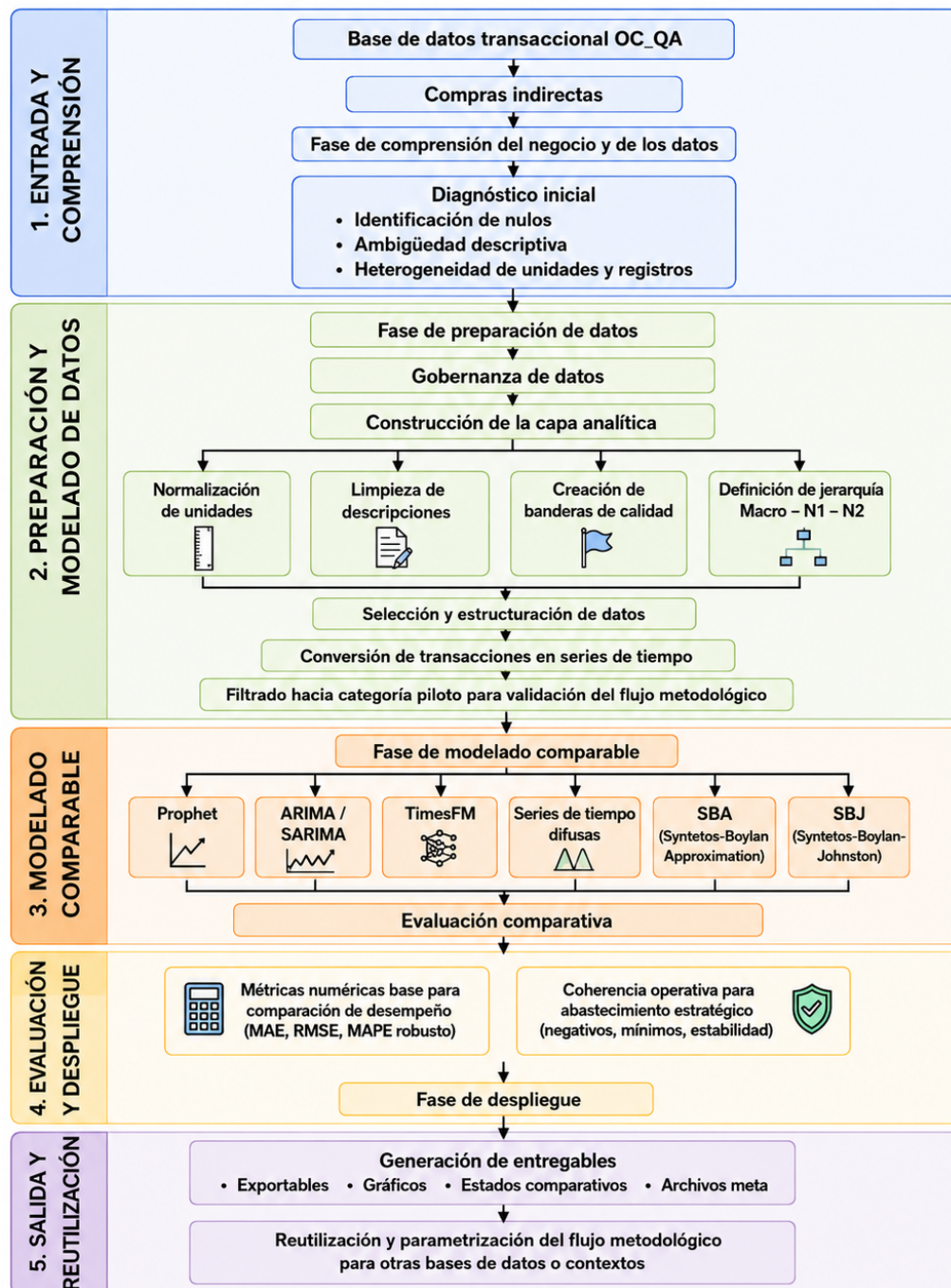


FIGURA 3.2: Flujo metodológico aplicado al caso OC_QA

Fuente: elaboración propia.

La Figura 3.2 muestra cómo se adaptaron las fases de CRISP-DM al trabajo realizado con OC_QA. El flujo comienza con la comprensión del problema y la revisión de los datos. Después incorpora la preparación, la categorización, el modelado y la evaluación de los resultados.

3.5 SELECCIÓN DE LA CATEGORÍA PILOTO

LP_RE02 no se eligió al azar. Su historial permitía construir una serie semanal y reservar un periodo para evaluar los pronósticos. Al mismo tiempo, presentaba semanas sin consumo y algunos movimientos de magnitud considerable.

Esta combinación era útil porque permitía comparar modelos generales con métodos desarrollados para demanda intermitente.

La serie se dividió respetando el orden de las semanas. No se utilizó una separación aleatoria porque eso habría mezclado observaciones anteriores y posteriores, creando una condición diferente a la que existe al pronosticar el futuro.

Los datos más antiguos se utilizaron para entrenar los modelos y las últimas 21 semanas quedaron reservadas para la prueba. Este procedimiento sigue la lógica de evaluación temporal descrita por Hyndman y Athanasopoulos (2021).

La selección de LP_RE02 no implica que sea la categoría más importante de la base ni que represente el comportamiento de los demás rubros. Se utilizó como caso de validación porque reunía las condiciones necesarias para probar el procedimiento.

Antes de modelarla, la serie se revisó visualmente y se analizaron sus valores semanales. También se verificaron la presencia de ceros, la separación entre los movimientos positivos y los cambios en la magnitud del consumo. Estas características influyeron en la elección de los modelos.

3.6 VARIABLES Y REGISTRO DE LAS TRANSFORMACIONES

El archivo analítico conservó variables procedentes de OC_QA y añadió otras creadas durante la preparación. Entre los campos originales se mantuvieron los identificadores necesarios para relacionar cada registro con su fuente.

Las variables derivadas incluyeron la descripción limpia, la unidad de medida homologada, las banderas de calidad y los niveles de categoría. También se generaron campos temporales para agrupar el consumo por semana.

Las transformaciones no se aplicaron directamente sobre el archivo original. Cada ajuste se realizó en una versión de trabajo y quedó asociado con una regla definida. De esta forma se podía revisar cómo un registro inicial terminó dentro de una categoría o de una serie temporal.

La gobernanza de datos necesita reglas y responsabilidades que ayuden a conservar la consistencia de la información, como explican Bernardo *et al.* (2024). La OECD también plantea que los datos adquieren mayor utilidad cuando su administración sigue criterios conocidos (OECD, 2022).

En OC_QA, estas ideas se llevaron a reglas sencillas. Las unidades equivalentes se asignaron a una forma común. Las descripciones se limpiaron sin eliminar el vínculo con el texto original. Las nuevas categorías se agregaron como campos adicionales y la información de origen permaneció disponible para cualquier revisión posterior.

Para los modelos se utilizó una estructura con fecha y cantidad semanal. El formato fue el mismo para todos los enfoques. Así se evitó que un modelo recibiera una versión diferente de la serie.

3.7 PROTOCOLO DE COMPARACIÓN

Los criterios de comparación se definieron antes de conocer los resultados de los modelos. Todos los enfoques utilizaron la misma categoría, la misma frecuencia semanal, la misma partición temporal y un horizonte de prueba de 21 semanas.

Esta decisión evitó que un modelo tuviera más observaciones de entrenamiento o un periodo de prueba distinto. Los resultados se compararon sobre las mismas semanas reales.

Las métricas principales fueron el error absoluto medio, la raíz del error cuadrático medio y el error porcentual absoluto ponderado. El error porcentual absoluto medio no se utilizó como criterio de selección porque la serie contiene varias observaciones iguales a cero. En esas condiciones, su valor puede crecer de forma desproporcionada.

El error porcentual absoluto medio simétrico se conservó como una medida descriptiva complementaria. También se contó el número de pronósticos negativos y se revisaron los valores mínimo y máximo generados por cada modelo.

La selección final siguió un orden definido. Primero se comparó la cantidad de pronósticos negativos. Después se revisaron RMSE, MAE y WAPE. La lectura visual se utilizó para entender las diferencias entre los modelos, pero no sustituyó las métricas.

Prophet generó intervalos de incertidumbre del 80 %. ARIMA/SARIMA produjo intervalos de predicción del 95 %. FTS-Chen, TimesFM, SBA y SBJ no entregaron bandas directamente comparables en los archivos utilizados para la evaluación.

Los intervalos se analizaron dentro de cada modelo, pero no formaron parte del criterio común de selección. La comparación integral se realizó con los pronósticos

puntuales de las mismas 21 semanas.

La configuración de cada modelo se conservó en archivos de apoyo. También se generaron archivos separados para la serie completa, el entrenamiento, la prueba, los pronósticos y las métricas. Esto facilitó revisar los resultados sin ejecutar nuevamente todo el proceso.

3.8 APLICACIÓN DE LOS MODELOS

3.8.1 PROPHET

Prophet se ejecutó con la serie semanal de LP_RE02. Para la variante base se evaluaron 64 combinaciones de parámetros. También se preparó una versión con crecimiento logístico para observar si una restricción sobre el crecimiento ayudaba a mantener valores más razonables.

Taylor y Letham (2018) plantean que Prophet separa la serie en componentes como tendencia y estacionalidad. En este estudio interesaba revisar si esa estructura podía representar una categoría con semanas en cero y picos aislados.

La configuración base utilizó una tendencia flexible y no activó componentes semanales ni anuales. La variante logística incorporó un límite inferior de cero y un límite superior calculado a partir del máximo observado en entrenamiento.

Las dos versiones se evaluaron con las mismas 21 semanas. La selección interna consideró las métricas y el comportamiento del pronóstico. La configuración base se conservó como representante de Prophet porque presentó menor MAE y WAPE que la variante logística.

3.8.2 ARIMA/SARIMA

ARIMA/SARIMA se trabajó mediante una búsqueda de 64 configuraciones. Cada alternativa se entrenó con el mismo tramo histórico y después generó un pronóstico para las 21 semanas de prueba.

Estos modelos utilizan la relación entre observaciones anteriores y errores pasados. La extensión SARIMA incorpora componentes para representar recurrencias estacionales, como explican Hyndman y Athanasopoulos (2021).

La selección priorizó la ausencia de pronósticos negativos. Entre las configuraciones que cumplieron esa condición se compararon RMSE, MAE y WAPE. El AIC se mantuvo como una referencia de ajuste, pero no fue el criterio principal para escoger el modelo final.

Además del error, se revisaron los residuos, los intervalos de predicción y la forma del pronóstico. El propósito era identificar una configuración que respondiera bien fuera del periodo utilizado para ajustarla.

3.8.3 SERIES DE TIEMPO DIFUSAS

El modelo difuso se incluyó porque trabaja con estados e intervalos en lugar de depender únicamente de relaciones estadísticas lineales. Song y Chissom (1993) presentaron la base de este enfoque y Chen (1996) propuso un procedimiento más sencillo para construir pronósticos.

La implementación utilizada fue FTS-Chen con siete intervalos iguales y reglas difusas de primer orden. Esta configuración se definió para el experimento y no formó parte de una búsqueda entre diferentes cantidades de intervalos.

Para LP_RE02 se construyeron los intervalos, los estados difusos y las relaciones entre ellos. Después se generaron de forma recursiva los valores correspondientes a

las 21 semanas de prueba.

La configuración quedó documentada porque el número y el tamaño de los intervalos influyen directamente en el resultado. Un esquema demasiado amplio puede producir una línea poco sensible a los cambios. Uno demasiado detallado puede responder de manera inestable.

El interés estaba en observar cómo reaccionaba el enfoque ante las semanas sin consumo y los picos de la serie.

3.8.4 TIMESFM

TimesFM se incorporó como una alternativa preentrenada. Das *et al.* (2024) describen este tipo de modelo como una propuesta que utiliza patrones aprendidos a partir de grandes colecciones de series temporales.

La configuración utilizada correspondió a TimesFM 2.5, con 200 millones de parámetros. La serie de LP_RE02 se entregó como contexto histórico y el modelo produjo directamente las 21 semanas siguientes.

No se compararon distintos tamaños, checkpoints o combinaciones de parámetros. TimesFM se utilizó con una configuración fija para observar cómo respondía ante una serie corta, irregular y con varios periodos en cero.

La inclusión del modelo no partió de la idea de que una alternativa más reciente debía obtener un mejor resultado. Su desempeño se evaluó con las mismas métricas y las mismas fechas utilizadas para los demás enfoques.

3.8.5 SBA Y SBJ PARA DEMANDA INTERMITENTE

La presencia de semanas sin consumo llevó a incorporar métodos diseñados para demanda intermitente. El procedimiento original de Croston (1972) separa la estimación del tamaño de la demanda y el tiempo que transcurre entre sus ocurrencias.

En esta investigación, Croston se utilizó como base conceptual. Los métodos ejecutados fueron SBA y SBJ, dos variantes que introducen correcciones sobre el procedimiento original.

Antes de aplicarlos se calcularon el intervalo promedio entre demandas, conocido como ADI, y el cuadrado del coeficiente de variación, identificado como CV^2 . Estos indicadores ayudaron a describir la separación entre los movimientos positivos y la variabilidad de sus tamaños. Los resultados de esta caracterización se presentan en el Capítulo 4.

SBA aplica una corrección al método original para reducir el sesgo de sus estimaciones, de acuerdo con Syntetos y Boylan (2001). Un trabajo posterior de Syntetos y Boylan (2005) evaluó esta propuesta frente a otros métodos con series reales de demanda intermitente.

SBJ pertenece a la misma familia de ajustes. Shale *et al.* (2006) desarrollaron una corrección orientada a obtener una estimación no sesgada del promedio bajo determinadas condiciones.

Ambos métodos utilizaron la misma serie semanal. Para seleccionar el parámetro de suavización α , se probaron 19 valores entre 0.05 y 0.95, con incrementos de 0.05.

La calibración se realizó con las últimas seis semanas de la serie. SBA obtuvo su menor RMSE con $\alpha = 0.55$ y SBJ con $\alpha = 0.40$. Esta ventana se utilizó únicamente para elegir el parámetro.

Después de fijar los valores de α , los dos métodos generaron pronósticos para las 21 semanas del conjunto de prueba. Las métricas finales se recalcularon sobre ese horizonte común y se compararon con Prophet, ARIMA/SARIMA, FTS-Chen y TimesFM.

3.9 HERRAMIENTAS DE TRABAJO

El proceso se realizó en cuadernos computacionales con Python como lenguaje principal. La preparación de los datos utilizó principalmente pandas y NumPy. Matplotlib se empleó para las gráficas y scikit-learn para algunas tareas de evaluación y transformación.

Statsmodels se utilizó en la familia ARIMA/SARIMA. Prophet y TimesFM necesitaron sus propias dependencias. Para las series difusas y los métodos de demanda intermitente se prepararon rutinas específicas que ayudaron a controlar los parámetros y conservar los resultados.

Los cuadernos se organizaron por etapas. La limpieza, el análisis exploratorio y los modelos no quedaron mezclados dentro de un solo archivo. Esta separación hizo más sencilla la revisión de errores y evitó repetir tareas ya completadas.

Cada ejecución generó archivos de salida. Se guardaron las series, las divisiones de entrenamiento y prueba, los pronósticos, las métricas, las figuras y la configuración utilizada. Así fue posible regresar a un resultado y revisar cómo se había obtenido.

Las herramientas se utilizaron como apoyo para ejecutar y documentar el análisis. Las decisiones sobre categorías, unidades y tratamiento de los registros no fueron delegadas automáticamente a los programas.

3.10 ALCANCES METODOLÓGICOS Y POSIBILIDAD DE REUTILIZACIÓN

El procedimiento fue probado con una sola categoría. Esto permitió revisar cada etapa con detalle, pero también limita el alcance de los resultados. El modelo que funcione mejor para LP_RE02 no necesariamente será el más adecuado para otro rubro.

La calidad del análisis depende de la información disponible. La limpieza puede corregir diferencias conocidas, pero no puede recuperar datos que nunca fueron registrados. Una descripción general puede seguir limitando la identificación exacta de una compra.

Los modelos también tienen límites. Algunos pueden responder bien en promedio y fallar en los picos. Otros pueden producir una trayectoria visualmente razonable y conservar errores elevados. Por eso, la selección consideró tanto las métricas como el comportamiento de la serie.

SBA y SBJ se añadieron porque LP_RE02 tenía varios periodos en cero. Su inclusión no significa que deban aplicarse a todas las categorías. Antes de utilizarlos es necesario revisar la frecuencia de los movimientos y la variabilidad de los valores positivos.

El procedimiento puede repetirse con otras categorías siempre que exista un historial suficiente y una clasificación consistente. La categoría puede cambiarse sin modificar la secuencia principal: revisar los datos, construir la serie, separar entrenamiento y prueba, ejecutar los modelos y comparar los resultados.

Esto no elimina el trabajo de revisión. Cada categoría debe analizarse antes de modelarse. Puede ser necesario cambiar la frecuencia, ajustar parámetros o excluir métodos que no correspondan con su comportamiento.

El valor metodológico del trabajo está en el orden seguido. La tesis no comienza con el pronóstico. Primero revisa la información que lo alimenta. Después organiza los registros y, solo entonces, compara los modelos en una categoría real de compras indirectas.

CAPÍTULO 4

RESULTADOS

Este capítulo presenta lo que se encontró al revisar la base OC_QA y al comparar los modelos aplicados a la categoría piloto LP_RE02. Primero se muestran los problemas detectados en los registros y los cambios logrados con la estandarización. Después se describe la serie semanal utilizada en el experimento. La última parte reúne los resultados de Prophet, ARIMA/SARIMA, FTS-Chen, TimesFM, SBA y SBJ bajo un mismo horizonte de prueba.

El análisis comenzó con los datos, no con los modelos. Esta decisión fue necesaria porque las diferencias en las descripciones, las unidades de medida y las categorías afectaban directamente la posibilidad de estudiar el consumo por rubro. Una vez organizada la información, fue posible construir una serie temporal y evaluar los pronósticos bajo las mismas condiciones.

4.1 RESULTADOS DEL ANÁLISIS EXPLORATORIO

4.1.1 COBERTURA DE LA CATEGORIZACIÓN

La revisión inicial mostró que el campo original de categoría tenía una cobertura limitada. Antes de la preparación, solo el 42.54 % de los registros tenía información en ese campo. Después de la limpieza, la cobertura original aumentó a 45.47 %, pero siguió siendo insuficiente para analizar toda la base por rubro.

La jerarquía creada durante la investigación alcanzó una cobertura del 100 % en los tres niveles utilizados: macrocategoría, categoría LP_N1 y categoría LP_N2. La comparación se presenta en la Figura 4.1.

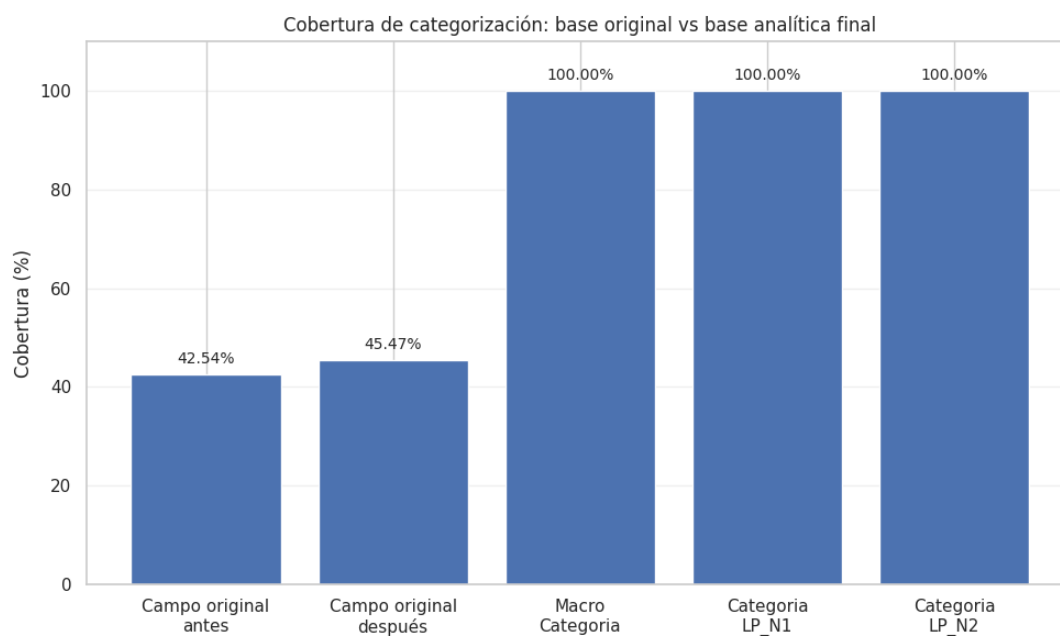


FIGURA 4.1: Cobertura de categorización antes y después de la preparación de OC_QA

Fuente: elaboración propia.

La mejora no provino de completar el campo anterior de manera automática. Se construyeron tres variables nuevas que ayudaron a ubicar cada registro dentro de

una estructura común. Esto hizo posible agrupar el consumo con mayor detalle y conservar una relación entre la transacción original y la categoría asignada.

El resultado coincide con el propósito de la clasificación del gasto descrito por Guida *et al.* (2025): convertir transacciones dispersas en información que pueda utilizarse para analizar compras. La estructura también facilita una lectura por categoría, como plantea Fontes *et al.* (2025).

4.1.2 ESTANDARIZACIÓN DE LAS UNIDADES DE MEDIDA

La base original contenía 116 formas distintas de registrar las unidades de medida. Después de la homologación quedaron 10 unidades estandarizadas. La reducción fue de 91.38 %.

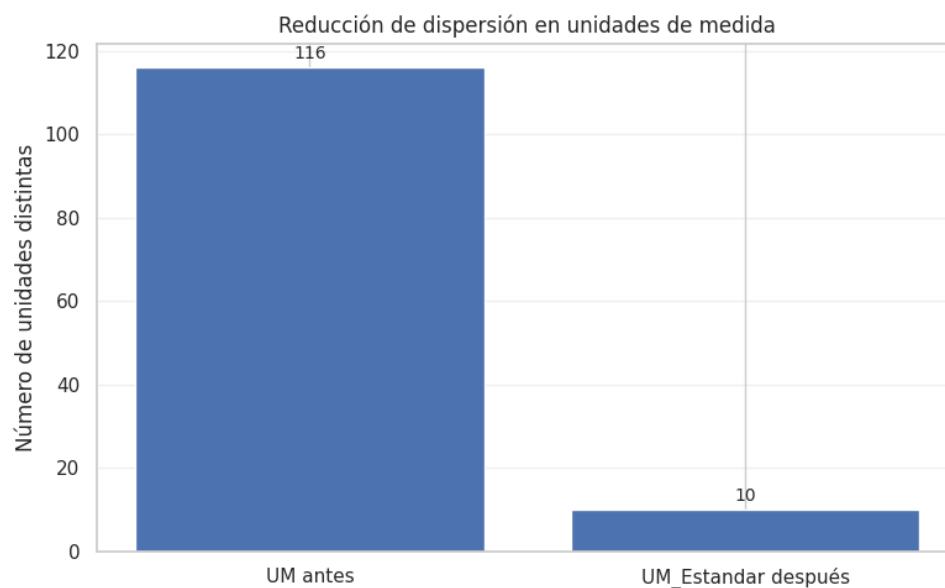


FIGURA 4.2: Reducción de unidades de medida diferentes

Fuente: elaboración propia.

La reducción evitó tratar como diferentes unidades que representaban lo mismo. También facilitó la comparación entre registros y disminuyó el trabajo de corrección necesario antes de cada análisis.

Este resultado muestra una aplicación concreta de la gobernanza del dato. Las reglas de homologación ayudaron a conservar una forma común de registrar la información, en lugar de corregir las mismas diferencias cada vez que se utilizaba la base (Bernardo *et al.*, 2024).

4.1.3 COMPORTAMIENTO MENSUAL POR CATEGORÍA

La Figura 4.3 presenta el comportamiento mensual de cinco categorías. Para facilitar la lectura se excluyeron únicamente los valores superiores al percentil 99, cuyo límite fue de 10 500 unidades. Esta decisión se aplicó a la visualización y no a la serie LP_RE02 utilizada posteriormente en los modelos.

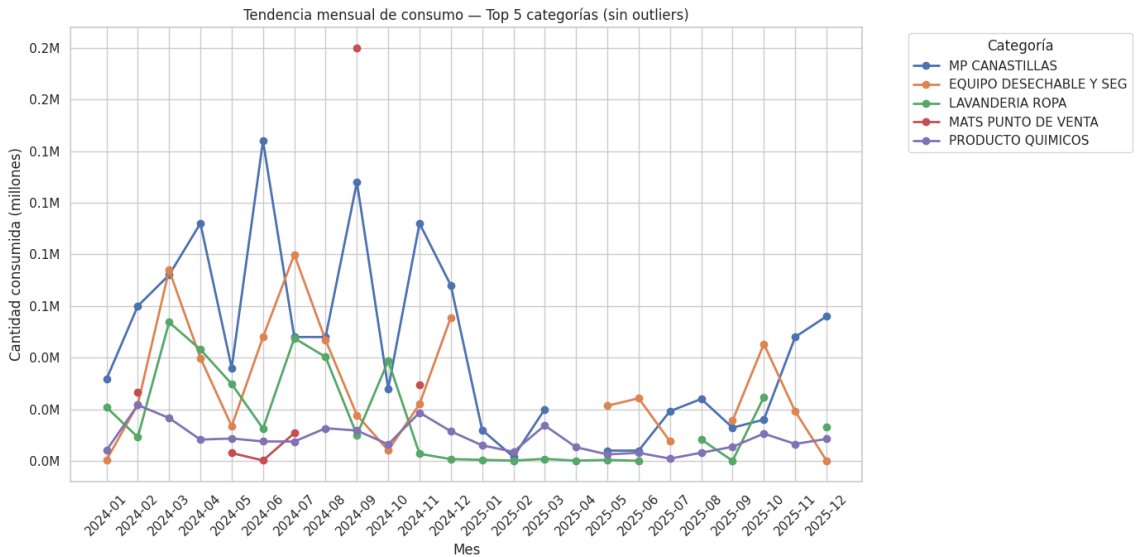


FIGURA 4.3: Consumo mensual de cinco categorías sin valores superiores al percentil 99

Fuente: elaboración propia.

Las categorías no siguieron una trayectoria uniforme. MP Canastillas presentó varios picos durante 2024 y volvió a crecer al cierre de 2025. Equipo Desechable y Seguridad registró movimientos de menor magnitud, pero con cambios frecuentes. Lavandería Ropa disminuyó después de 2024 y permaneció en niveles bajos durante

buena parte de 2025.

MATS PUNTO DE VENTA concentró un movimiento cercano a 200 mil unidades en septiembre de 2024. El valor se conservó porque la gráfica busca mostrar el comportamiento registrado. Su presencia no demuestra por sí sola que exista un error. Para clasificarlo sería necesario revisar la transacción y su contexto operativo.

4.1.4 CONCENTRACIÓN Y CONSUMO ACUMULADO

El Pareto de la Figura 4.4 se construyó con las cinco categorías seleccionadas después de excluir SIN_CATEGORIA y MATS PUNTO DE VENTA. El porcentaje acumulado corresponde únicamente a esas cinco categorías, no al total completo de OC_QA.

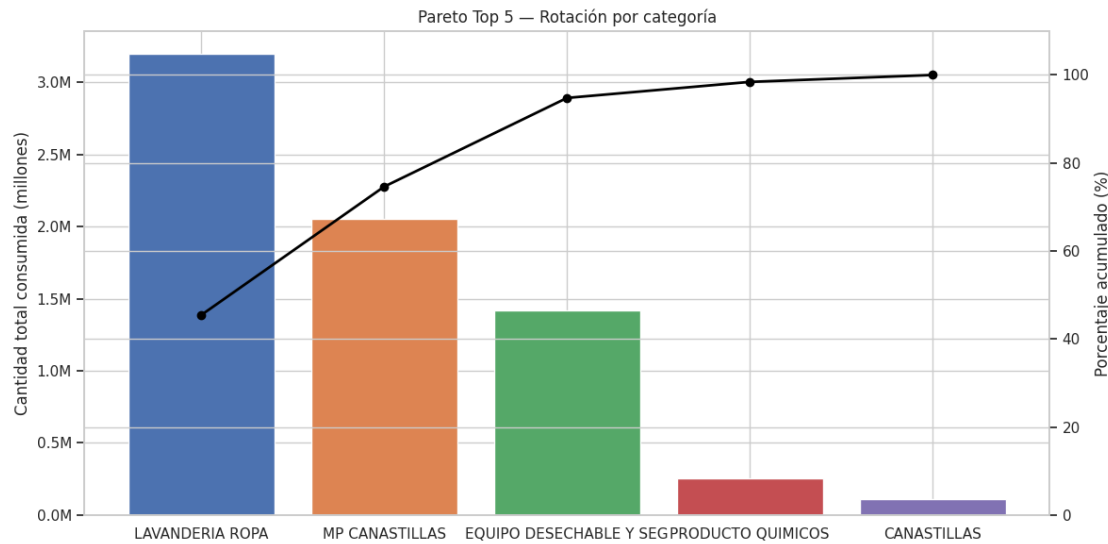


FIGURA 4.4: Distribución del consumo dentro de las cinco categorías seleccionadas

Fuente: elaboración propia.

Lavandería Ropa concentró el 45.45 % del consumo mostrado. Al agregar MP Canastillas, el porcentaje llegó a 74.62 %. Las tres primeras categorías reunieron 94.78 % del consumo de este grupo.

La concentración ayuda a reconocer qué rubros tienen mayor peso dentro del conjunto analizado. No significa que las categorías restantes carezcan de importancia. Indica que una parte considerable del volumen se encuentra en pocos grupos y que su seguimiento puede requerir mayor atención.

La Figura 4.5 muestra la suma acumulada del consumo. Por esta razón, se interpreta como consumo acumulado y no como una tasa de crecimiento.

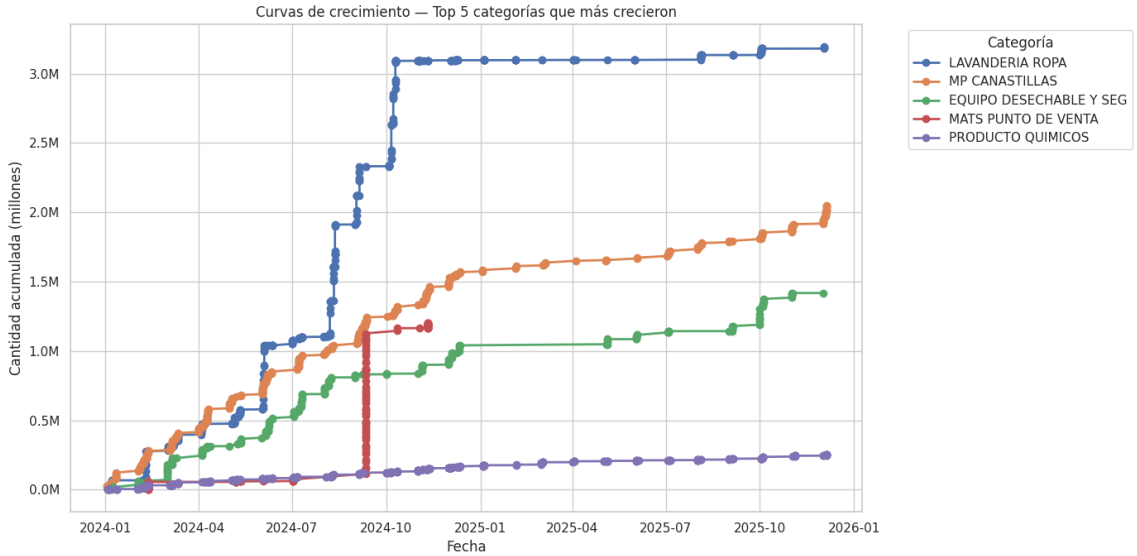


FIGURA 4.5: Consumo acumulado de cinco categorías

Fuente: elaboración propia.

Lavandería Ropa alcanzó el mayor volumen acumulado, con un incremento marcado durante 2024 y movimientos menores en 2025. MP Canastillas mostró una trayectoria más continua. Equipo Desechable y Seguridad avanzó de forma gradual, con nuevos incrementos al final del periodo.

Los saltos verticales de MATS PUNTO DE VENTA corresponden a varios registros concentrados en fechas cercanas. La figura ayuda a reconocer esa concentración, pero no debe utilizarse por sí sola para describir una tendencia estable.

4.1.5 DESCRIPCIONES DE ARTÍCULOS Y SERVICIOS

La revisión a nivel artículo encontró un problema adicional. El registro con mayor cantidad acumulada fue SIN_DESCRIPCION, con 3 251 630 unidades. Después aparecieron Polietileno alta densidad molido comprado, con 2 051 060, y Cubre-bocas blancos desechables, con 1 478 000.

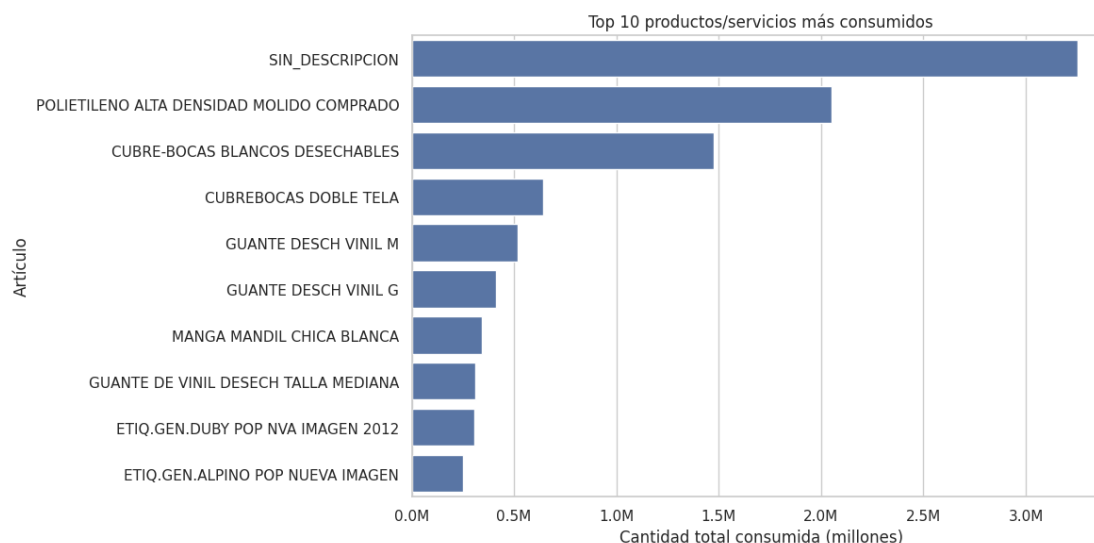


FIGURA 4.6: Diez productos o servicios con mayor cantidad acumulada

Fuente: elaboración propia.

El primer lugar no corresponde a un artículo identificable. Agrupa transacciones cuya descripción no permitía saber con claridad qué se había comprado. Esto limita el análisis por artículo, incluso cuando otros campos ayudan a recuperar la categoría.

La norma ISO/IEC 25012 relaciona la calidad del dato con su utilidad para el uso previsto (International Organization for Standardization y International Electrotechnical Commission, 2008). En OC_QA, una descripción incompleta puede ser suficiente para registrar una compra, pero no para analizarla con detalle después.

4.1.6 VALORES ATÍPICOS EN LA CANTIDAD

La variable Cantidad presentó una distribución muy concentrada cerca de valores pequeños y unos pocos extremos de gran magnitud. La mediana fue de 1 unidad. El primer cuartil también fue 1 y el tercero fue 8. El mínimo registrado fue de $-86\,407.88$ y el máximo de $180\,000$.

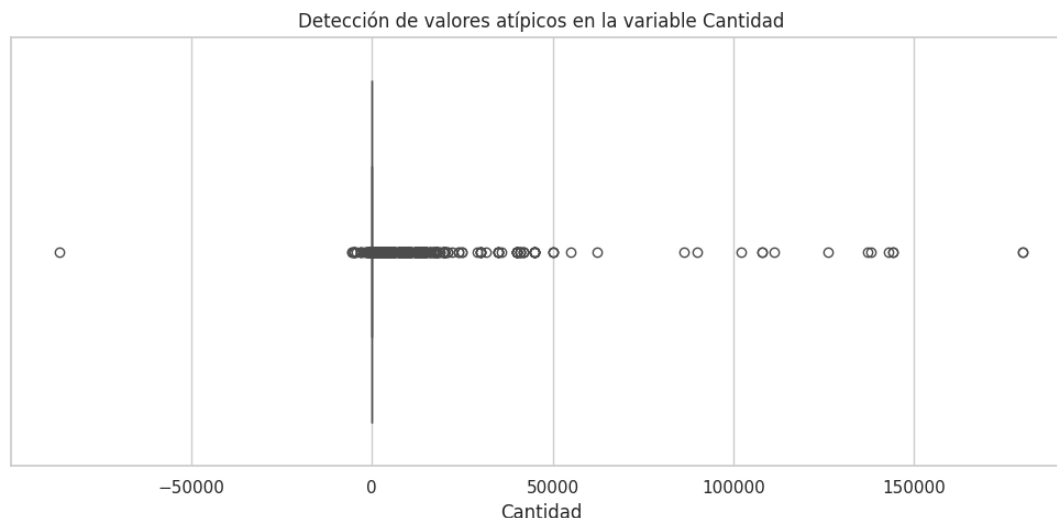


FIGURA 4.7: Distribución y valores atípicos de la variable Cantidad

Fuente: elaboración propia.

La escala queda dominada por los valores extremos y comprime la caja central. El valor negativo puede corresponder a una devolución, un ajuste, una cancelación o un error de captura. No se eliminó de forma automática porque su tratamiento depende del origen del movimiento.

La misma precaución se aplicó a los valores positivos altos. Un consumo grande puede ser excepcional y seguir siendo real. La revisión de atípicos se utilizó para reconocer casos que necesitaban contexto, no para borrar observaciones únicamente por su magnitud.

En conjunto, el EDA mostró que el problema de OC_QA no se reducía a elegir un modelo de pronóstico. Primero había que resolver diferencias en la

captura, completar una jerarquía de categorías y conservar el vínculo con el registro original. Estos resultados responden a los dos primeros objetivos específicos de la investigación.

4.2 PREPARACIÓN DE LA SERIE LP_RE02

La serie piloto se construyó mediante la suma semanal del consumo de LP_RE02. Las semanas sin movimiento se completaron con cero. No se aplicó ningún recorte de valores atípicos a la serie utilizada por los modelos.

El periodo total va del 7 de enero de 2024 al 7 de diciembre de 2025. La serie contiene 101 semanas. Las primeras 80 se utilizaron para entrenamiento y las últimas 21 para prueba. El conjunto de prueba comienza el 20 de julio de 2025 y termina el 7 de diciembre de 2025.

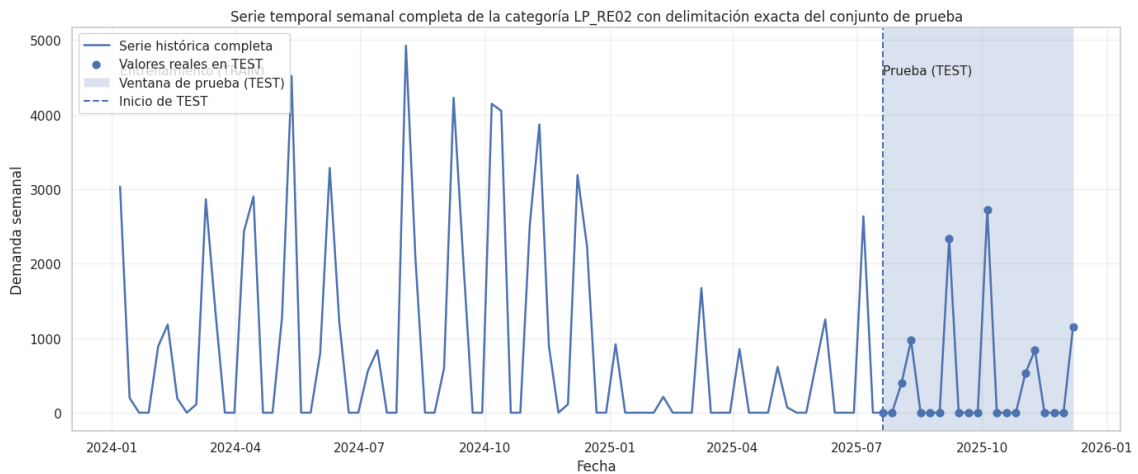


FIGURA 4.8: Serie semanal y partición temporal de LP_RE02

Fuente: elaboración propia.

La proporción de semanas con consumo igual a cero fue de 54.46 %. Esto confirma que la categoría no tiene una demanda continua. Los movimientos positivos aparecen separados por semanas sin consumo y presentan diferencias considerables de magnitud.

Todos los modelos utilizaron las mismas 21 semanas de prueba. Esta condición es importante porque la comparación final depende de que cada pronóstico se enfrente a las mismas observaciones reales.

4.3 RESULTADOS DE PROPHET

Prophet se utilizó como primera referencia. Se evaluaron 64 combinaciones de parámetros para la variante base y se preparó una segunda versión con crecimiento logístico.

La configuración base seleccionada utilizó un valor de 0.2 para *changepoint prior scale*, 0.8 para *changepoint range*, estacionalidad aditiva y un valor de 5 para *seasonality prior scale*. No se activaron componentes semanales ni anuales. Hubo configuraciones empatadas con las mismas métricas, por lo que la combinación conservada no debe entenderse como la única solución posible.

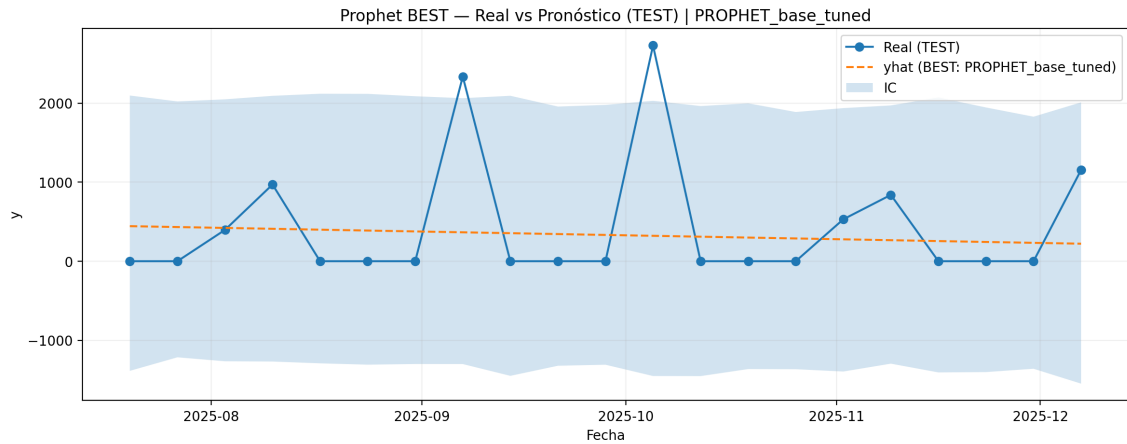


FIGURA 4.9: Pronóstico de Prophet base para LP_RE02

Fuente: elaboración propia.

El pronóstico base descendió de forma gradual a lo largo de las 21 semanas. La trayectoria fue estable, pero no reaccionó a los dos picos principales del conjunto de prueba. También mantuvo valores positivos durante las semanas con consumo igual a cero.

Las bandas corresponden a intervalos de incertidumbre del 80 %. Su amplitud refleja la dificultad para representar una serie con muchos ceros y repuntes aislados. El límite inferior fue negativo, por lo que la ausencia de valores negativos solo aplica al pronóstico puntual.

La variante logística utilizó un límite inferior de cero y un límite superior de 5 910 unidades, equivalente a 1.2 veces el máximo observado en entrenamiento.

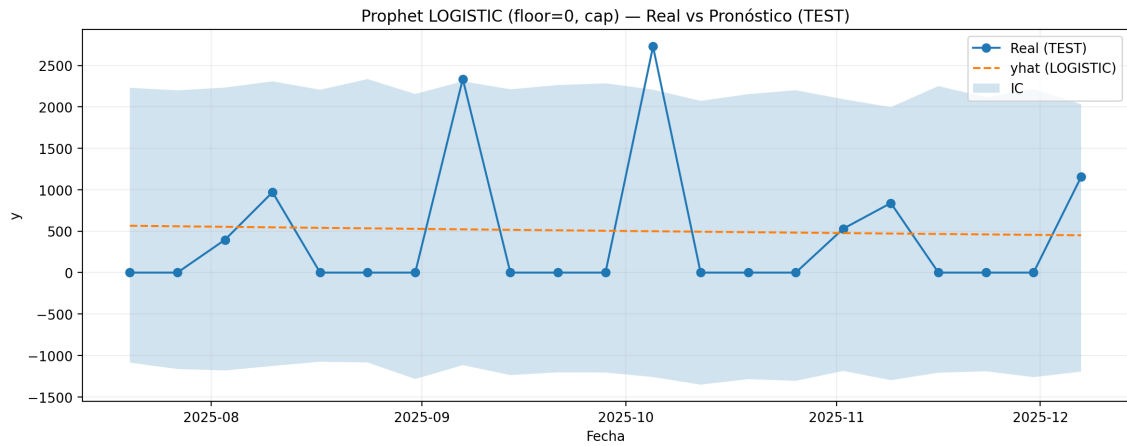


FIGURA 4.10: Pronóstico de Prophet logístico para LP_RE02

Fuente: elaboración propia.

La restricción logística elevó el nivel central del pronóstico. Esto redujo ligeramente el RMSE porque acercó la estimación a algunos consumos positivos. Al mismo tiempo, aumentó el error en las semanas sin movimiento.

TABLA 4.1: Resultados de las variantes de Prophet

Modelo	MAE	RMSE	WAPE (%)	sMAPE (%)	Neg.
Base ajustado	543.81	783.42	127.54	166.34	0
Logístico	612.02	778.20	143.54	157.54	0

La variante logística obtuvo un RMSE ligeramente menor. La base presentó menor MAE y menor WAPE. Por esa razón se conservó como representante de

Prophet en la comparación integral.

El resultado muestra una limitación clara. Prophet produjo una trayectoria interpretable y evitó valores centrales negativos, pero suavizó demasiado la serie. Durante las 14 semanas con consumo cero siguió pronosticando cantidades positivas y no logró aproximarse a los picos de septiembre y octubre.

4.4 RESULTADOS DE ARIMA/SARIMA

La familia ARIMA/SARIMA también se evaluó mediante 64 configuraciones. La selección siguió un orden definido previamente: primero se descartaron alternativas con pronósticos negativos y después se compararon RMSE, MAE y WAPE.

La Figura 4.11 documenta la separación temporal utilizada por esta familia. Esta partición es la misma presentada en la Figura 4.8.

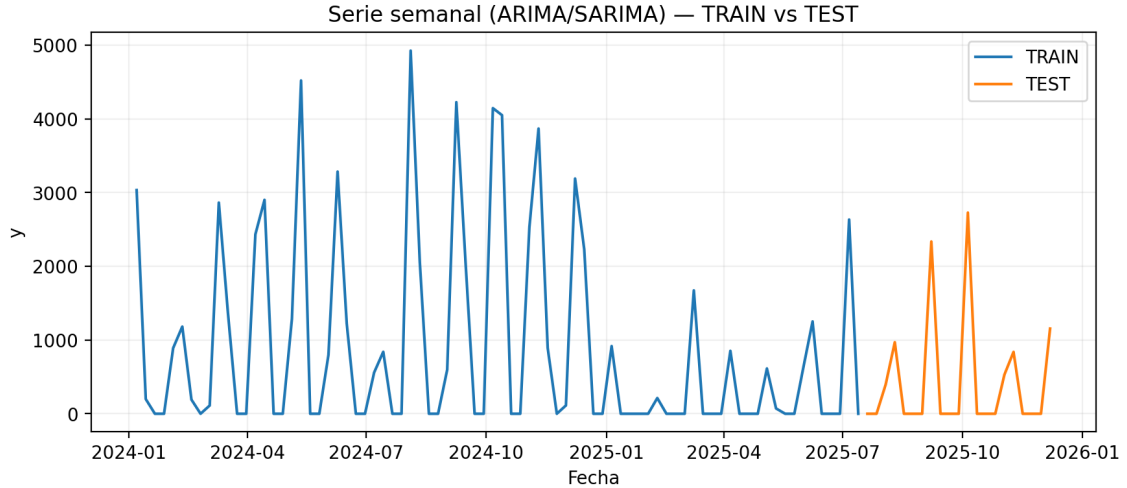


FIGURA 4.11: Entrenamiento y prueba utilizados en ARIMA/SARIMA

Fuente: elaboración propia.

La configuración elegida fue:

$$\text{SARIMA}(0, 0, 0) \times (1, 0, 0, 52)$$

El modelo no incorporó componentes no estacionales. La parte estacional utilizó un término autorregresivo con rezago de 52 semanas. El AIC fue de 440.42, pero este valor se mantuvo como referencia diagnóstica. No fue el criterio principal de selección.

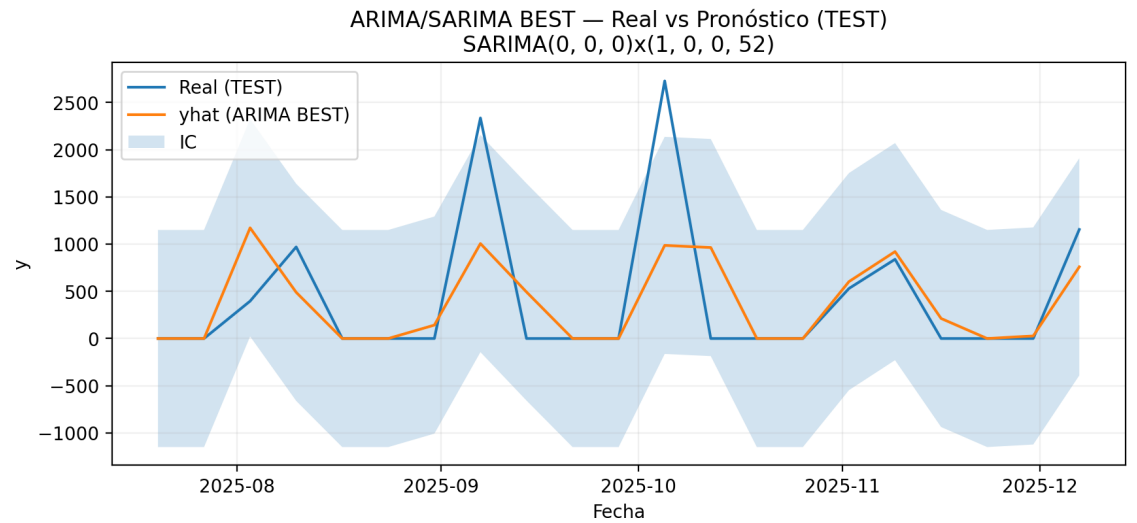


FIGURA 4.12: Pronóstico de SARIMA para LP_RE02

Fuente: elaboración propia.

SARIMA reaccionó con mayor claridad ante varios movimientos positivos. Se acercó especialmente a los consumos observados en noviembre y parte de diciembre. Los picos más altos siguieron siendo difíciles. Para el consumo de 2 336 unidades de septiembre, el modelo pronosticó cerca de 1 006. Para el pico de 2 728 unidades de octubre, la estimación quedó cerca de 986.

TABLA 4.2: Resultados de SARIMA para LP_RE02

Modelo	MAE	RMSE	WAPE (%)	sMAPE (%)	AIC	Neg.
SARIMA(0,0,0) × (1,0,0,52)	319.82	579.11	75.01	66.75	440.42	0

Los intervalos de predicción corresponden al 95 %. Cubrieron 19 de las 21 observaciones reales. Los dos valores fuera de la banda fueron los picos de septiembre y octubre.

El pronóstico central no produjo valores negativos. El límite inferior del intervalo sí fue negativo en casi todo el horizonte. Esto indica que la coherencia de no negatividad no se extiende a toda la distribución estimada.

La prueba de Ljung-Box obtuvo un valor p cercano a 1.49×10^{-16} . El resultado indica que quedó autocorrelación en los residuos. Por eso, el modelo no debe interpretarse como una representación completa de la serie. Fue la mejor alternativa entre las configuraciones evaluadas por su desempeño fuera de muestra.

El componente de 52 semanas también requiere cautela. El entrenamiento contiene 80 observaciones, por lo que cubre un ciclo anual completo y una parte del siguiente. La relación estacional fue útil para el pronóstico, pero el historial todavía es corto para afirmar que existe una estacionalidad anual estable.

4.5 RESULTADOS DE FTS-CHEN

La implementación difusa utilizó el método FTS-Chen con siete intervalos iguales y reglas de primer orden. Esta configuración fue definida directamente. No se realizó una búsqueda entre diferentes cantidades de intervalos u otros métodos de partición.

Los siete intervalos cubrieron el rango entre 0 y 4925 unidades. El primer intervalo tuvo un punto medio de 351.79. La última observación del entrenamiento fue cero y quedó asociada con el primer estado difuso.

La transición más frecuente desde ese estado fue $A_0 \rightarrow A_0$. Apareció en 34 de 50 salidas, equivalente al 68 %. Como el pronóstico se generó de forma recursiva, el modelo permaneció en el mismo estado durante las 21 semanas. Por eso, todos los valores estimados fueron de 351.79 unidades.

La salida constante evitó valores negativos, pero no distinguió entre semanas sin consumo y semanas con picos. En las 14 semanas con cero, sobreestimó la demanda.

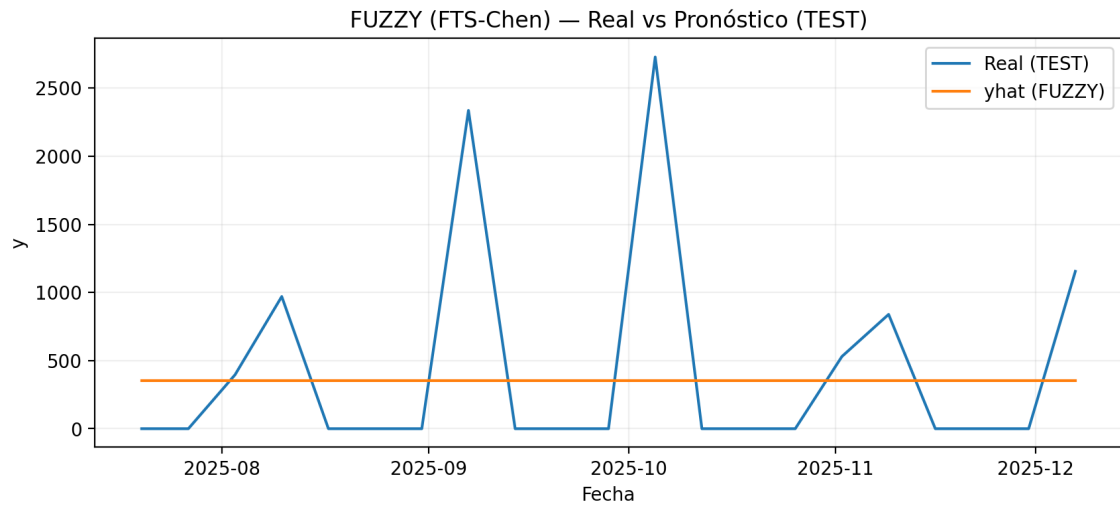


FIGURA 4.13: Pronóstico FTS-Chen para LP_RE02

Fuente: elaboración propia.

TABLA 4.3: Resultados de FTS-Chen para LP_RE02

Modelo	MAE	RMSE	WAPE (%)	sMAPE (%)	Neg.
FTS-Chen EQ7	543.64	775.04	127.50	163.63	0

En los movimientos de 2 336 y 2 728 unidades, la estimación siguió en 351.79.

La implementación no generó intervalos de predicción. El resultado se interpretó mediante la trayectoria central, las reglas construidas y las métricas de error.

4.6 RESULTADOS DE TIMESFM

TimesFM se ejecutó con el modelo preentrenado ‘google/timesfm-2.5-200m-pytorch’, de 200 millones de parámetros. Se utilizó una configuración fija con normalización de entradas y un horizonte de 21 semanas. No se compararon otros checkpoints ni tamaños del modelo.

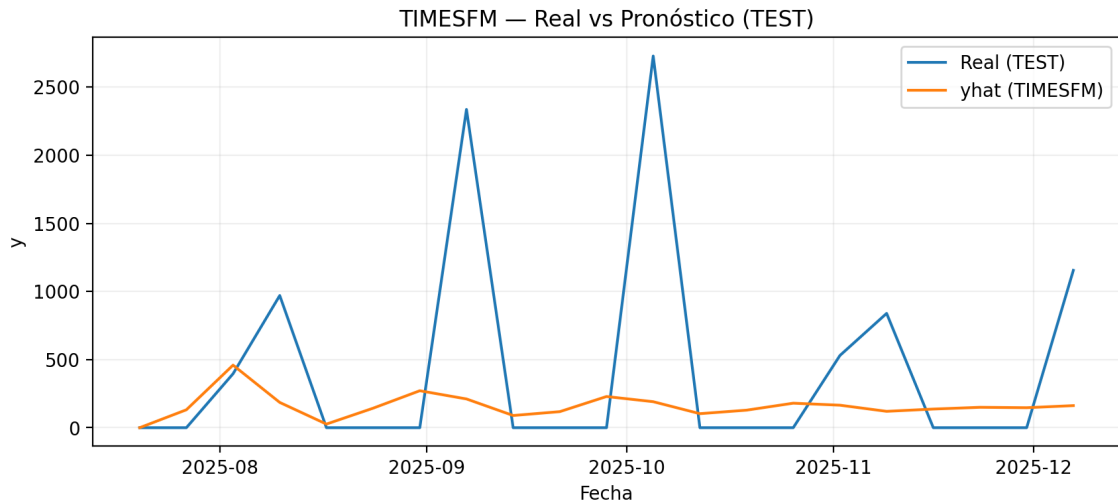


FIGURA 4.14: Pronóstico de TimesFM para LP_RE02

Fuente: elaboración propia.

La trayectoria tuvo más movimiento que FTS-Chen, pero permaneció en niveles bajos. En las semanas con consumo cero, el pronóstico promedio fue de 132.36 unidades. En las siete semanas con demanda positiva, el valor real promedio fue cercano a 1 279, mientras que TimesFM pronosticó cerca de 213.

Los errores más grandes aparecieron en septiembre y octubre. Frente a 2 336 unidades reales, el modelo estimó 210.98. Para el pico de 2 728, produjo 190.84.

TABLA 4.4: Resultados de TimesFM para LP_RE02

Modelo	MAE	RMSE	WAPE (%)	sMAPE (%)	Neg.
TimesFM 2.5, 200M	449.51	802.28	105.42	166.52	0

TimesFM obtuvo un MAE menor que Prophet, FTS-Chen, SBA y SBJ en la comparación común. Su RMSE fue el más alto. Esto ocurre porque el RMSE da mayor peso a los errores grandes y el modelo se alejó considerablemente de los dos picos principales.

La configuración tenía activa una salida de cuantiles, pero el archivo exportado conservó únicamente el pronóstico puntual. Por ello, no se utilizaron intervalos de TimesFM en la evaluación.

4.7 RESULTADOS DE SBA Y SBJ

Croston se utilizó como antecedente de los métodos especializados para demanda intermitente. En la implementación se evaluaron dos variantes derivadas de ese enfoque: SBA y SBJ.

La serie completa contiene 55 semanas con demanda cero y 46 con demanda positiva. El intervalo promedio entre movimientos, identificado como ADI, fue de 2.1957. El valor de CV^2 fue de 0.6062.

Con los umbrales empleados en el notebook, ADI de 1.32 y CV^2 de 0.49, LP_RE02 quedó clasificada como una serie *lumpy*. Esta categoría combina ocurrencias separadas y tamaños de demanda variables.

Para calibrar los métodos se probaron 19 valores de α , desde 0.05 hasta 0.95. La selección utilizó las últimas seis semanas de la serie. SBA obtuvo su menor RMSE

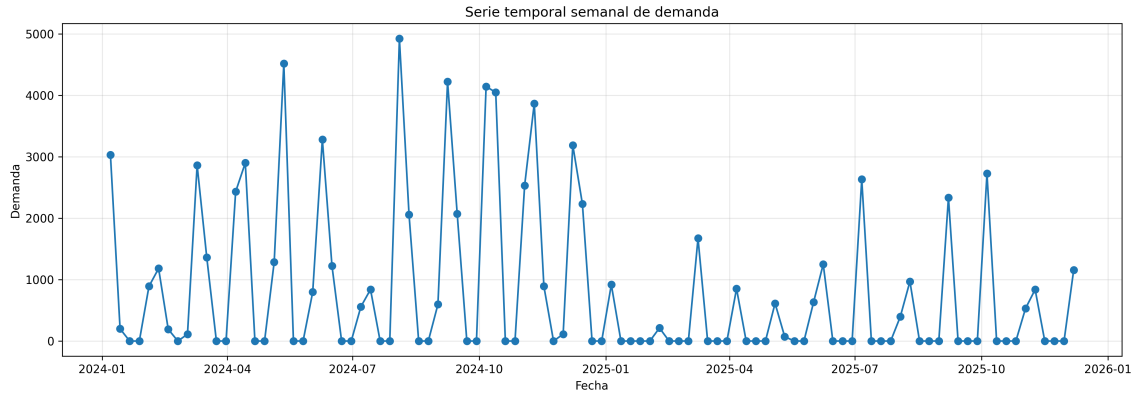


FIGURA 4.15: Serie semanal utilizada para caracterizar la demanda intermitente

Fuente: elaboración propia.

con $\alpha = 0.55$ y SBJ con $\alpha = 0.40$.

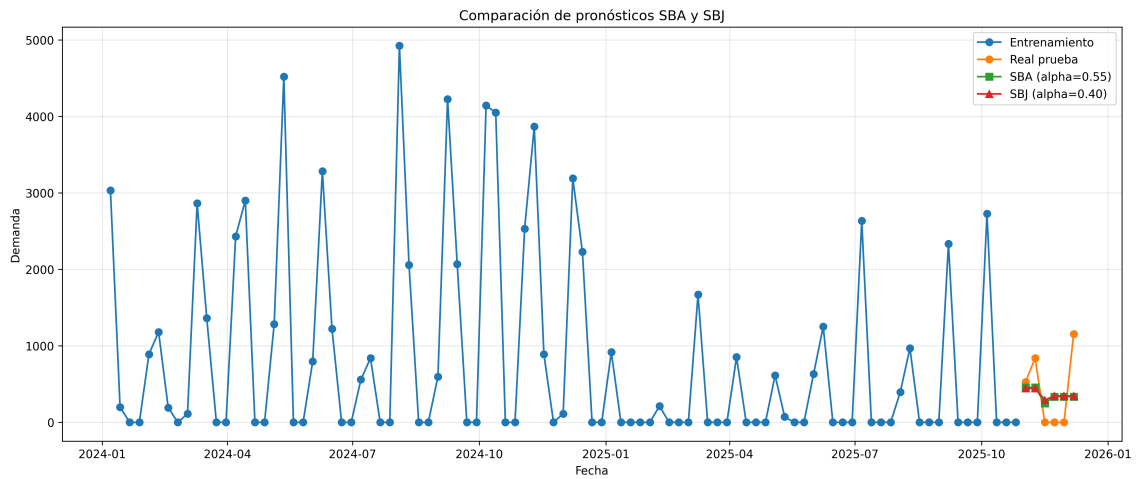


FIGURA 4.16: Calibración de SBA y SBJ en las últimas seis semanas

Fuente: elaboración propia.

Estas métricas corresponden a la ventana de seis semanas usada para elegir α . No deben compararse directamente con los resultados de Prophet, SARIMA, FTS-Chen o TimesFM.

Después de fijar los parámetros, SBA y SBJ se ejecutaron sobre las 21 semanas comunes. Ambos produjeron niveles relativamente estables y evitaron valores negativos. En las semanas con consumo cero mantuvieron pronósticos positivos. En los dos picos principales, sus estimaciones quedaron entre 330 y 400 unidades.

TABLA 4.5: Métricas utilizadas para calibrar SBA y SBJ

Modelo	α	MAE	RMSE	sMAPE (%)	WAPE (%)
SBA	0.55	366.11	430.18	130.61	87.03
SBJ	0.40	378.80	436.39	131.11	90.05

SBA tuvo mejor RMSE en la calibración y en la prueba común. SBJ terminó con un MAE y un WAPE ligeramente menores durante las 21 semanas. La diferencia entre ambos fue pequeña y ninguno superó a SARIMA en la evaluación final.

4.8 COMPARACIÓN INTEGRAL DE LOS MODELOS

La comparación final se realizó directamente con los pronósticos alineados sobre las mismas 21 semanas, del 20 de julio al 7 de diciembre de 2025. Esta salida sustituye cualquier comparación preliminar basada en ventanas diferentes.

El criterio de selección siguió este orden:

1. Menor cantidad de pronósticos negativos.
2. Menor RMSE.
3. Menor MAE.
4. Menor WAPE.

Todos los modelos produjeron valores centrales no negativos. La selección se definió entonces por el RMSE y las métricas siguientes.

MAPE no se utilizó porque la serie contiene numerosas observaciones iguales a cero. En esas condiciones, la métrica puede crecer de forma desproporcionada y

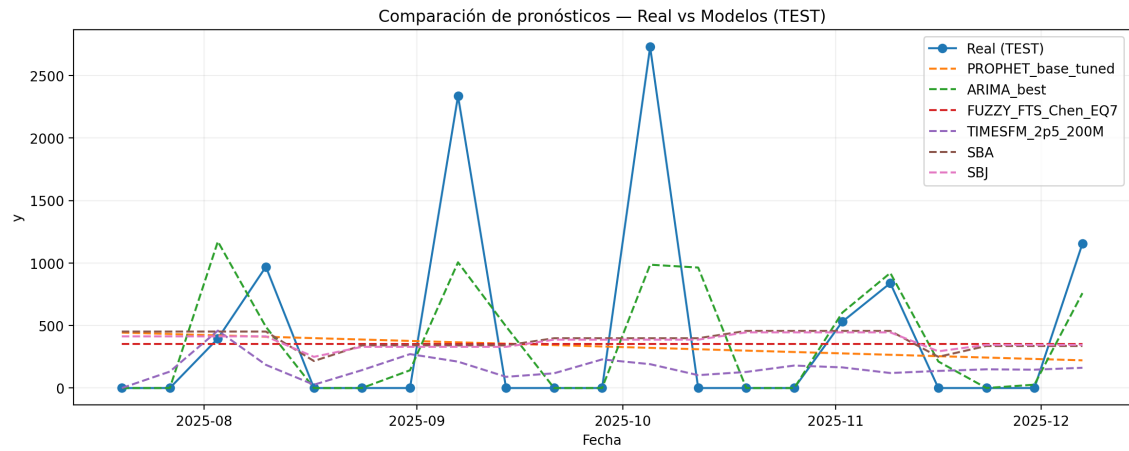


FIGURA 4.17: Comparación de modelos en las 21 semanas de prueba

Fuente: elaboración propia.

TABLA 4.6: Comparación final de modelos para LP_RE02

Modelo	MAE	RMSE	WAPE (%)	sMAPE (%)	Neg.
ARIMA/SARIMA	319.82	579.11	75.01	66.75	0
SBA	540.28	771.10	126.71	160.37	0
SBJ	538.54	772.59	126.31	160.59	0
FTS-Chen	543.64	775.04	127.50	163.63	0
Prophet	543.81	783.42	127.54	166.34	0
TimesFM	449.51	802.28	105.42	166.52	0

dejar de representar el desempeño general. sMAPE se conservó como información complementaria.

SARIMA obtuvo el menor RMSE, el menor MAE y el menor WAPE. También fue el modelo que respondió con mayor claridad a varios movimientos positivos. No consiguió anticipar completamente los dos picos principales y sus residuos conservaron autocorrelación. Aun con esas limitaciones, fue la alternativa con mejor desempeño fuera de muestra para LP_RE02.

TimesFM registró el segundo menor MAE, pero el RMSE más alto. Su trayectoria fue razonable en varias semanas y se alejó demasiado en los eventos de mayor magnitud.

SBA y SBJ ofrecieron niveles estables, pero esa estabilidad produjo sobreestimaciones durante los ceros y subestimaciones en los picos. FTS-Chen mantuvo un pronóstico constante. Prophet generó una trayectoria descendente que tampoco reaccionó a los cambios abruptos.

La selección de SARIMA corresponde únicamente a la categoría piloto y al horizonte evaluado. No implica que el mismo modelo deba utilizarse en todas las compras indirectas. Cada categoría puede tener una combinación distinta de continuidad, intermitencia, volumen e historial disponible.

El resultado principal está en el procedimiento completo. La comparación comenzó después de ordenar la información, construir la categoría y preparar una serie común. Bajo esas condiciones, fue posible evaluar métodos diferentes sin cambiar las fechas ni la variable observada. Para LP_RE02, SARIMA ofreció el mejor equilibrio entre precisión y coherencia operativa.

CAPÍTULO 5

CONCLUSIONES

5.1 CONCLUSIONES GENERALES

Al comenzar esta investigación, la empresa ya contaba con información histórica de sus compras indirectas. El problema era que esa información no estaba lista para utilizarse. Las descripciones no siempre ayudaban a reconocer lo comprado, las unidades de medida se registraban de distintas formas y el campo original de categoría dejaba una parte considerable de la base sin clasificar.

Por esta razón, el primer resultado del trabajo no fue un pronóstico. Fue una base más ordenada. La cobertura del campo original de categoría era de 42.54 % y, después de la preparación, llegó a 45.47 %. La mejora más amplia se obtuvo con la nueva jerarquía, ya que macrocategoría, LP_N1 y LP_N2 alcanzaron una cobertura del 100 %.

También se redujo la dispersión de las unidades de medida. La base original contenía 116 denominaciones distintas y, después de la homologación, quedaron 10. Con este cambio fue más sencillo comparar los registros y evitar que unidades equivalentes fueran tratadas como si representaran conceptos diferentes.

El análisis exploratorio reveló otros problemas que no podían ignorarse. El

registro con mayor cantidad acumulada correspondía a SIN_DESCRIPCION. También se encontraron cantidades negativas, valores extremos y una concentración importante del consumo en pocas categorías. Estos casos no se eliminaron de forma automática. Primero se consideró su posible relación con devoluciones, ajustes, cancelaciones o movimientos reales de la operación.

Después de preparar la información, se construyó la serie semanal de LP_RE02. La serie tuvo 101 observaciones. Las primeras 80 se utilizaron para entrenamiento y las últimas 21 para prueba. El 54.46 % de las semanas presentó consumo igual a cero. Este comportamiento confirmó que no se trataba de una demanda continua y justificó la comparación entre modelos generales y métodos desarrollados para series intermitentes.

Los modelos evaluados fueron Prophet, ARIMA/SARIMA, FTS-Chen, TimesFM, SBA y SBJ. Todos se compararon bajo las mismas condiciones y sobre las mismas 21 semanas, del 20 de julio al 7 de diciembre de 2025.

ARIMA/SARIMA presentó el mejor desempeño global para LP_RE02. El modelo obtuvo un MAE de 319.82, un RMSE de 579.11 y un WAPE de 75.01 %. También mantuvo cero pronósticos negativos. Ningún modelo logró anticipar por completo los dos picos más altos de la prueba, pero SARIMA reaccionó mejor que las demás alternativas ante varios movimientos positivos de la serie.

Este resultado debe entenderse dentro de los límites del caso. SARIMA fue el mejor modelo para LP_RE02 y para el periodo evaluado. Esto no significa que vaya a ser la mejor opción para todas las categorías de compras indirectas.

5.2 CUMPLIMIENTO DE LOS OBJETIVOS Y REVISIÓN DE LA HIPÓTESIS

El primer objetivo específico se atendió con el diagnóstico y la preparación de OC_QA. Se identificaron diferencias en la captura, campos incompletos, unidades heterogéneas, descripciones poco claras y valores que necesitaban revisión. La homologación redujo las unidades de medida de 116 a 10 y dejó una base con mejores condiciones para continuar el análisis.

El segundo objetivo se abordó mediante la categorización jerárquica. Con la estructura formada por macrocategoría, LP_N1 y LP_N2 se clasificó el total de los registros utilizados. Esto facilitó la lectura del consumo por rubro y la formación de grupos que antes no podían estudiarse de forma consistente.

El tercer objetivo se atendió mediante la comparación de los modelos. Prophet, SARIMA, FTS-Chen, TimesFM, SBA y SBJ recibieron la misma serie y el mismo horizonte de prueba. La selección final tomó en cuenta la cantidad de pronósticos negativos, el RMSE, el MAE y el WAPE.

Los datos respaldan la hipótesis planteada para el caso estudiado. La estandarización y la categorización mejoraron el seguimiento del consumo por rubro y facilitaron la construcción de una comparación común entre modelos.

Esta conclusión no debe interpretarse como una afirmación general para cualquier empresa o categoría. El estudio no realizó una prueba experimental para demostrar que un modelo sea superior en todos los escenarios. En OC_QA se encontró que el pronóstico pudo desarrollarse después de preparar la información, construir la jerarquía y formar una serie con criterios definidos.

5.3 APORTACIONES DEL TRABAJO

El trabajo dejó una base analítica construida a partir de registros que inicialmente tenían problemas de consistencia. La información original se conservó y las variables nuevas se añadieron sin perder el vínculo con la fuente.

También quedaron reglas para homologar unidades, limpiar descripciones y asignar categorías. Estas reglas pueden revisarse cuando aparezcan nuevos registros o cuando el área de compras necesite modificar la jerarquía.

Otro resultado fue la estructura de categorización formada por tres niveles. La cobertura completa no garantiza que cada asignación sea perfecta, pero ofrece un punto de partida más útil que el campo original. La revisión de los usuarios será necesaria para corregir casos particulares y mantener la clasificación vigente.

La investigación también produjo una serie semanal reproducible para LP_RE02. Las semanas sin movimiento no fueron eliminadas. Se completaron con cero para conservar el comportamiento observado y evitar que la demanda pareciera artificialmente continua.

Los cuadernos, scripts, archivos de configuración y resultados forman un flujo de trabajo que puede reutilizarse. No se construyó una plataforma única, pero sí un conjunto de herramientas organizado por etapas. Cada modelo generó pronósticos, métricas y archivos que después se reunieron dentro del comparativo final.

Otro resultado útil fue colocar todos los pronósticos sobre las mismas 21 semanas. Las primeras evaluaciones de algunos modelos se habían realizado con ventanas distintas. Al recalcular las métricas bajo un horizonte común fue posible tomar una decisión sin mezclar periodos ni condiciones de evaluación.

Esta distinción fue necesaria en SBA y SBJ. Durante la calibración de seis semanas, SBA obtuvo mejores métricas. En la evaluación final de 21 semanas, SBA

conservó un RMSE ligeramente menor, mientras que SBJ obtuvo un MAE y un WAPE un poco más bajos. La diferencia entre ambos fue pequeña y no permitió declarar un ganador absoluto dentro de esa familia para el horizonte final.

La comparación completa sí dejó una diferencia más clara. SARIMA obtuvo el menor MAE, RMSE y WAPE entre los seis modelos. Este resultado no elimina la utilidad de las demás alternativas. Ayuda a entender cómo respondió cada una frente a una serie corta, irregular y con una alta proporción de ceros.

5.4 IMPLICACIONES PRÁCTICAS Y LIMITACIONES

Los resultados pueden apoyar tareas concretas dentro del área de compras. La categorización ayuda a reconocer en qué rubros se concentra el consumo. La homologación evita que unidades equivalentes aparezcan como si fueran diferentes. La revisión de descripciones ayuda a localizar registros que todavía necesitan corrección.

La serie por categoría también ofrece otra forma de observar las necesidades. En lugar de revisar únicamente transacciones aisladas, el comprador puede seguir la frecuencia de los movimientos, reconocer periodos sin consumo y detectar semanas con cantidades fuera de lo habitual.

El pronóstico no debe convertirse en una orden automática de compra. Su función es ofrecer una referencia adicional para la planeación. La decisión final todavía necesita información que no estaba incluida en la serie, como inventario disponible, tiempos de entrega, cantidades mínimas, contratos, criticidad y conocimiento del área usuaria.

El estudio tiene varias limitaciones. La primera es que el flujo completo se probó con una sola categoría. Con LP_RE02 se revisó el procedimiento con detalle, pero su comportamiento no representa todas las compras indirectas.

La segunda limitación está relacionada con la longitud del historial. El entrenamiento tuvo 80 semanas. En el caso de SARIMA, la configuración seleccionada incorporó un rezago de 52 semanas, pero el historial solo cubre un ciclo anual completo y una parte del siguiente. El resultado fue útil para el horizonte evaluado, aunque todavía es pronto para afirmar que existe una estacionalidad anual estable.

La prueba de Ljung-Box indicó autocorrelación en los residuos de SARIMA. Esto significa que el modelo no capturó toda la estructura temporal. Se seleccionó porque obtuvo el mejor desempeño fuera de muestra entre las configuraciones evaluadas, no porque explicara por completo la serie.

Otra limitación se encuentra en la calidad de los registros. La limpieza puede corregir diferencias conocidas, pero no puede recuperar una descripción que nunca fue capturada. El caso de SIN_DESCRIPCION deja claro que todavía existen movimientos cuyo contenido no puede identificarse con suficiente detalle.

La jerarquía alcanzó cobertura completa, pero necesita validación con compradores y usuarios de la empresa. Una categoría asignada mediante reglas puede ser consistente desde el análisis y, aun así, necesitar un ajuste por conocimiento operativo.

El procedimiento tampoco se instaló de forma permanente dentro de la organización. Los resultados corresponden a una aplicación de tesis y no a una evaluación posterior a su puesta en marcha. Por eso, no fue posible medir si el uso del flujo reduce compras urgentes, faltantes, tiempos de análisis o costos.

5.5 TRABAJO FUTURO

El siguiente paso debería ser aplicar el procedimiento a más categorías. La comparación entre rubros con consumo continuo, intermitente, estacional o esporádico ayudaría a comprobar si SARIMA conserva el mejor desempeño o si otros modelos

se adaptan mejor.

También será necesario revisar la categorización con el área de compras. La participación de los usuarios ayudaría a confirmar los niveles creados, corregir registros dudosos y asignar responsables para el mantenimiento de la jerarquía.

Otra tarea pendiente es medir el efecto operativo. Un trabajo posterior podría comparar compras urgentes, faltantes, tiempos de análisis, frecuencia de pedido y niveles de inventario antes y después de utilizar el procedimiento. Con esa información se podría saber si la mejora analítica produce cambios reales en la operación.

La automatización puede desarrollarse después de validar estas etapas. Los scripts ya aceptan el cambio de categoría, pero todavía sería necesario integrar controles, reportes y revisiones antes de ejecutar el flujo de forma periódica.

También conviene probar otras configuraciones y métodos. FTS-Chen se aplicó con siete intervalos iguales y TimesFM con una configuración fija. Una evaluación más amplia podría revisar otras particiones difusas, diferentes modelos preentrenados, validación temporal cruzada y combinaciones de pronósticos.

5.6 REFLEXIÓN FINAL

El resultado más importante de esta tesis no fue encontrar un modelo ganador. Fue comprobar que la calidad del pronóstico estaba limitada desde antes por la forma en que se registraban las compras.

Al ordenar OC_QA se entendió qué información estaba disponible, qué partes necesitaban corrección y cuáles podían utilizarse para construir una serie. Solo después de ese trabajo fue posible comparar los modelos bajo las mismas condiciones.

Para LP_RE02, SARIMA ofreció el mejor resultado dentro del horizonte

analizado. La conclusión no termina con esa selección. En este caso, mejorar la decisión de compra comenzó por entender los registros y hacerlos utilizables. El pronóstico apareció después, como una herramienta de apoyo y no como sustituto del criterio operativo.

BIBLIOGRAFÍA

- BERNARDO, B. M. V., H. SÃO MAMEDE, J. M. P. BARROSO y V. M. P. DUARTE DOS SANTOS (2024), «Data governance & quality management—Innovation and breakthroughs across different fields», *Journal of Innovation & Knowledge*, **9**(4), 100598.
- CHAPMAN, P., J. CLINTON, R. KERBER, T. KHABAZA, T. REINARTZ, C. SHEARER y R. WIRTH (2000), «CRISP-DM 1.0: Step-by-step data mining guide», *Informe técnico*, SPSS Inc., URL <https://public.dhe.ibm.com/software/analytics/spss/documentation/modeler/14.2/es/CRISP-DM.pdf>.
- CHEN, S.-M. (1996), «Forecasting enrollments based on fuzzy time series», *Fuzzy Sets and Systems*, **81**(3), págs. 311–319.
- CHOI, T.-M., S. W. WALLACE y Y. WANG (2018), «Big data analytics in operations management», *Production and Operations Management*, **27**(10), págs. 1868–1883.
- CROSTON, J. D. (1972), «Forecasting and stock control for intermittent demands», *Journal of the Operational Research Society*, **23**(3), págs. 289–303.
- DAS, A., W. KONG, R. SEN y Y. ZHOU (2024), «A decoder-only foundation model for time-series forecasting», en *Proceedings of the 41st International Conference on Machine Learning, Proceedings of Machine Learning Research*, tomo 235, PMLR, págs. 10 148–10 167, URL <https://proceedings.mlr.press/v235/das24c.html>.

- DELKE, V., H. SCHIELE y W. BUCHHOLZ (2023), «Differentiating between direct and indirect procurement: Roles, skills, and Industry 4.0», *International Journal of Procurement Management*, **16**(1), págs. 1–30.
- DI MAURO, C., E. PETERS y S. CARNOVALE (2024), «A seat at the table: The future of purchasing and supply management», *Journal of Purchasing and Supply Management*, **30**(1), 100908.
- FONTES, F. L., V. DELKE, H. SCHIELE, J. ROTMENSEN, F. GROBMAN y N. M. M. D. FOUTO (2025), «Purchasing category management: Portfolio management, sourcing levers and strategy formulation», *Journal of Purchasing and Supply Management*, 101086, advance online publication.
- GHOBBAR, A. A. y C. H. FRIEND (2002), «Sources of intermittent demand for aircraft spare parts within airline operations», *Journal of Air Transport Management*, **8**(4), págs. 221–231.
- GLOCK, C. H., K. SCHWINDL y M. Y. JABER (2012), «An EOQ model with fuzzy demand and learning in fuzziness», *International Journal of Services and Operations Management*, **12**(1), págs. 90–100.
- GUIDA, M., F. CANIATO y A. MORETTO (2025), «AI meets spend classification: A new frontier in information processing», *Journal of Purchasing and Supply Management*, **31**(3), 100993.
- GUTIERREZ, R. S., A. O. SOLIS y S. MUKHOPADHYAY (2008), «Lumpy demand forecasting using neural networks», *International Journal of Production Economics*, **111**(2), págs. 409–420.
- HANKE, J. E. y D. W. WICHERN (2006), *Pronósticos en los negocios*, 8ª edición, Pearson Educación, México, D.F.
- HYNDMAN, R. J. y G. ATHANASOPOULOS (2021), *Forecasting: Principles and Practice*, tercera edición, OTexts, Melbourne, Australia, URL <https://otexts.com/fpp3/>.

- INTERNATIONAL ORGANIZATION FOR STANDARDIZATION (2019), «ISO 8000-63:2019: Data quality—Part 63: Data quality management: Process measurement», International Standard, URL <https://www.iso.org/standard/65344.html>.
- INTERNATIONAL ORGANIZATION FOR STANDARDIZATION y INTERNATIONAL ELECTROTECHNICAL COMMISSION (2008), «ISO/IEC 25012:2008: Software engineering—Software product quality requirements and evaluation (SQuaRE)—Data quality model», International Standard, URL <https://www.iso.org/standard/35736.html>.
- KARTTUNEN, E., K. LINTUKANGAS y J. HALLIKAS (2023), «Digital transformation of the purchasing and supply management process», *International Journal of Physical Distribution & Logistics Management*, **53**(5/6), págs. 685–706.
- KUMAR, S. y P. T. EVERS (2015), «Setting safety stock based on imprecise records», *International Journal of Production Economics*, **169**, págs. 68–75.
- MARTÍNEZ-PLUMED, F., L. CONTRERAS-OCHANDO, C. FERRI, J. HERNÁNDEZ-ORALLO, M. KULL, N. LACHICHE, M. J. RAMÍREZ-QUINTANA y P. FLACH (2021), «CRISP-DM twenty years later: From data mining processes to data science trajectories», *IEEE Transactions on Knowledge and Data Engineering*, **33**(8), págs. 3048–3061.
- MICROSOFT (2022), «Grupo Bimbo takes a bite out of production costs with Azure IoT throughout factories», Microsoft Customer Story, URL <https://www.microsoft.com/en/customers/story/1496159330015555860-grupo-bimbo-consume-r-goods-azure-iot>.
- MICROSOFT NEWS CENTER EUROPE (2022), «Grupo Bimbo transforma el análisis de datos de su área comercial, gracias a las soluciones de Microsoft», Microsoft News Center Europe, URL <https://news.microsoft.com/es-es/2022/11/02/grupo-bimbo-transforma-el-analisis-de-datos-de-su-area-comercial-gracias-a-las-soluciones-de-microsoft/>.

- NISHAD, A. K., G. AGGARWAL y A. SINGH (2023), «Hesitant fuzzy time series forecasting model of higher order based on one and two-factor aggregate logical relationship», *Engineering Applications of Artificial Intelligence*, **126**, 106897.
- OECD (2022), *Going Digital to Advance Data Governance for Growth and Well-Being*, OECD Publishing, Paris, URL <https://doi.org/10.1787/e3d783b0-en>.
- PEGA (n.d.), «Unilever reduces vendor onboarding from days to hours», Customer case study, URL <https://www.pega.com/customers/unilever-platform>.
- SAMANTA, B. y S. A. AL-ARAIMI (2001), «An inventory control model using fuzzy logic», *International Journal of Production Economics*, **73**(3), págs. 217–226.
- SCHRÖER, C., F. KRUSE y J. MARX GÓMEZ (2021), «A systematic literature review on applying the CRISP-DM process model», *Procedia Computer Science*, **181**, págs. 526–534.
- SHALE, E. A., J. E. BOYLAN y F. R. JOHNSTON (2006), «Forecasting for intermittent demand: The estimation of an unbiased average», *Journal of the Operational Research Society*, **57**(5), págs. 588–592.
- SHAMETI, K., D. HYSENI, Y. MANOLOPOULOS y B. ÇIÇO (2026), «Comparison of CRISP-DM versus OSEMN methodologies using linear regression and statistical analysis», *Expert Systems*, **43**(2), e70185.
- SONG, Q. y B. S. CHISSOM (1993), «Fuzzy time series and its models», *Fuzzy Sets and Systems*, **54**(3), págs. 269–277.
- SPREITZENBARTH, J. M., C. BODE y H. STUCKENSCHMIDT (2024), «Artificial intelligence and machine learning in purchasing and supply management: A mixed-methods review of the state-of-the-art in literature and practice», *Journal of Purchasing and Supply Management*, **30**(1), 100896.
- SWAMIDASS, P. M. (2000), «Seven “rights” of logistics», en P. M. Swamidass (editor), *Encyclopedia of Production and Manufacturing Management*, Springer, Boston, MA, págs. 684–684.

- SYNTETOS, A. A. y J. E. BOYLAN (2001), «On the Bias of Intermittent Demand Estimates», *International Journal of Production Economics*, **71**(1–3), págs. 457–466.
- SYNTETOS, A. A. y J. E. BOYLAN (2005), «The accuracy of intermittent demand estimates», *International Journal of Forecasting*, **21**(2), págs. 303–314.
- TAYLOR, S. J. y B. LETHAM (2018), «Forecasting at scale», *The American Statistician*, **72**(1), págs. 37–45.
- UNILEVER (2025), «How Unilever’s digital transformation is driving operational excellence», Unilever corporate news, URL <https://www.unilever.com/news/news-search/2025/how-unilevers-digital-transformation-is-driving-operational-excellence/>.
- WIRTH, R. y J. HIPPE (2000), «CRISP-DM: Towards a standard process model for data mining», en *Proceedings of the 4th International Conference on the Practical Applications of Knowledge Discovery and Data Mining*, Manchester, United Kingdom, págs. 29–39.

RESUMEN AUTOBIOGRÁFICO

Javier Alfonso López Parra

Candidato para obtener el grado de
Maestría en Logística y Cadena de Suministro

Universidad Autónoma de Nuevo León
Facultad de Ingeniería Mecánica y Eléctrica

Tesis:

DECISIÓN INTELIGENTE EN COMPRAS INDIRECTAS: INTEGRACIÓN
DE CATEGORIZACIÓN Y ANÁLISIS PREDICTIVO

Mi nombre es Javier Alfonso López Parra. Nací el 4 de enero de 1993 en Ansermanuevo, Valle del Cauca, Colombia. Crecí junto a mis padres, Luis Alfonso López Yépez y Dioselina Parra Gaviria, y a mi hermano, Andrés Parra.

Mi padre falleció en 2012. Todavía pienso en él seguido, sobre todo cuando las cosas se complican y siento que necesito su consejo. Era un hombre de trabajo, de esos que no se quejaban aunque el día hubiera sido pesado, y de alguna manera eso se me quedó grabado.

Tiempo después decidí venirme a México. No fue una decisión fácil ni rápida. Lo pensé mucho antes de hacerlo. Llegar aquí significó aprender desde cosas tan simples como el acento y las costumbres hasta cosas más grandes, como estar lejos

de mi familia en fechas importantes. Hubo momentos buenos y otros bastante duros, pero con el tiempo este país se volvió también un poco mío.

Estudié Ingeniería Aeronáutica en la Universidad Politécnica de Apodaca y terminé la carrera en 2022. La aviación siempre me llamó la atención desde niño. No sé exactamente por qué, pero algo de los aviones me atrapó. Cuando tuve la oportunidad de estudiarla no lo dudé.

Empecé a trabajar en el área de carga aérea. Ahí aprendí rápido que en esa industria no hay espacio para los detalles sueltos. Todo tiene que cuadrar y, si algo se retrasa, se producen consecuencias en cadena. Con los años fui tomando más responsabilidades y entendí mejor todo lo que hay detrás de que un avión simplemente despegue.

Después entré a trabajar en Mercado Libre. Empecé haciendo picking en el centro de distribución, caminando por los pasillos y sacando productos de los racks para armar los pedidos. No era un trabajo muy glamoroso, pero ahí aprendí cómo funciona un almacén por dentro, los tiempos que se manejan, los errores que pasan cuando uno va con prisa, y por qué cada detalle cuenta cuando se repite miles de veces al día.

Con el tiempo fui avanzando dentro de la empresa hasta llegar al área de logística de la línea de transporte. El cambio fue grande, porque pasé de estar moviendo cajas físicamente a estar viendo rutas, tiempos de entrega y la coordinación entre distintos puntos. Viéndolo en retrospectiva, el empezar desde abajo y conocer la operación desde piso me sirvió mucho después, cuando ya estaba del otro lado tomando decisiones sobre esas mismas operaciones.

También tuve la oportunidad de dar clases a estudiantes relacionados con la aviación. No esperaba que me gustara tanto enseñar, pero terminé aprendiendo yo también bastante de esa experiencia, sobre todo a explicar las cosas de formas distintas según quién tuviera enfrente.

Hoy, entre Colombia y México, siento que tengo mi corazón de cada lado. Colombia es mi familia, mi infancia y el lugar de donde vengo. México es donde construí buena parte de lo que soy ahora, con sus retos y también con cosas buenas que no esperaba encontrar.