

Análisis cuantitativo y proyección a largo plazo: evaluación de la recolección de residuos domiciliarios en Bogotá y su impacto en la planificación urbana (2018-2029)

Edixon Javier Diaz Romero

Asesor

Rafael Gaitán Ospina

Universidad Nacional Abierta y a Distancia UNAD
Escuela de Ciencias Básicas, Tecnología e Ingeniería ECBTI
Especialización en Ciencia de Datos y Analítica
2025

Nota de Aceptación

Director de Trabajo de Grado

Rafael Gaitán Ospina

Jurado

Jurado

Dedicatoria

Dedico la culminación de este proyecto a las personas esenciales en mi vida y a quienes hicieron de este recorrido una experiencia posible y gratificante.

A mis amados padres Héctor y Bertha, por su apoyo incondicional, su sacrificio y por enseñarme que la perseverancia es la clave para alcanzar cualquier meta. Sin su amor y confianza, nada de esto hubiera sido posible.

A mi querida hermana Leidy, por ser mi primera y más leal amiga de vida y por su dedicación a la preparación y enseñanza, palabra de ánimo y admiración que me impulsó a seguir adelante.

De manera muy especial, a mi tía Miryam Romero, por ser un faro de conocimiento y por inculcarme con tanta convicción el valor incalculable de la educación y el estudio. Sus lecciones resonaron en cada página y me recordaron el propósito de este esfuerzo.

A mis amigos, quienes con sus palabras de motivación y ánimo constante, me impulsaron a ponerle todas las ganas y la dedicación necesaria para ver este proyecto terminado. Su creencia en mí fue un gran incentivo.

Y finalmente, a mi fiel compañera de cuatro patas, mi adorada Mafalda, quien con su presencia silenciosa y cariñosa me ofreció el consuelo y la calma necesarios durante las largas noches, madrugadas y fines de semana dedicados a la elaboración de este trabajo.

Agradecimientos

Mi sincero agradecimiento se extiende a quienes, con su guía y sabiduría, enriquecieron significativamente este proyecto:

En primer lugar, expreso mi profunda gratitud al profesor Rafael Gaitán Ospina, cuya vasta experiencia y profundos conocimientos en modelos de predicción fueron fundamentales. Su orientación especializada fue clave para que la aplicación práctica de este proyecto pudiera materializarse con rigor y éxito.

Extiendo mi reconocimiento a todos los docentes de la Universidad Nacional Abierta y a Distancia (UNAD), por la calidad de los conocimientos compartidos a lo largo de mi formación.

Su dedicación no solo me proporcionó las bases teóricas necesarias, sino que también me despertó la vocación y la invitación constante a seguir aprendiendo y desarrollándome en el gigantesco y apasionante campo de la Ciencia de Datos.

Resumen

Esta investigación presenta un análisis exhaustivo de la generación de residuos domiciliarios en Bogotá, desglosado por localidad, desde 2018 hasta 2024. Además, se incluyen predicciones para los próximos cinco años, basadas en la recopilación, procesamiento y análisis de datos históricos.

El estudio se fundamenta en la metodología CRISP-DM (Proceso Estándar Inter-Industria para la Minería de Datos), un enfoque ampliamente reconocido y riguroso en proyectos de análisis de datos. Esta metodología permitió una exploración y análisis detallado de los datos históricos de recolección de residuos domiciliarios, utilizando la herramienta Power BI para la visualización y el entendimiento profundo de los patrones.

Para la proyección de tendencias futuras con alta precisión, se emplearon los modelos predictivos de series de tiempo SARIMAX (Seasonal Autoregressive Integrated Moving Average with Exogenous Regressors) y Prophet, los resultados de esta investigación son cruciales, ya que identifican las localidades con mayor tendencia a la generación de residuos. Esta información detallada, obtenida mediante el análisis en Power BI y las predicciones de SARIMAX y Prophet, proporciona un insumo de valor incalculable para el Gobierno de la Ciudad de Bogotá y otras entidades interesadas. Esto posibilitará enfocar esfuerzos y recursos en áreas específicas de la ciudad, permitiendo una intervención urgente y efectiva orientada a la reducción de la generación de residuos domiciliarios.

Palabras claves: Economía circular, Gestión de residuos, Minería de datos, Recolección de residuos domiciliarios, Relleno sanitario.

Abstract

This research presents an exhaustive analysis of household waste generation in Bogotá, broken down by localidad (district), from 2018 to 2024. Additionally, it includes predictions for the next five years, based on the collection, processing, and analysis of historical data.

The study is based on the CRISP-DM (Cross-Industry Standard Process for Data Mining) methodology, a widely recognized and rigorous approach in data analysis projects. This methodology allowed for a detailed exploration and analysis of historical household waste collection data, utilizing the Power BI tool for visualization and a deep understanding of the patterns.

For projecting future trends with high precision, the time series predictive models SARIMAX (Seasonal Autoregressive Integrated Moving Average with Exogenous Regressors) and Prophet were employed. The results of this research are crucial, as they identify the localidades with the highest tendency for waste generation. This detailed information, obtained through analysis in Power BI and the predictions from SARIMAX and Prophet, provides an invaluable input for the Bogotá City Government and other interested entities. This will make it possible to focus efforts and resources on specific areas of the city, allowing for an urgent and effective intervention aimed at reducing household waste generation.

Keywords: Circular economy, Data mining, Household waste collection, Sanitary landfill, Waste management.

Tabla de Contenido

Introducción	16
Descripción del Problema	18
Planteamiento del Problema.....	18
Delimitación del Alcance.....	19
Sistematización del Problema	19
Identificación del Problema Central.....	19
Análisis de la Generación de Residuos	20
Causas Principales del Problema.....	20
Consecuencias del Problema	21
Actores Involucrados.....	21
Estrategias Actuales de Intervención	22
Identificación de Vacíos y Áreas de Mejora	22
Justificación	23
Objetivos.....	24
Objetivo General	24
Objetivos Específicos.....	24
Marco de Referencia	25
Estado del Arte.....	25
Marco Contextual.....	27
Marco Teórico.....	31
Marco Conceptual	32
Marco Normativo	34

Metodología	37
Método	37
Comprensión del Problema	37
Comprensión de los Datos.....	37
Preparación de los Datos	38
Análisis Exploratorio de Datos (EDA).....	38
Modelado.....	39
Evaluación de los Resultados	39
Visualización y Comunicación.....	39
Despliegue y Recomendaciones	40
Resultados	41
Comprensión del Problema	41
Comprensión de los Datos	42
Preparación de los Datos.....	42
Análisis Exploratorio de Datos (EDA)	43
Introducción y Resumen de Datos.....	44
Análisis de Tendencias Temporales	45
Análisis Geográfico y Segmentación por Localidad.....	48
Modelado	57
Bases de Datos para el Modelado.....	57
Selección de Técnicas de Modelado	57
Análisis Descriptivo y Descomposición de la Serie de Tiempo	58
Resultado del Modelado	74

Conclusiones	112
Recomendaciones	114
Referencias Bibliográficas	116

Lista de Tablas

Tabla 1 <i>Evaluación y Comparación del Rendimiento del Modelo MAE Y RMSE Base de Datos Bogotá</i>	84
Tabla 2 <i>Evaluación y Comparación del Rendimiento del Modelo R² Y MAPE Base de Datos Bogotá</i>	85
Tabla 3 <i>Predicción Anual de Residuos Domiciliarios Base de Datos Bogotá</i>	88
Tabla 4 <i>Evaluación y Comparación del Rendimiento del Modelo MAE Y RMSE Base de Datos Suba</i>	96
Tabla 5 <i>Evaluación y Comparación del Rendimiento del Modelo R² Y MAPE Base de Datos Suba</i>	97
Tabla 6 <i>Predicción Anual de Residuos Domiciliarios Base de Datos Suba</i>	100
Tabla 7 <i>Evaluación y Comparación del Rendimiento del Modelo MAE Y RMSE Base de Datos Kennedy</i>	108
Tabla 8 <i>Evaluación y Comparación del Rendimiento del Modelo R² Y MAPE Base de Datos Kennedy</i>	108
Tabla 9 <i>Predicción Anual de Residuos Domiciliarios Base de Datos Kennedy</i>	111

Lista de Figuras

Figura 1 <i>Base de Datos para Análisis Exploratorio (EDA)</i>	43
Figura 2 <i>Total Toneladas Recolectadas por Año (12 de febrero del 2018 al 31 de octubre de 2024)</i>	44
Figura 3 <i>Grafica Toneladas Recolectadas por Mes (12 de febrero del 2018 al 31 de octubre de 2024)</i>	45
Figura 4 <i>Grafica Toneladas Recolectadas por Día de la Semana Bogotá (12 de febrero del 2018 al 31 de octubre de 2024)</i>	46
Figura 5 <i>Residuos Domiciliarios Recolectados por Día de la semana, Mes y Año (12 de febrero del 2018 al 31 de octubre de 2024)</i>	47
Figura 6 <i>Residuos Domiciliarios Recolectados por Localidad (12 de febrero del 2018 al 31 de octubre de 2024)</i>	49
Figura 7 <i>Localidades con Mayor Cantidad de Residuos Domiciliarios Recolectados (12 de febrero del 2018 al 31 de octubre de 2024)</i>	50
Figura 8 <i>Toneladas Recogidas por mes en Localidades Especificas (12 de febrero del 2018 al 31 de diciembre de 2021)</i>	50
Figura 9 <i>Toneladas Recogidas por Mes en Localidades Especificas (1 de enero del 2021 al 31 de octubre de 2024)</i>	51
Figura 10 <i>Toneladas Recogidas por Día de la Semana en la Localidad de Suba (12 de febrero del 2018 al 31 de octubre de 2024)</i>	52
Figura 11 <i>Toneladas Recogidas por Día de la Semana en la Localidad de Kennedy (12 de febrero del 2018 al 31 de octubre de 2024)</i>	53

Figura 12 <i>Toneladas Recogidas por Día de la Semana en la Localidad de Engativá (12 de febrero del 2018 al 31 de octubre de 2024)</i>	54
Figura 13 <i>Toneladas Recogidas por Día de la Semana en la Localidad de Bosa (12 de febrero del 2018 al 31 de octubre de 2024)</i>	55
Figura 14 <i>Toneladas Recogidas por Día de la Semana en la Localidad de Usaquén (12 de febrero del 2018 al 31 de octubre de 2024)</i>	56
Figura 15 <i>Descomposición de la Serie de Tiempo Base de Datos Bogotá 2018 (Toneladas Diarias de Residuos Domiciliarios)</i>	59
Figura 16 <i>Descomposición de la Serie de Tiempo Base de Datos Bogotá 2019 (Toneladas Diarias de Residuos Domiciliarios)</i>	59
Figura 17 <i>Descomposición de la Serie de Tiempo Base de Datos Bogotá 2020 (Toneladas Diarias de Residuos Domiciliarios)</i>	60
Figura 18 <i>Descomposición de la Serie de Tiempo Base de Datos Bogotá 2021 (Toneladas Diarias de Residuos Domiciliarios)</i>	60
Figura 19 <i>Descomposición de la Serie de Tiempo Base de Datos Bogotá 2022 (Toneladas Diarias de Residuos Domiciliarios)</i>	61
Figura 20 <i>Descomposición de la Serie de Tiempo Base de Datos Bogotá 2023 (Toneladas Diarias de Residuos Domiciliarios)</i>	61
Figura 21 <i>Descomposición de la Serie de Tiempo Base de Datos Bogotá 2024 (Toneladas Diarias de Residuos Domiciliarios)</i>	62
Figura 22 <i>Descomposición de la Serie de Tiempo Base de Datos Suba 2018 (Toneladas Diarias de Residuos Domiciliarios)</i>	64

Figura 23 <i>Descomposición de la Serie de Tiempo Base de Datos Suba 2019 (Toneladas Diarias de Residuos Domiciliarios)</i>	64
Figura 24 <i>Descomposición de la Serie de Tiempo Base de Datos Suba 2020 (Toneladas Diarias de Residuos Domiciliarios)</i>	65
Figura 25 <i>Descomposición de la Serie de Tiempo Base de Datos Suba 2021 (Toneladas Diarias de Residuos Domiciliarios)</i>	65
Figura 26 <i>Descomposición de la Serie de Tiempo Base de Datos Suba 2022 (Toneladas Diarias de Residuos Domiciliarios)</i>	66
Figura 27 <i>Descomposición de la Serie de Tiempo Base de Datos Suba 2023 (Toneladas Diarias de Residuos Domiciliarios)</i>	66
Figura 28 <i>Descomposición de la Serie de Tiempo Base de Datos Suba 2024 (Toneladas Diarias de Residuos Domiciliarios)</i>	67
Figura 29 <i>Descomposición de la Serie de Tiempo Base de Datos Kennedy 2018 (Toneladas Diarias de Residuos Domiciliarios)</i>	69
Figura 30 <i>Descomposición de la Serie de Tiempo Base de Datos Kennedy 2019 (Toneladas Diarias de Residuos Domiciliarios)</i>	70
Figura 31 <i>Descomposición de la serie de Tiempo Base de Datos Kennedy 2020 (Toneladas Diarias de Residuos Domiciliarios)</i>	70
Figura 32 <i>Descomposición de la Serie de Tiempo Base de Datos Kennedy 2021 (Toneladas Diarias de Residuos Domiciliarios)</i>	71
Figura 33 <i>Descomposición de la Serie de Tiempo Base de Datos Kennedy 2022 (Toneladas Diarias de Residuos Domiciliarios)</i>	71

Figura 34 <i>Descomposición de la Serie de Tiempo Base de Datos Kennedy 2023 (Toneladas Diarias de Residuos Domiciliarios)</i>	72
Figura 35 <i>Descomposición de la Serie de Tiempo Base de Datos Kennedy 2024 (Toneladas Diarias de Residuos Domiciliarios)</i>	72
Figura 36 <i>Resultados Estadísticos del Modelo SARIMAX Base de Datos Bogotá (Toneladas Diarias de Residuos Domiciliarios)</i>	77
Figura 37 <i>Predicción del Modelo SARIMAX sobre el Periodo de Prueba Base de Datos Bogotá (Toneladas Diarias de Residuos Domiciliarios)</i>	79
Figura 38 <i>Predicción del Modelo Prophet sobre el Periodo de Prueba Base de Datos Bogotá (Toneladas Diarias de Residuos Domiciliarios)</i>	82
Figura 39 <i>Predicciones Base de Datos Bogotá 01/11/2024 al 31/12/2029 (Toneladas Diarias de Residuos Domiciliarios)</i>	86
Figura 40 <i>Resultados Estadísticos del Modelo SARIMAX Base de Datos Suba (Toneladas Diarias de Residuos Domiciliarios)</i>	89
Figura 41 <i>Predicción del Modelo SARIMAX sobre el Periodo de Prueba Base de Datos Suba (Toneladas Diarias de Residuos Domiciliarios)</i>	91
Figura 42 <i>Predicción del Modelo Prophet sobre el Periodo de Prueba Base de Datos Suba (Toneladas Diarias de Residuos Domiciliarios)</i>	94
Figura 43 <i>Predicciones Base de Datos Suba 01/11/2024 al 31/12/2029 (Toneladas Diarias de Residuos Domiciliarios)</i>	98
Figura 44 <i>Resultados Estadísticos del Modelo SARIMAX Base de Datos Kennedy (Toneladas Diarias de Residuos Domiciliarios)</i>	101

Figura 45 <i>Predicción del Modelo SARIMAX sobre el Periodo de Prueba Base de Datos Kennedy (Toneladas Diarias de Residuos Domiciliarios)</i>	103
Figura 46 <i>Predicción del Modelo Prophet sobre el Periodo de Prueba Base de Datos Kennedy (Toneladas Diarias de Residuos Domiciliarios)</i>	106
Figura 47 <i>Predicciones Base de Datos Kennedy 01/11/2024 al 31/12/2029 (Toneladas Diarias de Residuos Domiciliarios)</i>	109

Introducción

Bogotá, como epicentro de Colombia, ha experimentado un crecimiento demográfico y urbano sin precedentes en las últimas décadas. Este desarrollo, evidenciado por el aumento del 94% en su población entre 1985 y 2018 equivalente a 3.607.930, ha impulsado la expansión de viviendas, infraestructura y la demanda de servicios básicos como agua, gas y energía. Sin embargo, este crecimiento exponencial ha generado un desafío crítico: la gestión de residuos domiciliarios y la presión que esto ejerce sobre los servicios de recolección.

La transición desde vertederos a cielo abierto en barrios como El Cortijo y Patio Bonito hasta la implementación del Relleno Sanitario Parque de Innovación Doña Juana en 1988 marcó un hito en la gestión de residuos de la ciudad (Camargo, 1987). No obstante, el aumento constante en la generación de residuos, impulsado por el crecimiento poblacional y la actividad económica, ha ejercido una presión creciente sobre la capacidad del relleno sanitario Parque de innovación Doña Juana, que actualmente se encuentra bajo constante evaluación.

Una gestión de residuos ineficaz puede desencadenar graves problemas de salud pública y ambiental, incluyendo la propagación de enfermedades y la contaminación del suelo y del agua, deteriorando la calidad de vida urbana.

La composición de los residuos, con una proporción significativa de materia orgánica, presenta tanto oportunidades para la valorización y el compostaje como desafíos logísticos. Para abordar esta situación, se requiere un enfoque integral que priorice la reducción en la fuente, la separación y el reciclaje, el aprovechamiento de residuos orgánicos y la disposición adecuada de los residuos no aprovechables.

La implementación de tecnologías innovadoras, la promoción de la economía circular y la participación de la ciudadanía son fundamentales para lograr una gestión sostenible de

residuos en Bogotá. Este desafío requiere una planificación cuidadosa y un compromiso firme con la sostenibilidad ambiental.

Para abordar los desafíos actuales y futuros en la gestión de residuos domiciliarios en Bogotá, se requiere un esfuerzo concertado y la aplicación precisa de técnicas de ciencia de datos, específicamente minería de datos. Es fundamental examinar el conjunto de datos que posee la Unidad Administrativa Especial de Servicios Públicos (UAESP) con información relacionada con los residuos domiciliarios recogidos y depositados en el relleno sanitario Doña Juana, con el fin de guiar decisiones óptimas. Este análisis detallado de la recolección de residuos no solo permitirá comprender las tendencias y desafíos actuales, sino que también justificará la importancia de proyectar la generación de residuos y las necesidades de recolección para los próximos años. De esta manera, se podrán planificar y optimizar los servicios, asegurando una gestión de residuos más eficiente y sostenible.

Este Proyecto aplicado tiene como objetivo analizar los datos históricos de la recolección de residuos domiciliarios en Bogotá (2018-2024), con el fin de proyectar tendencias y comportamientos esperados para el período 2025-2029 que contribuya a la solución de los desafíos actuales y futuros en la gestión de residuos domiciliarios en la ciudad de Bogotá, mediante la aplicación precisa de técnicas de ciencia de datos.

Descripción del Problema

La gestión de residuos domiciliarios en Bogotá ha sido históricamente un tema central en las campañas electorales, reflejando una problemática que se agudiza con el tiempo y que desafía la eficacia de las políticas y estrategias gubernamentales. A pesar de los esfuerzos de cada administración por reducir la producción de residuos, los resultados han sido insuficientes. A finales de 2023, Bogotá generaba aproximadamente 6.300 toneladas de basura diarias, con un aprovechamiento de solo el 14% al 15%, según datos del Ministerio de Ambiente y Desarrollo Sostenible, un indicador preocupante de la limitada eficacia en la valorización de residuos (Armando Gutiérrez González & Concejo de Bogotá, 2020).

Adicional, el relleno sanitario Parque de Innovación Doña Juana enfrenta un horizonte preocupante, con una vida útil limitada, condicionada a la implementación efectiva de estrategias de reciclaje y separación en la fuente. La ausencia de procesos de separación desde 1988 ha llevado a la acumulación de vastas cantidades de residuos, incluyendo materiales reciclables, lo que acelera el agotamiento del relleno sanitario y agrava la contaminación del entorno.

Ahora bien, Para que la Alcaldía Mayor de Bogotá pueda implementar estrategias efectivas y focalizadas en la resolución del problema de la gestión de residuos, es imperativo realizar un análisis exhaustivo del comportamiento y la evolución de los residuos domiciliarios. Este análisis debe abarcar datos históricos, la situación actual y proyecciones futuras, permitiendo identificar y rastrear a los usuarios con mayor producción de residuos y evaluar su comportamiento a lo largo del tiempo.

Planteamiento del Problema

Considerando la situación que supone retos y actividades urgentes para las entidades gubernamentales y privadas en el manejo de residuos domiciliarios que produce la ciudad de

Bogotá, este proyecto aplicado tiene como objetivo principal analizar y predecir la generación de residuos domiciliarios en Bogotá, desglosando el estudio por cada localidad. Para ello, se recopilaron y procesaron datos de recolección de residuos desde 2018 hasta 2024. Mediante técnicas de análisis exploratorio de datos (EDA), se identificaron patrones y tendencias relevantes en cada localidad. Posteriormente, se desarrolló un modelo predictivo que proyecte la cantidad de residuos generados entre 2025 y 2029. Este análisis detallado permitirá al gobierno de la ciudad enfocar sus esfuerzos de reciclaje y separación en la fuente en aquellas localidades con mayor generación de residuos.

Delimitación del Alcance

El área de estudio se centra en Bogotá, dividida en sus 20 localidades para analizar detalladamente el comportamiento sectorizado de la generación y recolección de residuos domiciliarios. Se utilizaron datos recopilados durante un período de seis años, de 2018 a 2024, para predecir las tendencias de los próximos cinco años.

Sistematización del Problema

La sistematización del problema de la gestión de residuos domiciliarios en Bogotá implica un proceso estructurado para comprender y organizar las diversas dimensiones que conforman este fenómeno. A continuación, se describen los elementos clave para su sistematización, abordando tanto las causas como las consecuencias, así como las interacciones entre actores y territorios involucrados.

Identificación del Problema Central

La ciudad de Bogotá, que cuenta con cerca de 7.9 millones de habitantes, enfrenta el desafío de gestionar aproximadamente 6,300 toneladas diarias de residuos que son depositados en el Relleno Sanitario Doña Juana, cuya vida útil se estima entre 5 y 10 años. Aunque la

recolección y el transporte están a cargo de cinco operadores privados, su rol no incluye el aprovechamiento de los residuos, lo que deja la recuperación de materiales reciclables en manos de los recicladores de oficio; dada la deficiente separación en la fuente, una cantidad considerable de residuos con potencial de reciclaje termina innecesariamente en el relleno sanitario.

Análisis de la Generación de Residuos

Este proyecto tomo bases de datos otorgados por la Unidad Administrativa Especial de Servicios Públicos (UAESP). Relacionados únicamente con aquellos tipificados entre las empresas de aseo y la UAESP como residuos domiciliarios, por lo que no se tendrá en cuenta por ejemplo aquellos residuos de construcción y demolición RCD propios del producto de construcción y remodelación de infraestructura civil como escombros y arrojados clandestinos, residuos de corte de césped y poda de árboles, barrido de calles y grandes generadores.

Causas Principales del Problema

Bogotá enfrenta un desafío de gestión de residuos que está constantemente sobrepasado por el aumento exponencial de su población y consumo. Esta problemática compleja se origina en el crecimiento demográfico y la urbanización que superan la capacidad de la infraestructura actual. El problema se profundiza debido a la insuficiente separación en la fuente por parte de la ciudadanía y la limitada capacidad tecnológica de los sistemas de recolección y tratamiento, lo que resulta en un volumen excesivo de materiales reciclables enviados al relleno sanitario. A esto se suman factores como los comportamientos inadecuados de los ciudadanos, las deficiencias en la normativa y su aplicación, las desigualdades socioeconómicas y la falta de educación ambiental. En última instancia, la ausencia de un análisis de información adecuado por parte de

la Administración dificulta la toma de decisiones estratégicas para fomentar la economía circular y resolver la acumulación crítica de basura en espacios públicos.

Consecuencias del Problema

La deficiente gestión de residuos domiciliarios en Bogotá se traduce principalmente en la ineficacia en la toma de decisiones debido a la falta de datos precisos, lo que impide diseñar políticas efectivas y resulta en la pérdida de materiales reciclables y la asignación ineficiente de recursos. Esto, a su vez, genera dificultades para la planificación y optimización de servicios, acelerando la saturación del Relleno Sanitario Parque de Innovación Doña Juana y aumentando los costos de gestión. Además, la falta de información detallada obstaculiza la economía circular y la transición de residuos a recursos, lo que conlleva una pérdida de oportunidades económicas y de creación de empleo. Finalmente, el manejo inadecuado provoca un aumento de los impactos ambientales, incluyendo la contaminación del suelo, agua y aire, que culmina en una grave afectación de la salud pública por la proliferación de enfermedades.

Actores Involucrados

La gestión de residuos en Bogotá es un proceso complejo que involucra a múltiples actores de diversos sectores. La Alcaldía Mayor y sus secretarías (Ambiente, Hábitat) son responsables de la formulación de políticas, mientras que la UAESP se encarga específicamente de la coordinación y supervisión de los servicios de aseo. En la operación, participan las Empresas de Servicios Públicos (recolección, transporte y tratamiento) y las Organizaciones de Recicladores, que son fundamentales para la recuperación y el aprovechamiento de materiales. Finalmente, la Ciudadanía es un actor clave mediante la separación en la fuente y la disposición adecuada, complementada por el Sector Privado, que invierte en tecnologías y promueve la economía circular, y las ONGs, que cumplen un rol esencial de veeduría y concientización.

Estrategias Actuales de Intervención

Para abordar la gestión de residuos, Bogotá cuenta con la Unidad Administrativa Especial de Servicios Públicos (UAESP), entidad encargada de coordinar y supervisar los servicios de aseo, incluyendo la recolección, el transporte y la disposición final de residuos. A través de la contratación de estos servicios, se han implementado metodologías que integran a los recicladores de oficio antes de que las empresas de aseo realicen la recolección y el traslado de los residuos al relleno sanitario. Además, mediante la colaboración entre las empresas prestadoras del servicio de aseo y la interventoría, se ha recopilado información valiosa sobre la cantidad de toneladas de residuos generadas diariamente y su origen, lo que permitirá analizar y comprender el comportamiento de la producción de residuos en la ciudad.

Identificación de Vacíos y Áreas de Mejora

La gestión eficaz de los residuos domiciliarios en Bogotá exige una precisión crítica en la proyección de su aumento, el crecimiento demográfico y el consumo, ya que la falta de esta exactitud obstaculiza las estrategias adecuadas. Actualmente, persisten vacíos como la insuficiente separación en la fuente, la limitada capacidad de infraestructura y los comportamientos inadecuados de la ciudadanía, que acortan la vida útil del relleno sanitario Doña Juana. Por ello, es imperativo abordar estos desafíos mediante una transición hacia la economía circular, el fortalecimiento de la educación ambiental, la formalización de los recicladores y, fundamentalmente, la aplicación formal de técnicas de ciencia de datos a lo largo de todo el ciclo de vida de los residuos para obtener información valiosa y asegurar una toma de decisiones oportuna de cara al futuro.

Justificación

La gestión de residuos sólidos siempre ha sido un desafío ambiental y social para los gobiernos de las grandes ciudades. Bogotá es una de ellas, la cual tiene el Parque de innovación de Doña Juana (relleno sanitario) que recibe diariamente más de 6,300 toneladas de residuos generados por la ciudad, esta cantidad de residuos lo coloca al borde de su capacidad operativa, se estima una durabilidad de entre 5 a 10 años dependiendo si esta cantidad de residuos diarios se reducen considerablemente o si se prorroga su licencia ambiental (Procuraduría General de la Nación, 2023).

Siendo esta una gran preocupación para el gobierno de la Ciudad, pues no puede prescindir de un lugar a donde llevar los residuos que genera constantemente su población, para ello el gobierno debe atender con urgencia a través de políticas de economía circular, iniciativas de reciclaje y gestión de residuos la problemática presente.

La relevancia de este proyecto aplicado radica en la urgencia de comprender a fondo el comportamiento de los residuos domiciliarios generados en Bogotá entre 2018 y 2024, y en la necesidad de predecir su evolución durante los próximos 5 años. Mediante el análisis con herramientas de la ciencia de datos, se busca proporcionar información crucial para la toma de decisiones oportunas, previniendo así las complicaciones ambientales, de salud pública y económicas que se avecinan.

Objetivos

Objetivo General

Analizar los datos históricos de la recolección de residuos domiciliarios en Bogotá (2018-2024) con el fin de proyectar tendencias y comportamientos esperados para el período 2025-2029.

Objetivos Específicos

Analizar el comportamiento histórico de los residuos domiciliarios en Bogotá (2018-2024) mediante el procesamiento y exploración de datos.

Desarrollar un modelo predictivo basado en datos históricos para proyectar las tendencias de generación de residuos en Bogotá (2025-2029).

Proponer recomendaciones estratégicas para la gestión de residuos basadas en los hallazgos del análisis histórico y las proyecciones futuras.

Marco de Referencia

Estado del Arte

El continuo crecimiento de la población y la generación de residuos con el pasar de los años ha tomado gran importancia debido a la preocupación que genera la carencia de métodos y alternativas que solucionen el impacto a la salud y ambiental, por lo tanto, la cantidad de investigaciones relacionadas con el manejo de residuos domiciliarios son muy extensas, para este caso específico nos enfocaremos a las investigaciones y trabajos relacionados para la Ciudad de Bogotá.

Inicialmente es necesario conocer que La Unidad Administrativa Especial de Servicios Públicos (UAESP) es una entidad del gobierno de Bogotá (Capital de Colombia), en la que su principal objetivo es garantizar la prestación, coordinación, supervisión y control de los servicios públicos en la ciudad, entre uno de estos servicios públicos se haya el Aseo, que incluye la recolección, transporte, disposición final, reciclaje y aprovechamiento de residuos sólidos, así como el barrido y limpieza de vías y áreas públicas.

La Unidad Administrativa Especial de Servicios Públicos (UAESP) fue creada el 30 de noviembre de 1994, mediante el Decreto 782 de 1994 considerando que esta surgió como una transformación de la Unidad Ejecutiva de Servicios Públicos (UESP), que a su vez se creó después de la liquidación de la Empresa Distrital de Aseo (EDIS), la cual había sido creada mucho antes, en 1958. (ALCALDIA MAYOR DE SANTAFE DE BOGOTA, 1994)

Los residuos en Bogotá inicialmente se depositaban en botaderos como El Cortijo y Gibraltar, sin embargo, estos generaron emergencias sanitarias en la década de los 80, que luego fueron cerrados en 1985 y 1988, respectivamente. (Calderón, L 2022)

Esta situación generó en 1988 la apertura del Relleno Sanitario Doña Juana (RSDJ) y fue la puerta de entrada para que los privados gestionaran el servicio de aseo en Bogotá, dado que el Decreto 888 de 1988 declaró la emergencia y autorizó que el 40% del servicio de aseo se contratara con privados y con la Ley 143 de 1994 llegó hasta el 100%.

Desde entonces el servicio de aseo en Bogotá se ha venido prestando con concesiones a empresas privadas y públicas, actualmente mediante la licitación pública de aseo para la prestación de los servicios en Bogotá que se llevó a cabo en 2017, se generaron contratos que comenzaron a ejecutarse en 2018. Esta licitación dividió la ciudad en cinco Áreas de Servicio Exclusivo (ASE) y adjudicó la prestación del servicio a diferentes empresas por un periodo de 8 años con una fecha de terminación prevista para febrero de 2026 (bogota.gov.co, n.d.).

Adicionalmente la empresa que actualmente presta el servicio de administración, operación y mantenimiento integral del Relleno Sanitario Parque de innovación Doña Juana es la Compañía de Gerenciamiento de Residuos (CGR) DOÑA JUANA S.A. ESP. empresa que ha estado a cargo de la gestión integral del relleno desde el 16 de diciembre de 2010, la cual también se encuentra bajo la supervisión y vigilancia de la Unidad Administrativa Especial de Servicios Públicos (UAESP), que es la propietaria de los predios del relleno. (www.uaesp.gov.co, n.d.)

Estas empresas mediante sus métodos de operación generan diferentes datos, entre esos la cantidad de toneladas recolectadas en la ciudad día a día, mucha de esta información se encuentra disponible en páginas como Datos Abiertos Colombia (Superintendencia de Servicios Públicos Domiciliarios, 2024) y Datos Abiertos Bogotá (Datos Abiertos Bogotá, 2024).

Marco Contextual

La gestión de residuos es un desafío global con graves implicaciones para el cambio climático, la contaminación y la pérdida de biodiversidad. La Agenda 2030 de la ONU aborda este problema a través de los Objetivos de Desarrollo Sostenible (ODS), especialmente el ODS 6 (Agua Limpia y Saneamiento), que busca reducir la contaminación de fuentes de agua dulce por desechos mal manejados y mejorar el tratamiento de aguas residuales; el ODS 11 (Ciudades Sostenibles), que exige que las ciudades mejoren la recolección, el manejo y la promoción de la economía circular para reducir el impacto ambiental per cápita; y el ODS 12 (Producción y Consumo Responsables), cuya meta 12.5 busca reducir significativamente los desechos mediante la prevención, reutilización y reciclaje, promoviendo el ecodiseño y la separación en la fuente. El éxito de todas estas metas requiere un trabajo coordinado entre gobiernos, sector privado y ciudadanos, y depende fundamentalmente de un análisis de datos exhaustivo que permita tomar decisiones informadas sobre la generación de residuos.

Colombia a través de la Ley 142 de 1994 regula los servicios públicos domiciliarios en su territorio nacional, en los que se incluye el servicio público domiciliario de aseo, el cual lo define como la recolección, transporte, tratamiento, aprovechamiento y disposición final de residuos sólidos, por medio de esta fomenta la participación de empresas públicas, privadas y mixtas para garantizar la eficiencia en la prestación de los servicios, incluye un marco tarifario en el que estas se definen bajo criterios de eficiencia económica, suficiencia financiera, y un principio de solidaridad que permite subsidios para los usuarios de menores ingresos, crea organismos clave como la Superintendencia de Servicios Públicos Domiciliarios (SSPD), encargada de la inspección, vigilancia y control, y las Comisiones de Regulación, como la CRA (Comisión de Regulación de Agua Potable y Saneamiento Básico), que definen las normas técnicas y tarifarias.

Bogotá ha mantenido un crecimiento poblacional sostenido a través de los censos nacionales (1985 a 2018), aunque la tasa ha mostrado una desaceleración progresiva, con proyecciones que anticipan un leve repunte para el periodo 2018-2031. El censo de 2018 registró una población de 7.181.469 personas, la gran mayoría residiendo en hogares particulares. No obstante, el DANE estimó una omisión censal de 231.097 personas, lo que elevó el total poblacional real de Bogotá en 2018 a 7.412.566 personas.

La capital colombiana es reconocida como la principal potencia económica de Colombia y un motor de crecimiento crucial para el país, contribuyendo con aproximadamente una cuarta parte del Producto Interno Bruto (PIB) nacional y cerca del 20% del empleo. La economía de Bogotá es altamente diversificada y orientada al sector terciario (servicios). Los sectores de servicios financieros e inmobiliarios son los mayores aportantes al PIB, consolidándola como el centro financiero de Colombia, mientras que el comercio y los servicios sociales (salud, educación y administración pública) son pilares fundamentales, siendo el comercio el mayor generador de empleo.

La gestión de residuos en Bogotá ha evolucionado a través de las administraciones recientes: el gobierno de Enrique Peñalosa (2016-2019) se centró en un cambio del modelo de recolección de público a privado, promoviendo el reciclaje y mejorando la infraestructura; la alcaldía de Claudia López (2020-2023) adoptó el Plan de Desarrollo "Un nuevo contrato social y ambiental", enfocándose en la economía circular mediante la actualización del PGIRS, la intensificación de la limpieza, el impulso a la separación en la fuente con el uso de bolsas de colores y el fortalecimiento de la inclusión y formalización de los recicladores de oficio; y, más recientemente, el gobierno de Carlos Fernando Galán ha implementado un plan de choque con tres pilares clave: operativo (mejorar recolección en puntos críticos), pedagógico (promover la

educación y separación en la fuente) y sancionatorio (establecer multas más rigurosas), promoviendo además la participación ciudadana.

La Unidad Administrativa Especial de Servicios Públicos (UAESP) es la entidad distrital de Bogotá encargada de garantizar, coordinar, supervisar y controlar la prestación de servicios esenciales. Sus funciones principales incluyen la gestión integral de residuos sólidos (recolección, transporte, disposición final y reciclaje), el barrido y limpieza de áreas públicas, la supervisión del alumbrado público y la gestión de servicios funerarios. Para el servicio de aseo, La UAESP ha implementado el modelo de Áreas de Servicio Exclusivo (ASE), que divide la ciudad en zonas geográficas específicas asignadas a un único operador privado mediante licitación pública. Este esquema busca asegurar una cobertura total y una calidad uniforme del servicio en todas las áreas, incluyendo las menos rentables.

En el 2018, la UAESP adjudicó la licitación bajo el modelo ASE, resultando ganadoras cinco empresas privadas que se repartieron las zonas de la ciudad. Las compañías adjudicatarias fueron:

Promoambiental Distrito: Se encargó de las zonas de Usaquén, Chapinero, Santa Fe, La Candelaria, San Cristóbal, Usme y Sumapaz.

Limpieza Metropolitana (LIME): Su área de operación fue Antonio Nariño, Ciudad Bolívar, Bosa, Los Mártires, Puente Aranda, Rafael Uribe, Teusaquillo y Tunjuelito.

Ciudad Limpia: Adjudicada para las localidades de Fontibón y Kennedy.

Bogotá Limpia: Cubrió las zonas de Engativá y Barrios Unidos.

Área Limpia: Obtuvo la concesión para la localidad de Suba.

Esta licitación marcó un cambio significativo en la gestión del servicio de aseo en Bogotá, ya que la empresa pública Aguas de Bogotá dejó de prestar el servicio, el cual fue asumido por las nuevas empresas privadas.

La gestión de residuos sólidos en Bogotá opera mediante la recolección de desechos domiciliarios y comerciales con vehículos compactadores, siguiendo horarios y frecuencias definidas, donde la ciudadanía tiene la responsabilidad clave de separar en la fuente utilizando el código de colores (blanca, verde y negra) para facilitar el aprovechamiento.

Complementariamente, los operadores de aseo se encargan del barrido y limpieza de vías y áreas públicas, utilizando métodos manuales y mecánicos, lo que incluye la limpieza de sumideros y la atención a puntos críticos de acumulación. Finalmente, todos los residuos no aprovechables recolectados tienen como destino final el Relleno Sanitario Parque de innovación Doña Juana, considerado un componente central y crítico en todo el sistema de gestión de la ciudad.

Un aspecto distintivo y fundamental del modelo de Aseo de Bogotá es la inclusión de los recicladores de oficio. Tras un largo proceso legal y social, la Corte Constitucional y la regulación nacional han ordenado su formalización y remuneración por su labor.

Los recicladores son reconocidos como prestadores del servicio público de aseo en la actividad de aprovechamiento. Se encargan de la recolección selectiva de residuos sólidos aprovechables (plástico, papel, cartón, vidrio, etc.).

Para compensar su labor, la UAESP ha implementado un esquema de pago que les permite recibir una tarifa por cada kilogramo de material aprovechable que recuperan y comercializan, lo que les brinda un ingreso formal y dignifica su trabajo. Esta labor no solo reduce la cantidad de residuos que llegan al relleno sanitario, sino que también contribuye significativamente a la economía circular de la ciudad.

Por lo tanto, el análisis histórico (2018-2024), es un período crucial porque abarca el comportamiento de la generación de residuos durante dos administraciones anteriores que implementaron estrategias para reducir la cantidad de residuos llevada al Relleno Sanitario Doña Juana, del cual se obtiene un análisis descriptivo que permita observar cómo la ciudad de Bogotá se ha comportado durante los últimos 5 años.

Con esta base de base de datos, la proyección para los siguientes 5 años (2025-2029) se basará en las tendencias identificadas en el análisis histórico, el objetivo es ofrecer una visión prospectiva sobre la cantidad de residuos que producirá el sistema de recolección de residuos en Bogotá y dimensionar sobre los retos que esta propone.

Marco Teórico

El proyecto de investigación emplea la metodología CRISP-DM (Comprensión del problema, Comprensión de los datos, Preparación de los datos, Modelado, Evaluación y Despliegue) para asegurar un enfoque estructurado en el análisis de datos históricos de residuos. Se utilizó el Análisis Exploratorio de Datos (EDA) con herramientas de Python (pandas, matplotlib, seaborn) para identificar inicialmente patrones y anomalías. Finalmente, para una predicción profunda de la generación de residuos por localidad, se aplicaron modelos matemáticos y estadísticos de series de tiempo, como SARIMAX o Prophet, que son adecuados para proyectar valores futuros basándose en patrones históricos. (Boschetti & Massaron, 2018)

Una vez finalizado el análisis de datos, la comunicación efectiva de los resultados es crucial, por lo que se emplean herramientas de visualización interactivas como Power BI y Tableau para presentar los hallazgos a los responsables de políticas públicas. El uso de la ciencia de datos en la gestión de residuos está comprobado que mejora la planificación y eficiencia del servicio; estudios en ciudades como Medellín y Quito han demostrado cómo la aplicación de

modelos de series temporales predice patrones de generación de residuos, permitiendo optimizar las rutas de recolección y reducir costos operativos de manera significativa.

Marco Conceptual

En el marco de la gestión de residuos, es importante conocer que un residuo sólido ordinario es todo residuo sólido de características no peligrosas que por su naturaleza, composición, tamaño, volumen y peso es recolectado, manejado, tratado o dispuesto normalmente por la persona concesionaria del servicio público de aseo (Unidad Administrativa de Servicios Públicos – Alcaldía Mayor de Bogotá, 2017) este servicio por lo general se hace bajo la modalidad de recolección puerta a puerta que consiste en la recolección de los residuos sólidos en el andén de la vía pública frente al predio del usuario.

El generador o productor es la persona que crea los residuos sólidos y los entrega al concesionario del servicio público de aseo para su recolección, siendo, por lo tanto, el usuario de dicho servicio. La presentación de los residuos sólidos es el acto en el que los usuarios exponen estos residuos, debidamente empacados, en los sitios e infraestructura específicamente previstos, para que el concesionario pueda proceder con su recolección y transporte. (Unidad Administrativa de Servicios Públicos – Alcaldía Mayor de Bogotá, 2017)

La Gestión Integral de Residuos Sólidos abarca el conjunto de actividades destinadas a reducir la generación de residuos, incluyendo su aprovechamiento (considerando características, volumen, procedencia, costos, y posibilidades de comercialización/valorización energética), así como el tratamiento y disposición final de los residuos no aprovechables.

La Persona Concesionaria (empresa de aseo) del servicio público de aseo es aquella encargada de una o varias actividades de la prestación del servicio público de aseo, en los términos del artículo 15 de la Ley 142 de 1994 y demás que la modifiquen o complementen, para

el caso de Bogotá en el 2025 se encuentra 5 empresas privadas que son Bogotá Limpia, Área Limpia, Ciudad Limpia, Promoambiental y Limpieza Metropolitana, cada una de ellas gana por licitación una ASE.

Localidad en Bogotá es una división administrativa de la ciudad. Es como un pequeño municipio dentro de la gran ciudad, con su propia identidad, características y organización. Cada localidad tiene un alcalde local y una junta administradora local que se encargan de tomar decisiones y gestionar los servicios públicos en su territorio, estas cuentan con una superficie determinada y límites establecidos entre las diferentes localidades.

Lo anterior debe incluirse en un Plan de gestión integral de residuos sólidos (PGIRS) el cual funciona como un instrumento de planeación municipal o regional que contiene un conjunto ordenado de objetivos, metas, programas, proyectos, actividades y recursos definidos por uno o más entes territoriales para el manejo de los residuos sólidos, basado en la política de gestión integral de los mismos, el cual se ejecutará durante un período determinado, basándose en un diagnóstico inicial, en su proyección hacia el futuro y en un plan financiero viable que permita garantizar el mejoramiento continuo del manejo de residuos y la prestación del servicio de aseo a nivel municipal o regional, evaluado a través de la medición de resultados. Corresponde a la Entidad territorial en este caso a la Alcaldía de Bogotá, la formulación, implementación, evaluación, seguimiento y control y actualización del PGIRS. (Unidad Administrativa de Servicios Públicos – Alcaldía Mayor de Bogotá, 2017)

Por otro lado, para la realización de los objetivos de este proyecto aplicado es necesario conocer la ciencia de datos y su alcance, esta se entiende como un dominio del conocimiento relativamente nuevo, aunque sus componentes centrales han sido estudiados e investigados por la comunidad de ciencias de la computación durante muchos años. Sus componentes incluyen

álgebra lineal, modelado estadístico, visualización, lingüística computacional, análisis de gráficos, aprendizaje automático, inteligencia de negocios y almacenamiento y recuperación de datos. (Boschetti & Massaron, 2018)

La implementación de la ciencia de datos incluye el uso de modelos predictivos, que son algoritmos de inteligencia artificial que aprovechan datos históricos y patrones identificados para proyectar valores o eventos futuros. Complementando esto, la visualización de datos es esencial, ya que es la herramienta que permite representar gráficamente la información y los hallazgos del análisis, lo que facilita enormemente la interpretación y la comunicación de los resultados a todas las partes interesadas. (Kane, 2017)

Marco Normativo

Históricamente, la forma en que hemos manejado nuestros desechos ha evolucionado de un simple acto de "desaparecer la basura" a un proceso complejo y multidisciplinario que involucra tecnología, economía y, fundamentalmente, un marco legal robusto. No podemos hablar de una gestión ambientalmente responsable y sostenible sin comprender las normas y leyes que la rigen.

Para ello no solo es necesario explorar los principios fundamentales que sustentan la legislación en materia de residuos, las normativas clave a nivel nacional y, si aplica, local, y cómo estas se traducen en responsabilidades y oportunidades para todos los actores. El manejo de los residuos sólidos no es solo un problema ambiental, sino también un desafío de salud pública y un motor potencial para la economía circular. Por tanto, la normativa legal no es un obstáculo, sino una herramienta indispensable para construir un futuro más limpio y sostenible.

A continuación, se indica el entramado legal que nos guía hacia una gestión de residuos más eficiente y consciente:

Resolución 799 del 31 de diciembre de 2012: Unidad Administrativa Especial de Servicios Públicos. “por la cual se establece el listado detallado los materiales reciclables y no reciclables para la separación en la fuente de los residuos sólidos domésticos en el distrito capital”

Decreto 555 de 2021: Alcaldía Mayor de Bogotá, Responde a las necesidades de recolección, transporte, aprovechamiento, tratamiento y fortalecimiento de las cadenas de gestión de residuos en el Distrito Capital, en el marco del concepto de economía circular y de desarrollo urbano sostenible.

Resolución 0631 del 17 de marzo de 2015: Ministerio de Ambiente y Desarrollo Sostenible, Por la cual se establecen los parámetros y los valores límites máximos permisibles en los vertimientos puntuales a cuerpos de aguas superficiales y a los sistemas de alcantarillado público y se dictan otras disposiciones.”

Decreto 1077 de 2015 1 de abril de 2016. Por el cual se modifica y adiciona el en lo relativo con el esquema de la actividad de aprovechamiento del servicio público de aseo y el régimen transitorio para la formalización de los recicladores de oficio, y se dictan otras disposiciones"

Resolución CRA- Comisión Reguladora del Agua Potable- 720 de 2015: Por la cual se establece el régimen de regulación tarifaria al que deben someterse las personas prestadoras del servicio público de aseo que atiendan en municipios de más de 5.000 suscriptores en áreas urbanas, la metodología que deben utilizar para el cálculo de las tarifas del servicio público de aseo y se dictan otras disposiciones.

Resolución 276 del 29 de abril de 2016: MINISTERIO DE VIVIENDA, CIUDAD Y TERRITORIO. “Por la cual se reglamentan los lineamientos del esquema operativo de la

actividad de aprovechamiento del servicio público de aseo y del régimen transitorio para la formalización de los recicladores de oficio acorde con lo establecido en el capítulo 5 del título 2 de la parte 3 del Decreto 1077 de 2015 adicionado por el Decreto 596 del 11 de abril de 2016”

Resolución 114 de 2003: Unidad Ejecutiva de Servicios Públicos -UESP- Aquí define como la actividad de Recolección Selectiva como “El servicio de recolección selectiva domiciliaria o de residuos aprovechables será realizado por el concesionario de aseo, mediante un sistema manual o mecanizado, con vehículos apropiados e independientes a los del sistema de recolección de residuos no aprovechables, en las zonas de reciclaje en que se ha dividido la ciudad.

Decreto 596 del 11 de abril de 2016: Ministerio de Vivienda, Ciudad y Territorio. “Por el cual se modifica y adiciona el Decreto 1077 de 2015 en lo relativo con el esquema de la actividad de aprovechamiento del servicio público de aseo y el régimen transitorio para la formalización de los recicladores de oficio, y se dictan otras disposiciones.

Ley 142 de 1994: Ley de Servicios Públicos Domiciliarios, es una legislación fundamental en Colombia que establece el régimen legal para la prestación de estos servicios esenciales.

Resolución 026 de 2018 de la UAESP: Reglamento Técnico Operativo para la prestación del Servicio Público de Aseo en la Ciudad de Bogotá.

Metodología

Método

La metodología que se propone para este proyecto de investigación que busca analizar datos históricos de la recolección de residuos domiciliarios y proyectar tendencias futuras, consistirá en un enfoque basado en CRISP-DM (Cross Industry Standard Process for Data Mining) ya que es ampliamente utilizado en proyectos de análisis de datos.

Esta metodología se compone de las siguientes fases:

Comprensión del Problema

Inicialmente debemos definir claramente el propósito del proyecto y los requisitos de análisis, por lo que es necesario realizar preguntas claves que nos orienten en la investigación de los datos que deseamos obtener, algunas de estas preguntas son:

¿Qué información histórica existe relacionada en la recolección de residuos ordinarios?
¿Quién me suministra esta información? ¿Cuál es el detalle de la información recolectada relacionada con la cantidad de residuos domiciliarios por localidad en la ciudad de Bogotá?
¿Según la información obtenida, que posibilidades tengo de análisis? ¿Qué factores podrían influir en las proyecciones futuras?

Una vez resultan las anteriores preguntas, es importante considerar que podremos definir los resultados esperados: modelos predictivos, visualizaciones, y reportes, todo relacionado en contexto con el problema en el marco de la gestión de residuos en Bogotá.

Comprensión de los Datos

Una vez se obtienen estos datos es importante iniciar con la familiarización de estos, entender su estructura y evaluar su calidad, ver su utilidad y datos erróneos.

Se debe identificar las fuentes de datos, registros históricos de recolección de residuos por localidad (2018-2024), explorar el formato, las variables necesarias como “Localidad”, “Fecha” y “Residuos en Toneladas”, la cantidad y la calidad de los datos, detectar posibles problemas, como valores faltantes, inconsistencias o datos atípicos.

Preparación de los Datos

La fase de preparación de los datos es crucial y comienza con la limpieza de datos, que implica corregir o eliminar valores faltantes, duplicados e inconsistentes para garantizar su calidad. Posteriormente, es fundamental estandarizar los datos, asegurando que todas las unidades y formatos sean uniformes para un análisis coherente. Finalmente, para enriquecer el estudio, se requiere generar nuevas variables y métricas adicionales (como tasas de generación de residuos e indicadores por localidad, cantidad por día, semana, mes o año), y combinar datos de diferentes fuentes para construir una base de datos consolidada y completamente interpretable.

Análisis Exploratorio de Datos (EDA)

Una vez se tienen preparados los datos se procede con un análisis exploratorio que consiste en identificar patrones, tendencias y relaciones dentro de los datos, momento en el cual se analizan las variaciones en la cantidad de residuos en la actividad recolección en cada una de las localidades de la ciudad de Bogotá a lo largo del tiempo, se exploraran correlaciones entre la generación de residuos y variables contextuales, se detectan patrones estacionales o tendencias lineales y utilizan herramientas como gráficos, histogramas y mapas para visualizar los datos, usaremos la herramienta Power BI para el EDA ya que ofrece varias ventajas clave, con un enfoque de Business Intelligence (BI) que buscan un enfoque rápido, visual e interactivo .

Modelado

Culminado el análisis exploratorio de los datos, para predecir la cantidad de residuos que producirán las localidades de la ciudad de Bogotá debemos construir modelos predictivos para proyectar tendencias futuras, esto se logrará mediante técnicas de modelado como Series Temporales (SARIMAX, Prophet), herramientas computacionales claves que permitirán analizar y predecir tendencias basadas en patrones históricos.

Finalmente realizamos un entrenamiento de los modelos, dividimos los datos en conjuntos de entrenamiento y prueba, evaluamos su desempeño con métricas como MAE, RMSE, MAPE y R^2 para medir la precisión de los modelos.

Evaluación de los Resultados

Después de realizar la modelación de los datos y evaluar su desempeño, procedemos a validar que estos resultados del análisis sean útiles y precisos. Para ello debemos comparar las proyecciones con datos históricos recientes para evaluar la precisión, y solo así poder tener una mayor tranquilidad de nuestros resultados, hay que revisar si los hallazgos responden a las preguntas planteadas inicialmente, si éstos no cumple con las expectativas será necesario reajustar el modelado.

Visualización y Comunicación

El éxito de este proyecto de investigación radica en la capacidad de presentar sus resultados de manera clara y comprensible mediante la creación de dashboards interactivos. Utilizando herramientas como Power BI, Matplotlib o Seaborn, estos paneles deben mostrar tendencias históricas, la generación de residuos desglosada por localidad, y proyecciones clave para el periodo 2025-2029. El fin último es comunicar estos hallazgos (insights) del proyecto aplicado de tal forma que sean accesibles y sirvan como material de consulta fundamental para

las partes interesadas, como el gobierno local y las entidades dedicadas a la gestión de residuos domiciliarios.

Despliegue y Recomendaciones

Como fase final, los hallazgos del análisis nos permitirán sugerir mejor toma de decisiones, se entregarán las proyecciones y recomendaciones para mejorar la gestión de residuos sólidos domiciliarios, sugerir acciones específicas de incentivar el reciclaje sobre las localidades con más generación de residuos sólidos, como optimización de rutas de recolección o campañas de concientización.

Resultados

Comprensión del Problema

Este proyecto de investigación tiene como objetivo principal realizar un análisis exhaustivo y detallado de la generación de residuos domiciliarios en Bogotá por localidad, cubriendo el período 2018-2024 e incluyendo predicciones para los próximos cinco años. El análisis se centrará en la recopilación, procesamiento y estudio de datos históricos para identificar las localidades con mayor tendencia a generar residuos, con el fin de generar información valiosa para el Gobierno de la Ciudad y entidades relacionadas. La base de datos esencial para este estudio debe contener la cantidad en peso (toneladas) de residuos recogidos, la fecha y la localidad de origen, a pesar de que las bases de datos abiertas no poseen el nivel de detalle necesario; por ello, la fuente primaria y más detallada de esta información clave es la Unidad Administrativa Especial de Servicios Públicos (UAESP).

Por esta razón, como punto de partida del proyecto, se radicó la solicitud N°. 20247000565192 (16 de octubre de 2024) ante la UAESP a través de su plataforma de atención al público. La UAESP trasladó esta solicitud al Consorcio Proyección Capital (interventoría del servicio de aseo en Bogotá desde el 18 de febrero de 2018) mediante el comunicado N°. 20242000216691 (01 de noviembre de 2024).

En respuesta, el Consorcio Proyección Capital emitió el comunicado UAESP-CPC-TER-3135-24 (29 de noviembre de 2024) con el anexo "Anexo 20242000216691" el cual constituirá el sustento de nuestra base de datos para el análisis exhaustivo de la generación de residuos domiciliarios en Bogotá (desde el 18 de febrero de 2018 hasta el 31 de octubre de 2024).

Comprensión de los Datos

La base de datos proporcionada por el Consorcio Proyección Capital en el "Anexo 20242000216691" del comunicado con radicado UAESP-CPC-TER-3135-24 es un archivo Excel compuesto por siete hojas, Estas hojas detallan la información anual desde 2018 hasta 2024.

La base de datos utilizada para el análisis de residuos domiciliarios es un registro exhaustivo que detalla la operación de recolección y disposición final en el Relleno Sanitario Parque de Innovación Doña Juana. Contiene información temporal (Mes, Fecha de entrada y salida), de control vehicular (Consecutivo, Placa, Número de Vehículo), operativa (Concesión, Macrorruta, Microrruta, ASE, Tipo de Servicio y Zona de Descarga), y de pesaje (Peso de entrada, Peso de salida, Peso Residuos en Kg y Peso Residuos en Ton). Además, incluye la Localidad de origen, la cual es una de las 20 divisiones administrativas de Bogotá. De todas estas columnas, las variables clave seleccionadas para el análisis exhaustivo de la generación de residuos domiciliarios por localidad, y para la fundamentación de proyecciones futuras (2018-2024 y próximos cinco años), son la Fecha de entrada, el Peso Residuos (Ton) y la Localidad.

Preparación de los Datos

Para asegurar una correcta interpretación, la preparación de los datos se llevó a cabo utilizando Power Query (un motor de transformación de Microsoft) para limpiar y estandarizar la base de datos "Anexo 20242000216691", corrigiendo valores faltantes, duplicados e inconsistencias. Se creó una nueva base de datos reducida que solo contiene la información necesaria para el análisis, eliminando tipos de servicio irrelevantes, dejando únicamente la "Recolección Domiciliaria". Un paso crucial de limpieza fue corregir errores tipográficos en la columna "Localidad" (mayúsculas, tildes y espacios inconsistentes), lo que garantizó la

uniformidad y correcta categorización de los datos. Finalmente, se eliminaron todas las columnas no esenciales, dejando solo Fecha de entrada, Peso Residuos (Ton) y Localidad, lo que optimizó el rendimiento del análisis al reducir el tamaño de la base de datos.

Figura 1

Base de Datos para Análisis Exploratorio (EDA)

FECHA ENTRADA	LOCALIDAD	PESO RESIDUOS (Ton)
12/02/2018 12:57:06 a.m.	Barrios Unidos	4,62
12/02/2018 01:09:35 a.m.	Engativá	7,96
12/02/2018 01:11:42 a.m.	Barrios Unidos	5,92
12/02/2018 01:20:34 a.m.	Engativá	5,00
12/02/2018 01:32:41 a.m.	Chapinero	4,54
12/02/2018 01:35:49 a.m.	Engativá	6,73
12/02/2018 01:37:15 a.m.	Engativá	7,58
12/02/2018 02:07:40 a.m.	Chapinero	3,69
12/02/2018 02:31:12 a.m.	Chapinero	12,76
12/02/2018 02:32:38 a.m.	Engativá	10,63
12/02/2018 03:09:51 a.m.	Engativá	7,19
12/02/2018 03:41:24 a.m.	Chapinero	7,96

Adicionalmente, se creó una tabla calendario relacionada con nuestros datos en Power BI que actuó como una dimensión de tiempo central y unificada para un correcto análisis de datos a través de diferentes períodos de tiempo.

De igual manera, se desarrolló una medida en DAX usando la función COALESCE que reemplaza los valores nulos o blancos por un valor igual a cero (0), de esta manera obtenemos graficas correctas en el análisis de datos temporales, también nos permitirá generar nuevas bases de datos por localidad para las predicciones previstas en esta investigación.

Análisis Exploratorio de Datos (EDA)

Para el análisis de los datos usamos la aplicación Power BI que es un conjunto de herramientas de Microsoft para analizar y visualizar datos de una manera sencilla e intuitiva,

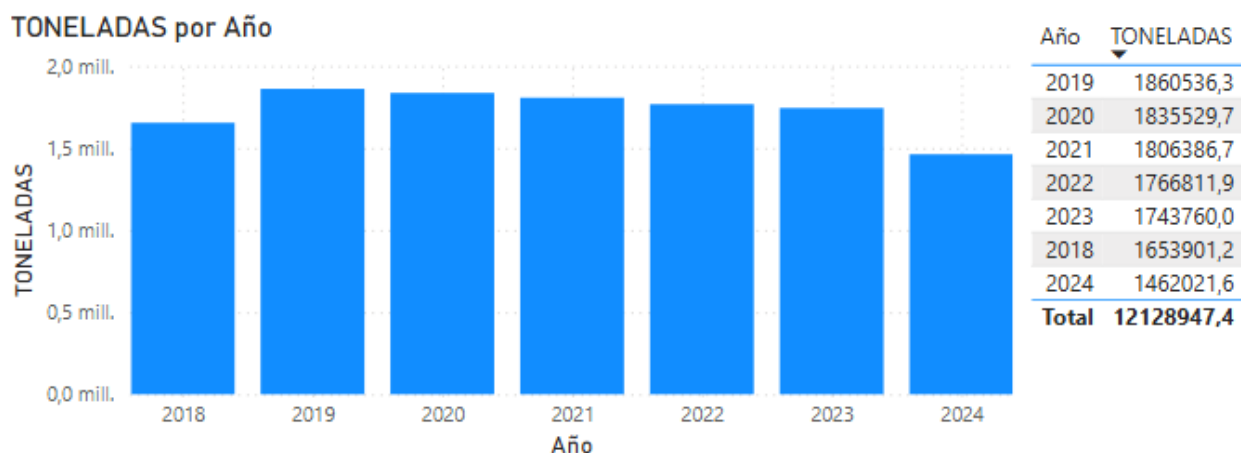
para ello se tomó la data que con anterioridad se le realizó una preparación preliminar y que contiene los datos necesarios para nuestro análisis.

Introducción y Resumen de Datos

El análisis abarca datos diarios desde el 12 de febrero del 2018 al 31 de octubre de 2024, como datos generales tenemos:

Figura 2

Total Toneladas Recolectadas por Año (12 de febrero del 2018 al 31 de octubre de 2024)

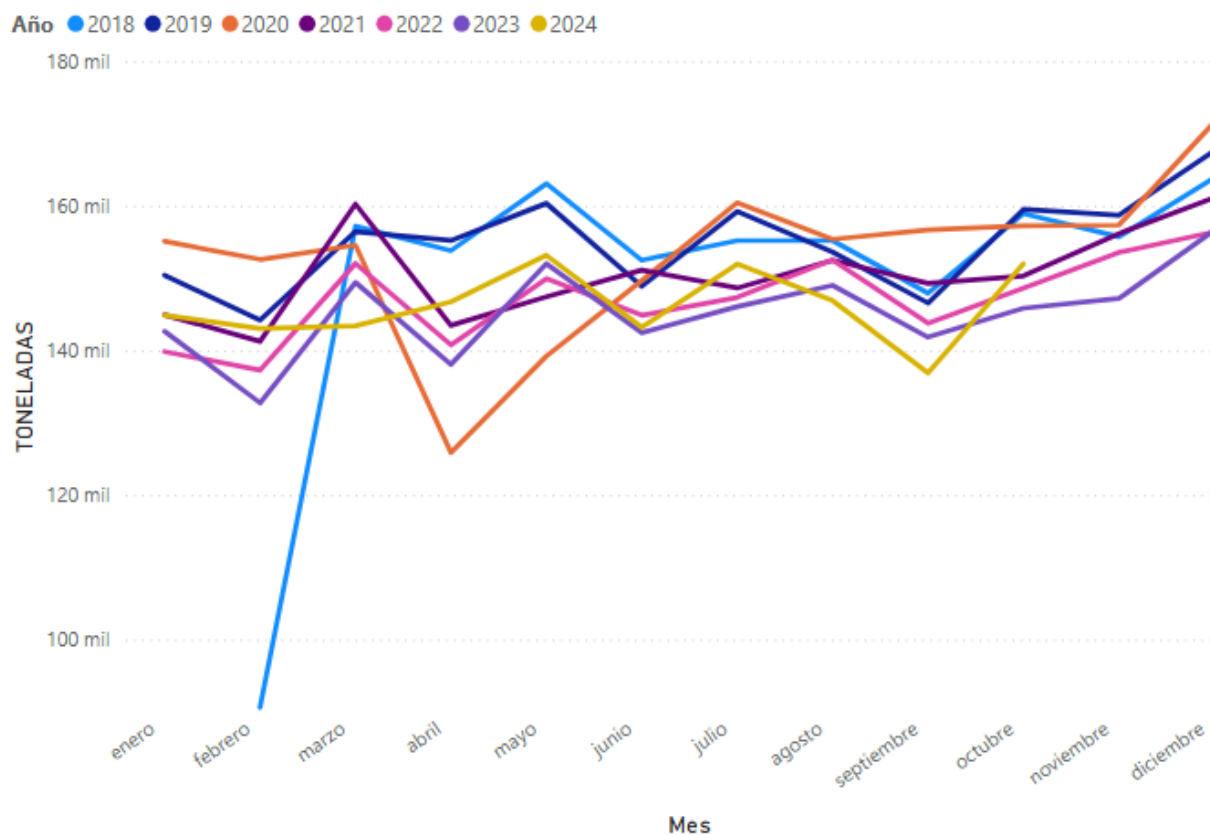


El análisis de la recolección de residuos domiciliarios en Bogotá, que abarca desde el 12 de febrero de 2018 hasta el 31 de octubre de 2024, revela una cantidad total recolectada de 12.128.947 toneladas, siendo 2019 el año con el mayor volumen registrado, con 1.860.536 toneladas. Durante este periodo, los promedios consolidados de recolección son de 4.942 toneladas diarias, 198.835 toneladas mensuales y 1.732.706 toneladas anuales. Es importante destacar que el alcance de los datos tiene la limitación de no incluir la recolección del 1 de enero al 11 de febrero de 2018, ni del 1 de noviembre al 31 de diciembre de 2024.

Análisis de Tendencias Temporales

Figura 3

Grafica Toneladas Recolectadas por Mes (12 de febrero del 2018 al 31 de octubre de 2024)



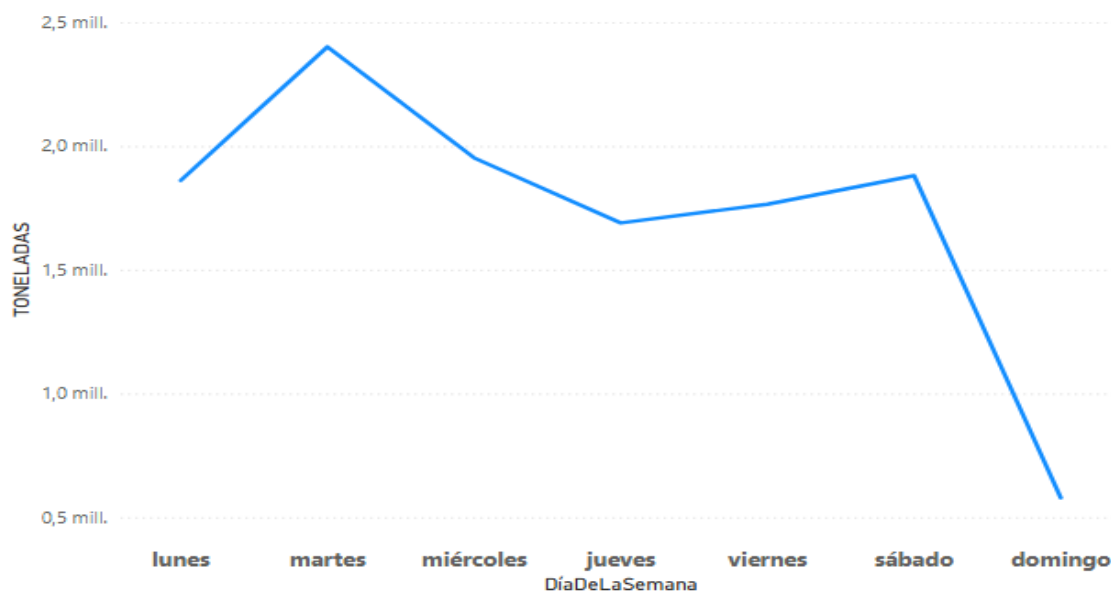
Como podemos observar en la figura anterior, describe la cantidad de toneladas recogidas por mes entre el 12 de febrero del 2018 al 31 de octubre de 2021, de esta podemos concluir que:

El análisis de la generación de residuos domiciliarios en Bogotá entre febrero de 2018 y octubre de 2024 revela tres patrones clave: una rápida consolidación inicial a mediados de 2018 por encima de las 150 mil toneladas mensuales, una fuerte estacionalidad recurrente, y un notable impacto por la pandemia de COVID-19. A pesar de las fluctuaciones (principalmente entre 145 mil y 165 mil toneladas), no existe una tendencia de crecimiento ascendente sostenida. La estacionalidad se caracteriza por picos consistentes a finales (noviembre/diciembre) y

principios de año (marzo/abril), asociados a festividades y consumo, y valles marcados en enero/febrero. El evento más extremo fue la caída a 125.882 toneladas en abril de 2020 debido a la pandemia, seguida de una rápida recuperación y el registro del pico histórico más alto en diciembre de 2020 (171.545 toneladas). Finalmente, los datos recientes de 2024 sugieren una aceleración que podría indicar un nuevo pico alto al cierre del año.

Figura 4

Grafica Toneladas Recolectadas por Día de la Semana Bogotá (12 de febrero del 2018 al 31 de octubre de 2024)



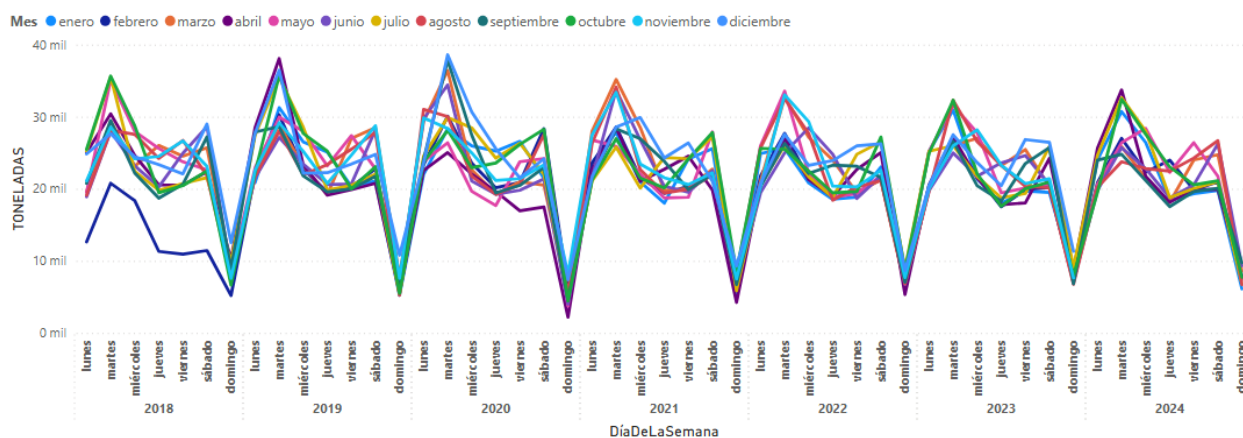
En la figura anterior se muestra la distribución consolidada del total de toneladas de residuos recolectadas por cada Día de la Semana a lo largo de todo el periodo de análisis, lo que indica que se trata de la suma total de toneladas para cada día de la semana a lo largo de los años.

El análisis de la recolección de residuos domiciliarios por día de la semana revela un ciclo operativo semanal marcado, donde la máxima actividad se concentra al inicio de la semana. El martes registra el mayor volumen total de recolección (más de 2.5 millones de toneladas),

seguido de cerca por el lunes (cerca de 1.9 millones de toneladas), lo que indica un esfuerzo concentrado para manejar los residuos acumulados. Los volúmenes disminuyen progresivamente a mitad de semana, alcanzando su punto más bajo en días laborales el jueves (por encima de 1.5 millones de toneladas). Se observa un repunte significativo el sábado (cercano a 2 millones de toneladas) para la recolección pre-dominical, y la caída más drástica y el punto más bajo de la serie se registra el domingo (0.5 millones de toneladas o menos), evidenciando que es un día de muy baja o nula operación de recolección en Bogotá.

Figura 5

Residuos Domiciliarios Recolectados por Día de la Semana, Mes y Año (12 de febrero del 2018 al 31 de octubre de 2024)



Como se puede observar en la figura anterior todas las líneas siguen casi exactamente la misma forma, lo que confirma que el horario y la logística de recolección semanal es el factor más fuerte en la determinación del volumen diario.

El análisis detallado por mes y día de la semana confirma un patrón semanal inmutable: el martes es consistentemente el día de máxima recolección en todos los meses, superando las 200 mil toneladas en periodos pico, lo que subraya la gestión de la carga acumulada del fin de

semana. Por el contrario, el domingo registra uniformemente la menor actividad operativa. Aunque el ciclo semanal (pico el martes, caída notable de miércoles a jueves y desplome el domingo) se mantiene, la posición vertical de las líneas indica la estacionalidad: diciembre, mayo y marzo son los meses más activos históricamente (combinando alta operación y mayor generación de residuos), mientras que enero y febrero son los meses menos activos con volúmenes totales significativamente inferiores.

Análisis Geográfico y Segmentación por Localidad

Para comprender la recolección de residuos domiciliarios en la ciudad de Bogotá, el análisis desglosa el comportamiento individual de la recolección en las 20 localidades que componen la ciudad de Bogotá. Esta segmentación se realiza sobre el total de residuos recogidos en cada área regional específica durante el periodo comprendido entre el 12 de febrero de 2018 y el 31 de octubre de 2024.

Figura 6

Residuos Domiciliarios Recolectados por Localidad (12 de febrero del 2018 al 31 de octubre de 2024)

LOCALIDAD	TONELADAS	Porcentaje Total Recogido
Suba	1771281,2	14,60 %
Kennedy	1721664,0	14,19 %
Engativá	1204472,7	9,93 %
Bosa	956501,2	7,89 %
Usaquén	895972,4	7,39 %
Ciudad Bolívar	717503,5	5,92 %
Fontibón	522808,4	4,31 %
San Cristóbal	516721,6	4,26 %
Puente Aranda	491292,9	4,05 %
Rafael Uribe Uribe	487484,4	4,02 %
Usme	465270,4	3,84 %
Chapinero	459856,1	3,79 %
Tunjuelito	394722,6	3,25 %
Barrios Unidos	349112,8	2,88 %
Teusaquillo	316758,1	2,61 %
Los Mártires	273087,3	2,25 %
Antonio Nariño	238879,8	1,97 %
Santa Fe	234707,5	1,94 %
La Candelaria	83942,4	0,69 %
Sin Definir	25901,0	0,21 %
Sumapaz	1007,1	0,01 %
Total	12128947,4	100,00 %

La localidad que más recolecto residuos domiciliarios es Suba con un millón setecientos setenta y un mil doscientos ochenta y una 1.771.281 toneladas, equivalentes al 14,60 % de la totalidad de los residuos domiciliarios recolectados en el periodo de análisis, seguido por la localidad de Kennedy, Engativá, Bosa y Usaquén.

La localidad que obtuvo menos residuos domiciliarios recolectados fue Sumapaz con mil siete 1.007 toneladas durante el periodo de análisis, con un equivalente al 0,01 % de la totalidad de los residuos recolectados en el periodo de análisis.

La base datos contiene toneladas de residuos domiciliarios que no fueron categorizadas en alguna de las 20 localidades y que son reportadas como localidad sin definir, estas contienen

veinticinco mil novecientos un toneladas 25.901 de residuos domiciliarios, equivalentes a 0,21 % de la totalidad de los residuos recolectados en el periodo de análisis.

Figura 7

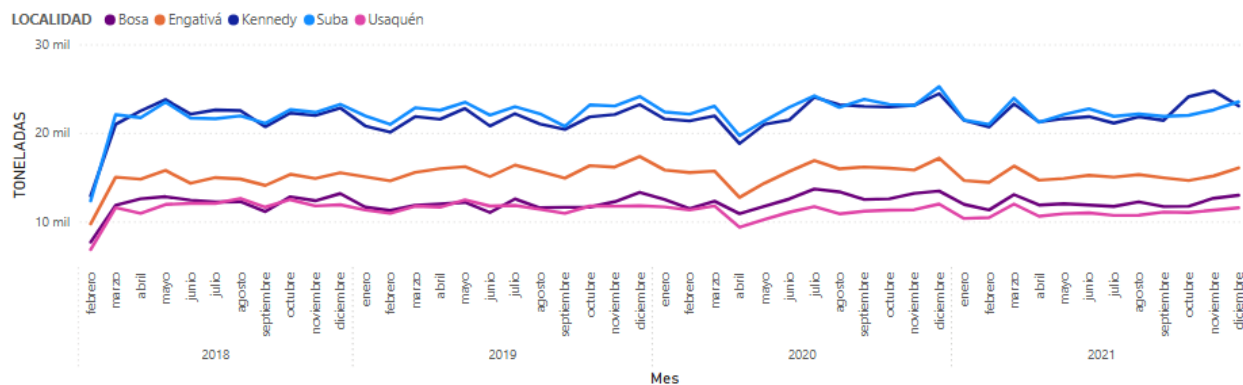
Localidades con Mayor Cantidad de Residuos Domiciliarios Recolectados (12 de febrero del 2018 al 31 de octubre de 2024)

LOCALIDAD	TONELADAS	Porcentaje Total Recogido
Suba	1771281,2	14,60 %
Kennedy	1721664,0	14,19 %
Engativá	1204472,7	9,93 %
Bosa	956501,2	7,89 %
Usaquén	895972,4	7,39 %
Total	6549891,5	54,00 %

La mayor cantidad de residuos domiciliarios recogidos en la ciudad de Bogotá provienen de las localidades Suba, Kennedy, Engativá, Bosa y Usaquén, equivalentes al 54 % de la totalidad de los residuos recolectados en el periodo de análisis, por lo que nos enfocamos en estas localidades para un análisis más detallado.

Figura 8

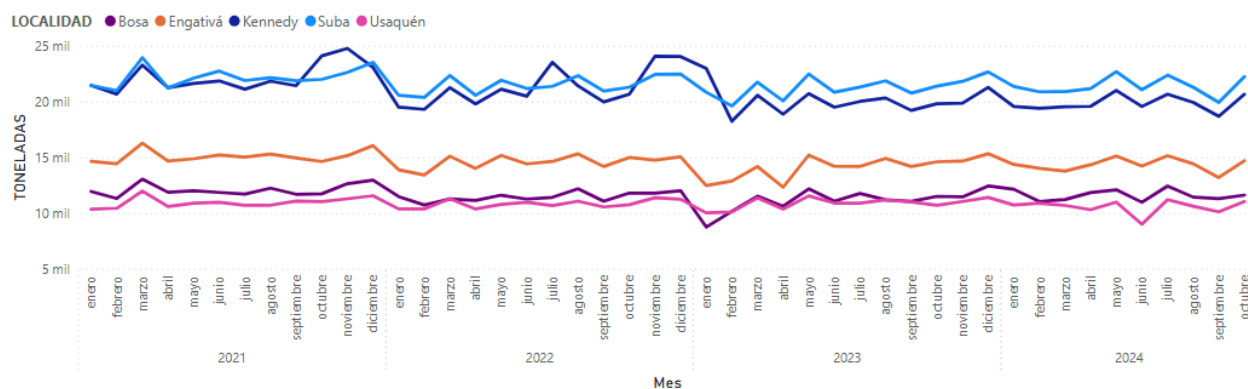
Toneladas Recogidas por Mes en Localidades Especificas (12 de febrero del 2018 al 31 de diciembre de 2021)



El análisis de la generación de residuos domiciliarios por localidad entre febrero de 2018 y diciembre de 2021 muestra una clara disparidad en los volúmenes, dominada consistentemente por Kennedy y Suba, que se mantienen como las mayores generadoras (superando las 20-22 mil toneladas), seguidas por Engativá en un rango intermedio (14-17 mil toneladas), y con Bosa y Usaquén registrando las menores cantidades (10-13 mil toneladas). Este patrón se enmarca en una fuerte estacionalidad cíclica, con picos de recolección sistemáticos a finales de año (noviembre y diciembre) debido a festividades, y valles notables alrededor de abril y mayo. Además, se registró una caída visible en todas las localidades durante marzo/abril de 2020 por el inicio de la pandemia, aunque los volúmenes se recuperaron rápidamente a niveles prepandemia hacia finales de 2020 y 2021, sin mostrar un crecimiento sostenido a lo largo del periodo.

Figura 9

Toneladas Recogidas por Mes en Localidades Específicas (1 de enero del 2021 al 31 de octubre de 2024)



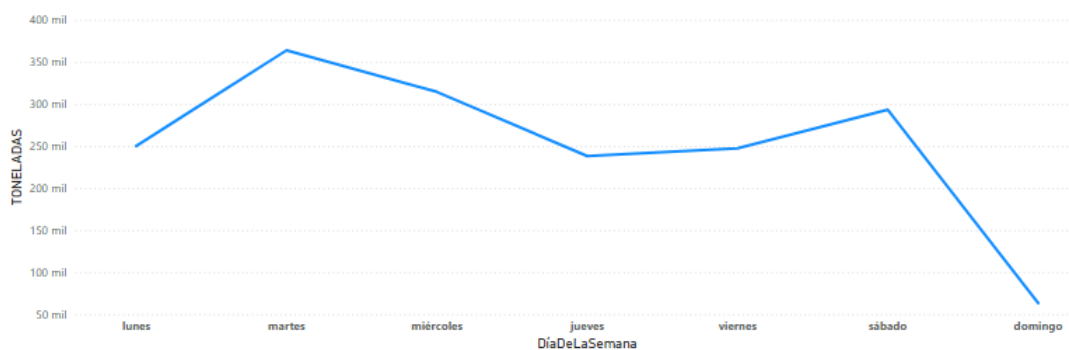
Entre 2021 y 2024, se mantuvo una jerarquía de residuos constante y marcada entre las localidades: Kennedy y Suba fueron consistentemente las mayores generadoras (20-25 mil toneladas), seguidas por Engativá en tercer lugar (14-17 mil toneladas), y con Bosa y Usaquén registrando los menores volúmenes (10-12 mil toneladas). Este patrón se explica por la densidad

demográfica y actividad económica de cada zona. La tendencia dominante es un fuerte ciclo estacional anual, caracterizado por picos de recolección en noviembre y diciembre (debido a festividades y consumo), y valles consistentes entre enero y abril (con el mínimo más bajo registrado en febrero de 2023). En general, las curvas de las localidades muestran una notable estabilidad en el volumen total de residuos a lo largo del periodo, sin evidenciarse un crecimiento sostenido o exponencial, más allá de la variabilidad estacional.

Suba. Es la localidad con mayor cantidad de residuos domiciliarios recolectados en la Ciudad de Bogotá, y su estacionalidad semanal se puede observar en la siguiente imagen que nos muestra la totalidad de la cantidad de residuos recogidos en el día de la semana durante el periodo de análisis:

Figura 10

Toneladas Recogidas por Día de la Semana en la Localidad de Suba (12 de febrero del 2018 al 31 de octubre de 2024)



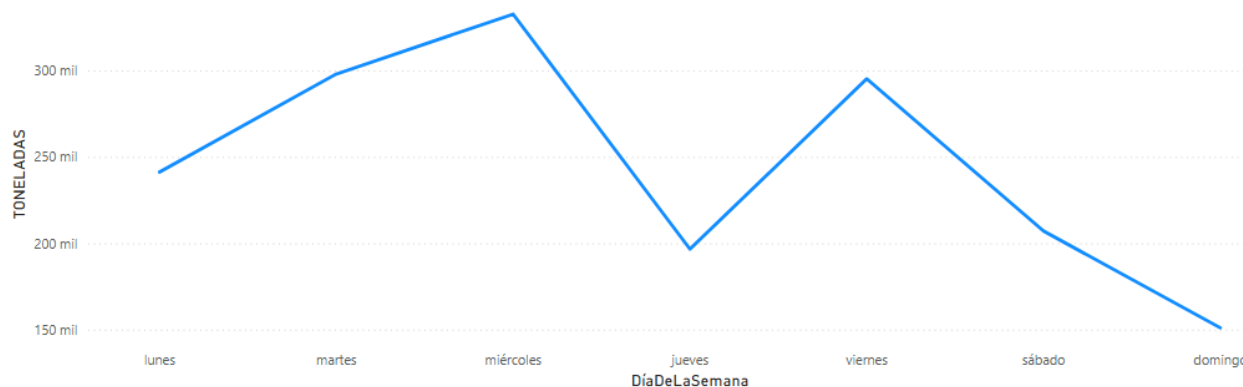
El análisis de la recolección diaria de residuos domiciliarios revela un patrón semanal altamente concentrado en el esquema operativo de Bogotá. La máxima actividad se registra el martes, alcanzando el pico con aproximadamente 360,000 toneladas, lo que sugiere la gestión del volumen acumulado. La recolección inicia fuerte el lunes (250,000 toneladas) y desciende

progresivamente de miércoles a viernes, manteniendo una relativa estabilidad. Hay un repunte notable el sábado, con cerca de 300,000 toneladas (el segundo pico), antes de la pausa dominical. Finalmente, el domingo presenta la caída más drástica y el punto más bajo, con solo unas 60,000 toneladas, confirmando que es el día de mínima operación o frecuencia de servicio. La disparidad es significativa, ya que el volumen del martes es seis veces mayor que el del domingo.

Kennedy. Es la segunda localidad con mayor cantidad de residuos domiciliarios recolectados en la Ciudad de Bogotá, y su estacionalidad semanal se puede observar en la siguiente imagen que nos muestra la totalidad de la cantidad de residuos recogidos en el día de la semana durante el periodo de análisis:

Figura 11

Toneladas Recogidas por Día de la Semana en la Localidad de Kennedy (12 de febrero del 2018 al 31 de octubre de 2024)



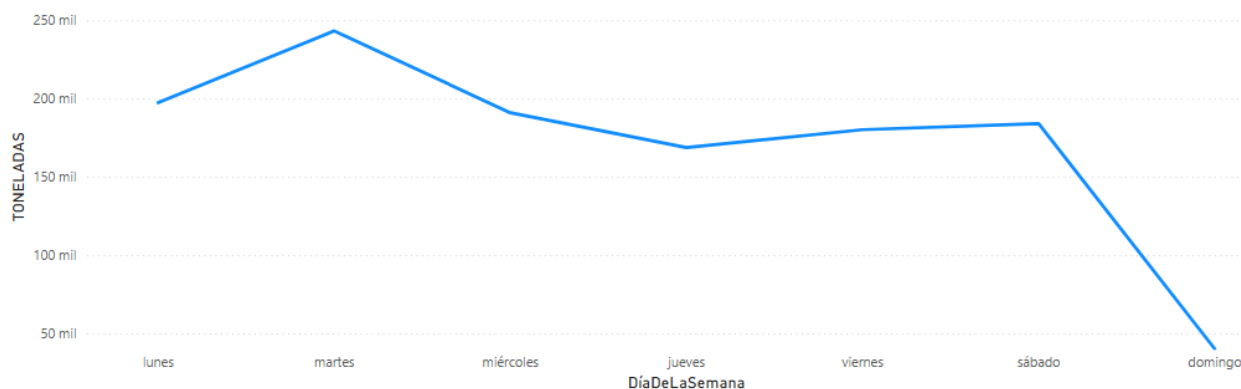
El patrón semanal de recolección de residuos en la localidad de Kennedy es bimodal y atípico, caracterizado por dos picos altos: el miércoles registra el mayor volumen con aproximadamente 335,000 toneladas, seguido por el viernes con cerca de 300,000 toneladas. Se observa una caída notable el jueves (aproximadamente 200,000 toneladas), creando un valle

intermedio entre los dos picos. La menor recolección se produce el domingo, cayendo drásticamente a unas 150,000 toneladas. Este esquema de dos picos intensos (miércoles y viernes) separados por un valle sugiere una programación de rutas de recolección altamente concentrada en días específicos, reflejando el calendario de servicio de la localidad.

Engativá. Es la tercera localidad con mayor cantidad de residuos domiciliarios recolectados en la Ciudad de Bogotá, y su estacionalidad semanal se puede observar en la siguiente imagen que nos muestra la totalidad de la cantidad de residuos recogidos en el día de la semana durante el periodo de análisis:

Figura 12

Toneladas Recogidas por Día de la Semana en la Localidad de Engativá (12 de febrero del 2018 al 31 de octubre de 2024)



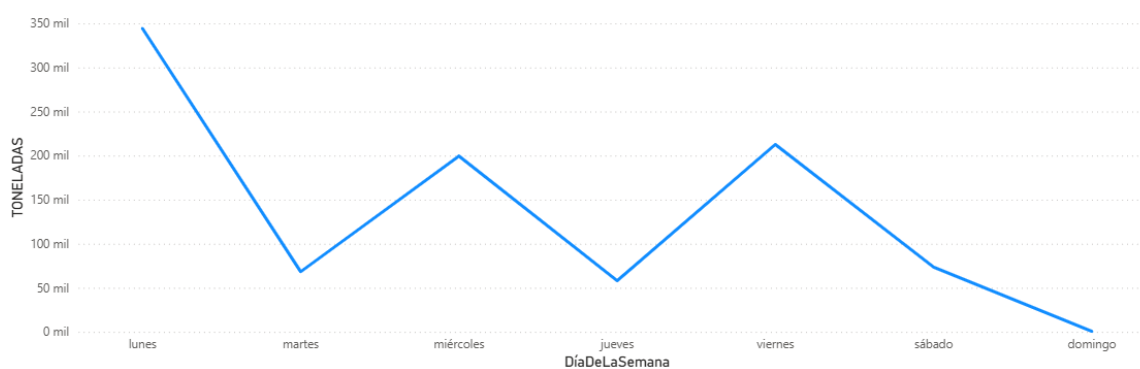
La Localidad de Engativá presenta un patrón semanal de recolección con una clara concentración al inicio de la semana laboral y una actividad mínima el domingo. La mayor cantidad de residuos se recoge el martes, alcanzando aproximadamente 245,000 toneladas. A lo largo de la semana, los niveles se mantienen en un rango medio a alto de lunes a sábado, pero la recolección cae dramáticamente a su punto más bajo el domingo, registrando solo unas 40,000 toneladas. Este patrón subraya que, aunque la actividad es constante, la máxima capacidad

operativa está dedicada a gestionar los residuos al comienzo de la semana, con un servicio de recolección casi nulo el día domingo.

Bosa. Es la cuarta localidad con mayor cantidad de residuos domiciliarios recolectados en la Ciudad de Bogotá, y su estacionalidad semanal se puede observar en la siguiente imagen que nos muestra la totalidad de la cantidad de residuos recogidos en el día de la semana durante el periodo de análisis:

Figura 13

Toneladas Recogidas por Día de la Semana en la Localidad de Bosa (12 de febrero del 2018 al 31 de octubre de 2024)

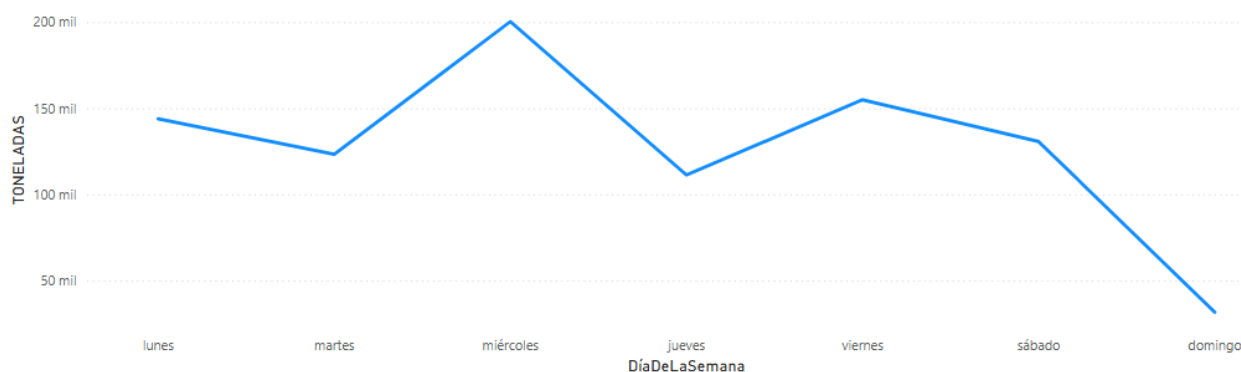


La recolección de residuos en la localidad de Bosa sigue un patrón de frecuencia alterna (día por medio), lo que genera picos y valles muy pronunciados, y una actividad casi nula el domingo. El día con la mayor cantidad de recolección es el lunes, alcanzando casi 350,000 toneladas, lo cual refleja la gestión de la acumulación de residuos del fin de semana. El viernes registra el segundo pico más alto, con aproximadamente 215,000 toneladas. En contraste, los volúmenes caen abruptamente en los días intermedios (el martes a 70,000 toneladas y el jueves a 60,000 toneladas), mientras que el domingo registra el nivel más bajo, con una recolección muy cercana a cero.

Usaquén. Es la quinta localidad con mayor cantidad de residuos domiciliarios recolectados en la Ciudad de Bogotá, y su estacionalidad semanal se puede observar en la siguiente imagen que nos muestra la totalidad de la cantidad de residuos recogidos en el día de la semana durante el periodo de análisis:

Figura 14

Toneladas Recogidas por Día de la Semana en la Localidad de Usaquén (12 de febrero del 2018 al 31 de octubre de 2024)



La localidad de Usaquén presenta un claro patrón bimodal de recolección semanal, con los picos de mayor actividad concentrados en la mitad y final de la semana. El miércoles es el día de mayor recolección, alcanzando las 200,000 toneladas, seguido de un pico secundario notable el viernes con aproximadamente 155,000 toneladas. La recolección inicia en un nivel alto el lunes (145,000 toneladas) y desciende a un valle intermedio el jueves (115,000 toneladas) antes del segundo pico. Finalmente, el domingo registra la caída más drástica y el punto más bajo, con apenas 30,000 toneladas, lo que sugiere una pausa casi total del servicio ese día y explica el volumen superior recolectado el lunes. Este patrón bimodal indica que las rutas de recolección en Usaquén priorizan y concentran la actividad en días específicos de la semana.

Modelado

Bases de Datos para el Modelado

El proyecto se enfoca en el modelado predictivo de series temporales para pronosticar la cantidad diaria de residuos sólidos (en toneladas) que ingresarán al relleno sanitario Doña Juana. El estudio abarca un periodo de 2,437 registros diarios, desde el 1 de marzo de 2018 hasta el 31 de octubre de 2024. Se seleccionaron tres bases de datos (BD) clave: la BD Bogotá (total de residuos diarios de la ciudad), la BD Suba (la mayor contribuyente) y la BD Kennedy (la segunda mayor contribuyente). A cada una de estas series se le aplicarán los modelos predictivos SARIMAX y Prophet para generar un pronóstico detallado que cubra el período del 1 de noviembre de 2024 al 31 de diciembre de 2029.

Selección de Técnicas de Modelado

Para el modelado predictivo, se han seleccionado los modelos SARIMAX y Prophet debido a sus fortalezas complementarias. SARIMAX (Seasonal Autoregressive Integrated Moving Average with eXogenous regressors) se elige por su solidez estadística y alta interpretabilidad, ya que incorpora términos estacionales (SARIMA) y, crucialmente, permite añadir variables exógenas (X), lo que potencialmente aumenta la precisión al modelar influencias externas junto con la estructura interna de la serie. Por otro lado, Prophet (desarrollado por Meta) se distingue por su facilidad de uso y robustez ante datos desordenados y anomalías; modela la serie como una regresión aditiva de tendencia, estacionalidad (semanal y anual modelada con series de Fourier) y efectos de días festivos/eventos, y está diseñado para ser intuitivo y automatizado, manejando de forma nativa valores atípicos y grandes cambios de tendencia.

Se descartaron varios modelos predictivos por no ser adecuados para series temporales con la fuerte estacionalidad presente en los datos de residuos. Se evitó el ARIMA simple por no

manejar la estacionalidad, y la Regresión Lineal Simple por violar el supuesto de independencia de observaciones y no poder capturar patrones complejos. El Suavizado Exponencial (como Holt-Winters) se consideró demasiado complejo para manejar la estacionalidad múltiple de los datos. Finalmente, los modelos de Machine Learning (ML) y Redes Neuronales (LSTM) fueron excluidos por ser "cajas negras" con baja interpretabilidad (crucial para la toma de decisiones), requerir altos costos computacionales y no manejar la estacionalidad de forma nativa. Por lo tanto, se eligieron SARIMAX y Prophet porque ofrecen un equilibrio superior: SARIMAX brinda precisión estadística y modelado explícito de la estacionalidad, y Prophet, robustez y un manejo intuitivo de los ciclos semanales y eventos.

Análisis Descriptivo y Descomposición de la Serie de Tiempo

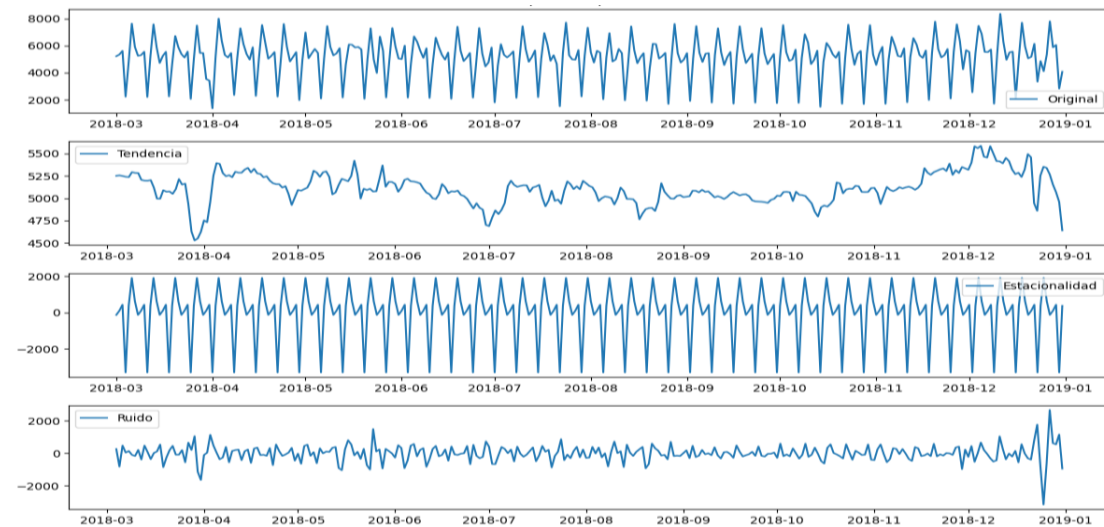
Se detalla el análisis descriptivo y la descomposición de las series de tiempo para cada base de datos. Este proceso es fundamental para comprender a fondo los patrones subyacentes (tendencia, estacionalidad y componente residual). Para una mejor interpretación visual, los resultados se proyectan con una desagregación anual, considerando el volumen total de los datos.

Base de Datos Bogotá (Toneladas Diarias de Residuos Domiciliarios). Para la descomposición de la serie de tiempo usamos Python en el cual se realiza un análisis exploratorio y una descomposición de series de tiempo sobre un archivo CSV que contiene las toneladas diarias de residuos domiciliarios llevadas por día al relleno sanitario Parque de innovación Doña Juana desde el 1 de marzo del 2018 al 31 de octubre de 2024 de la ciudad de Bogotá.

En Python cargamos los datos de toneladas a lo largo del tiempo, se limpian, se analizan descriptivamente y, finalmente, se visualizan sus componentes de tendencia, estacionalidad y residuos de manera anual.

Figura 15

Descomposición de la Serie de Tiempo Base de Datos Bogotá 2018 (Toneladas Diarias de Residuos Domiciliarios)

**Figura 16**

Descomposición de la Serie de Tiempo Base de Datos Bogotá 2019 (Toneladas Diarias de Residuos Domiciliarios)

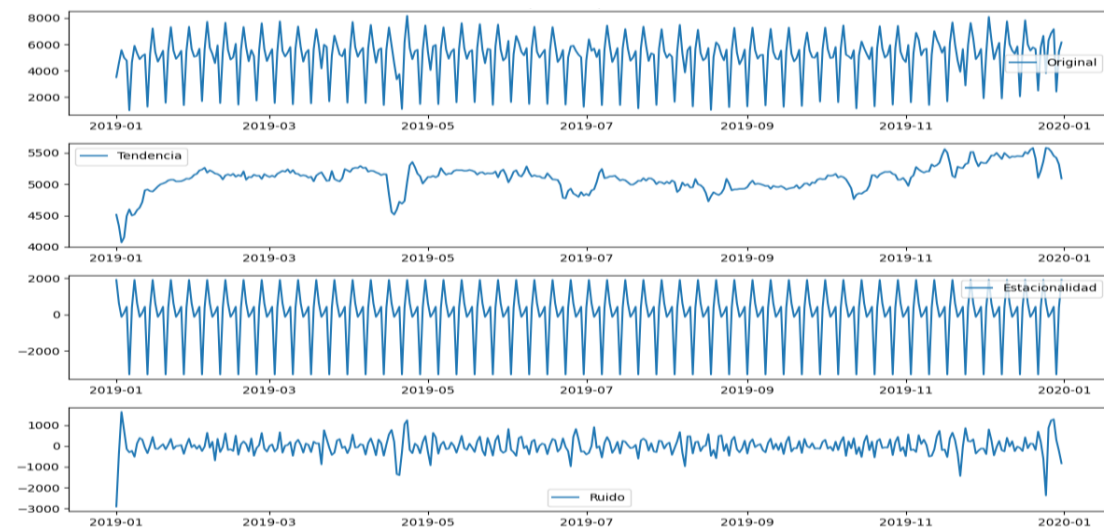
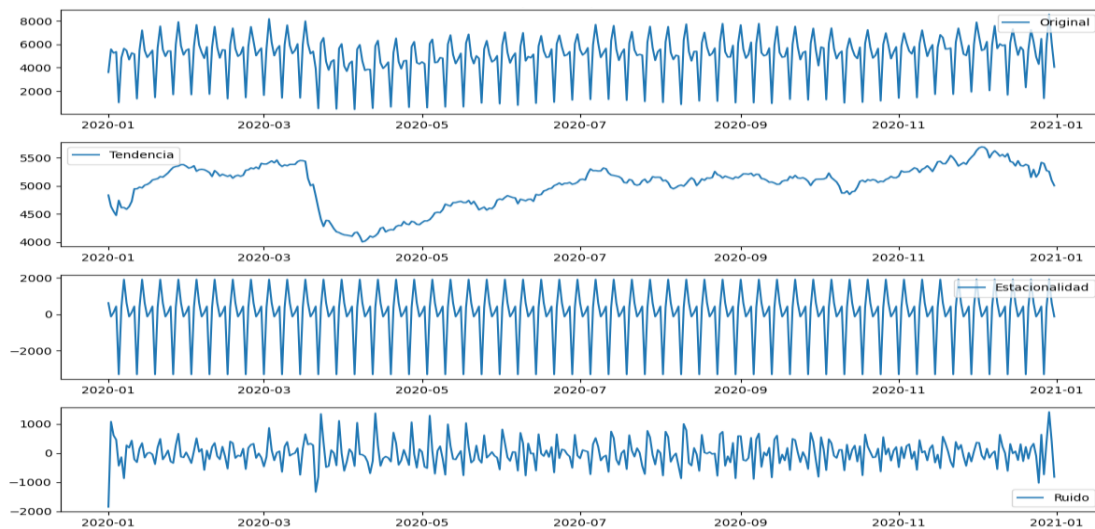


Figura 17

Descomposición de la Serie de Tiempo Base de Datos Bogotá 2020 (Toneladas Diarias de Residuos Domiciliarios)

**Figura 18**

Descomposición de la Serie de Tiempo Base de Datos Bogotá 2021 (Toneladas Diarias de Residuos Domiciliarios)

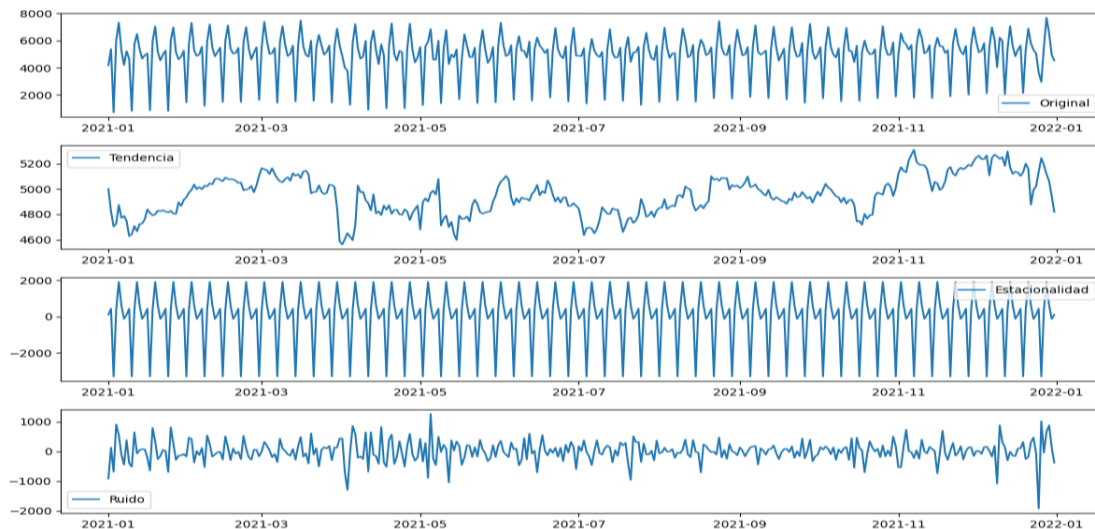
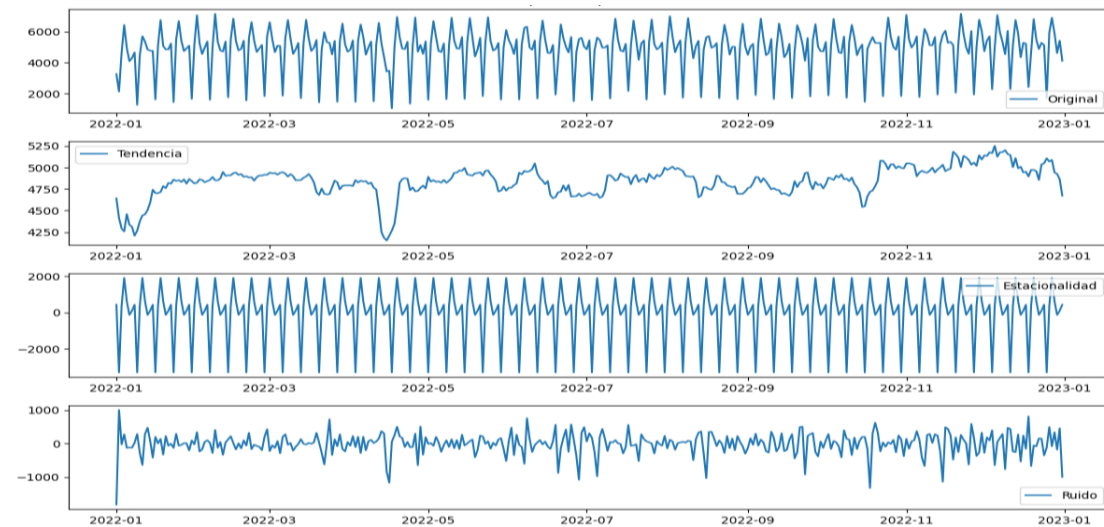


Figura 19

Descomposición de la Serie de Tiempo Base de Datos Bogotá 2022 (Toneladas Diarias de Residuos Domiciliarios)

**Figura 20**

Descomposición de la Serie de Tiempo Base de Datos Bogotá 2023 (Toneladas Diarias de Residuos Domiciliarios)

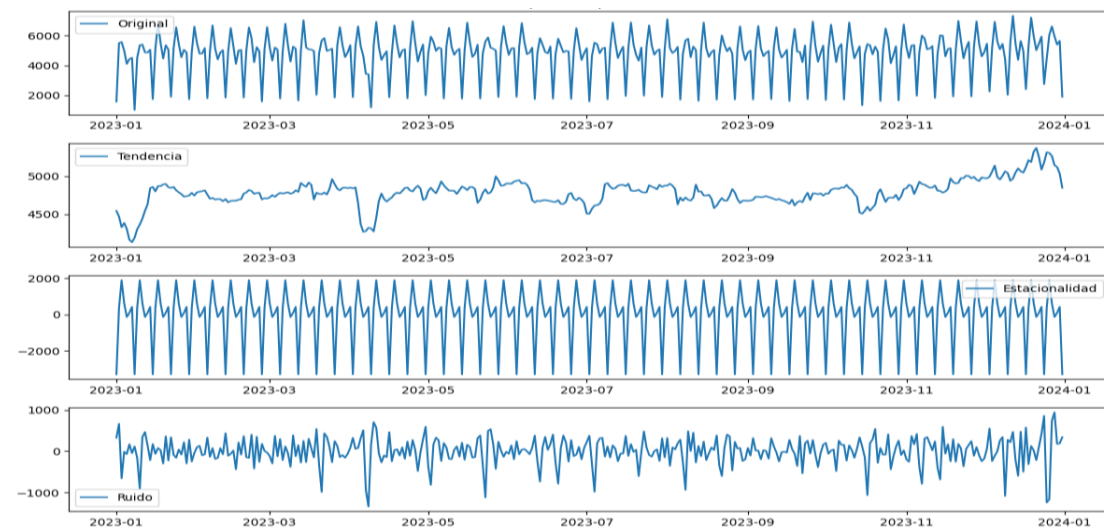
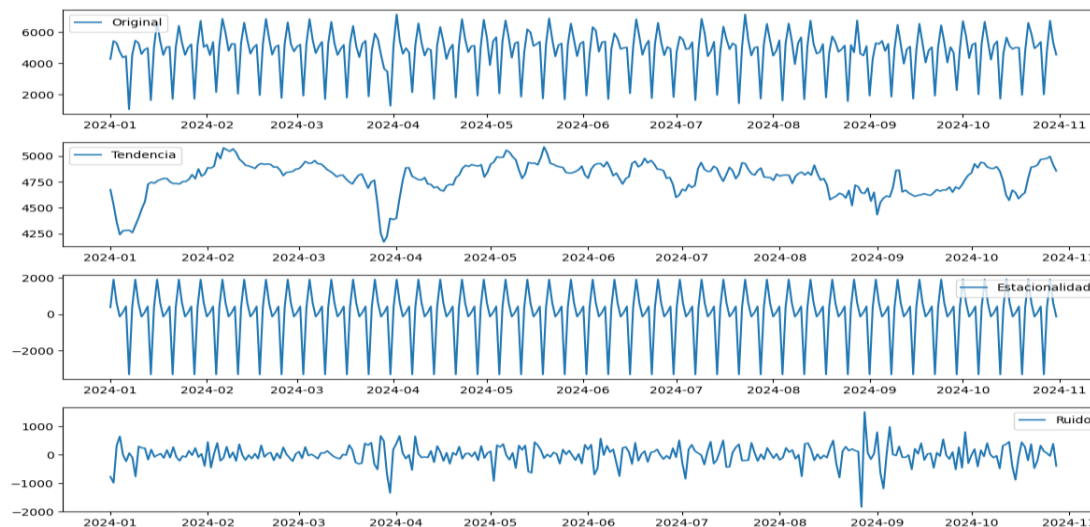


Figura 21

Descomposición de la Serie de Tiempo Base de Datos Bogotá 2024 (Toneladas Diarias de Residuos Domiciliarios)



El siguiente análisis se centra en la descomposición de la serie temporal del volumen diario de residuos domiciliarios en Bogotá (marzo de 2018 a octubre de 2024) en sus componentes fundamentales: Original, Tendencia, Estacionalidad y Ruido (Residual) observadas en las figuras 15 a la 21.

La serie original (entre 2,000 y 6,500 toneladas diarias) muestra una clara y fuerte Estacionalidad con subidas y bajadas diarias muy pronunciadas. El nivel base se mantiene relativamente estable, aunque se observa un ligero aumento hacia el final de cada año.

La Tendencia general es ligeramente positiva en el periodo analizado. A pesar de ser bastante plana y fluctuante durante gran parte del año (mayo a noviembre), se observan dinámicas específicas:

Un fuerte aumento inicial entre enero y febrero (de 4,250 a casi 5,000 toneladas).

Una estabilidad relativa a mediados de año (marzo a julio), con breves caídas notables.

Un claro repunte hacia el cierre del año (a partir de octubre), alcanzando los valores más altos (cerca de 5,250 toneladas), lo que indica que la serie concluye el año en un nivel superior al que comenzó.

La Estacionalidad es el factor más dominante y notable de la serie original. Presenta un patrón muy marcado y altamente consistente a lo largo de todo el año, con fluctuaciones regulares y de gran magnitud, oscilando consistentemente entre aproximadamente -2,000 y +2,000 toneladas. Esta extrema regularidad sugiere la presencia de un patrón repetitivo muy fuerte, probablemente de naturaleza diaria o semanal.

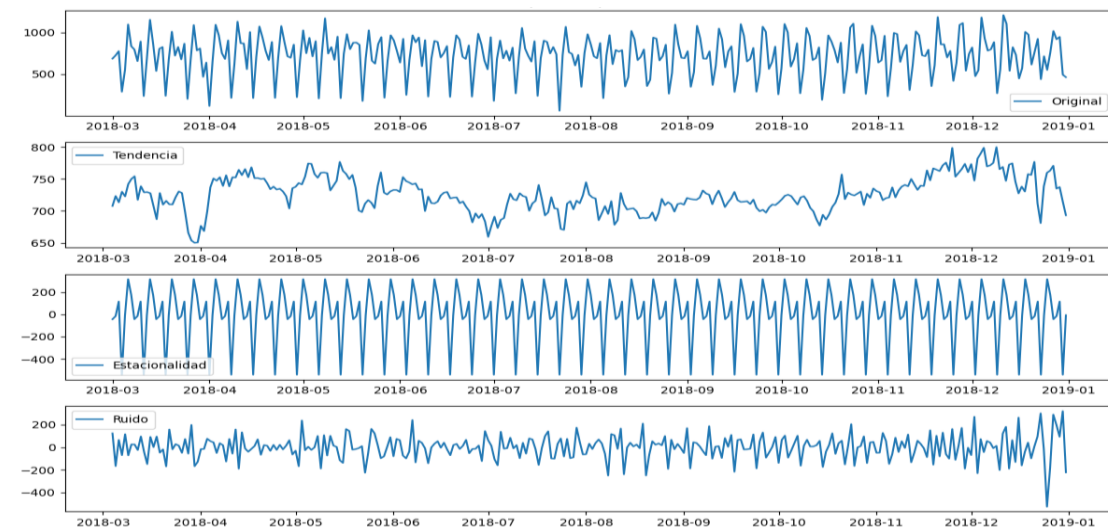
El Ruido se encuentra generalmente centrado en cero y tiene una amplitud mucho menor que la estacionalidad (mayormente entre -1,000 y 1,000 toneladas). Esto indica que la descomposición ha capturado la mayoría de las dinámicas importantes. Sin embargo, la presencia de algunos picos residuales (ej. a principios y finales del año) sugiere que ciertos eventos no explicados por la Tendencia o la Estacionalidad aún influyen en la serie.

Base de Datos Suba (Toneladas Diarias de Residuos Domiciliarios). Para la descomposición de la serie de tiempo usamos Python en el cual se realiza un análisis exploratorio y una descomposición de series de tiempo sobre un archivo CSV que contiene las toneladas diarias de residuos domiciliarios llevadas por día al relleno sanitario Parque de innovación Doña Juana desde el 1 de marzo del 2018 al 31 de octubre de 2024 de la localidad de Suba.

En Python cargamos los datos de toneladas a lo largo del tiempo, se limpian, se analizan descriptivamente y, finalmente, se visualizan sus componentes de tendencia, estacionalidad y residuos de manera anual.

Figura 22

Descomposición de la Serie de Tiempo Base de Datos Suba 2018 (Toneladas Diarias de Residuos Domiciliarios)

**Figura 23**

Descomposición de la Serie de Tiempo Base de Datos Suba 2019 (Toneladas Diarias de Residuos Domiciliarios)

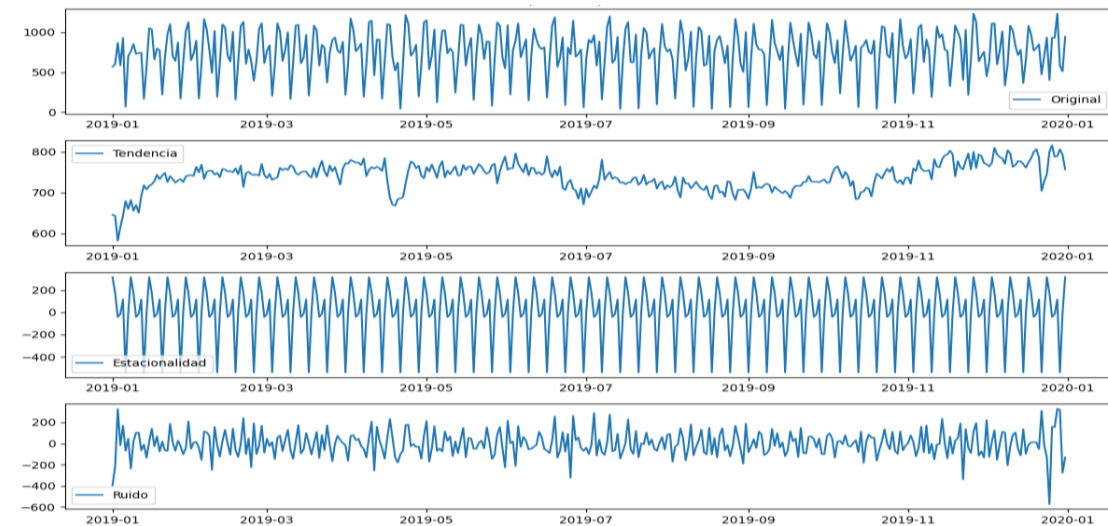
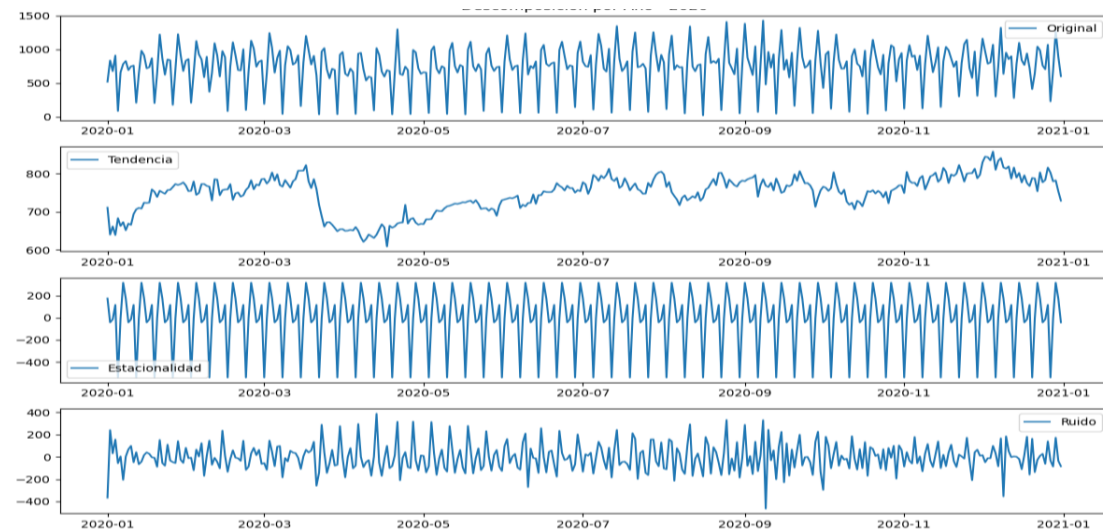


Figura 24

Descomposición de la Serie de Tiempo Base de Datos Suba 2020 (Toneladas Diarias de Residuos Domiciliarios)

**Figura 25**

Descomposición de la Serie de Tiempo Base de Datos Suba 2021 (Toneladas Diarias de Residuos Domiciliarios)

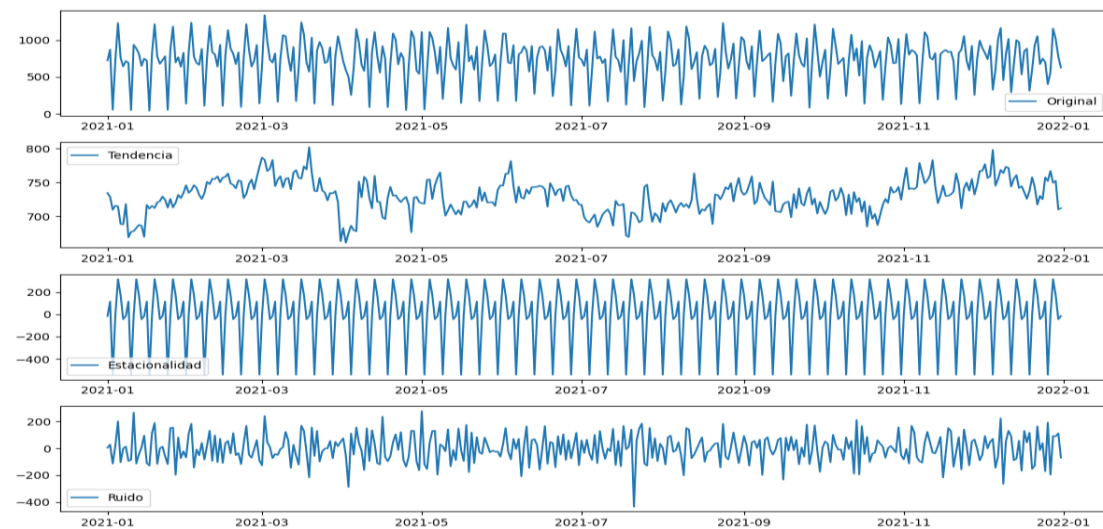
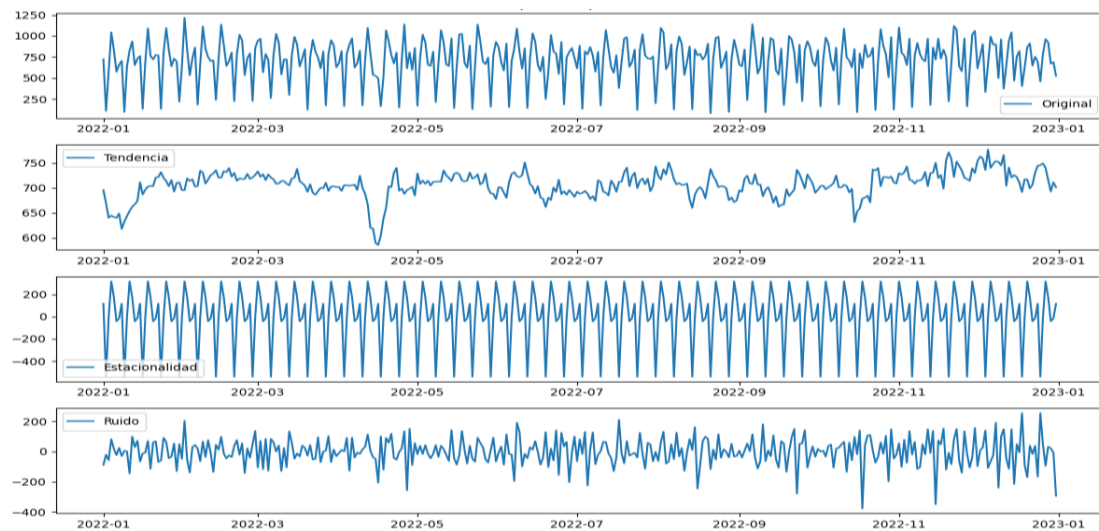


Figura 26

Descomposición de la Serie de Tiempo Base de Datos Suba 2022 (Toneladas Diarias de Residuos Domiciliarios)

**Figura 27**

Descomposición de la Serie de Tiempo Base de Datos Suba 2023 (Toneladas Diarias de Residuos Domiciliarios)

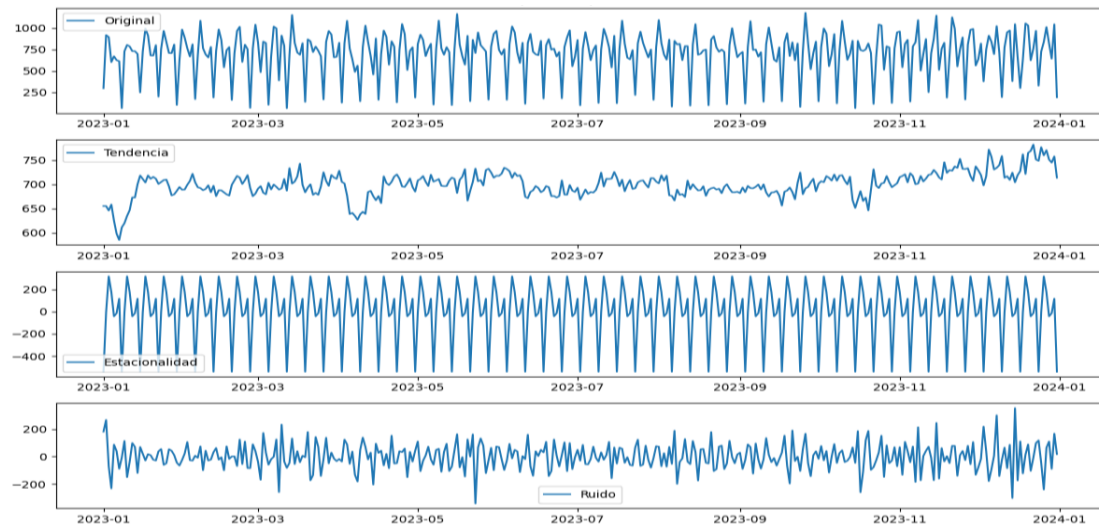
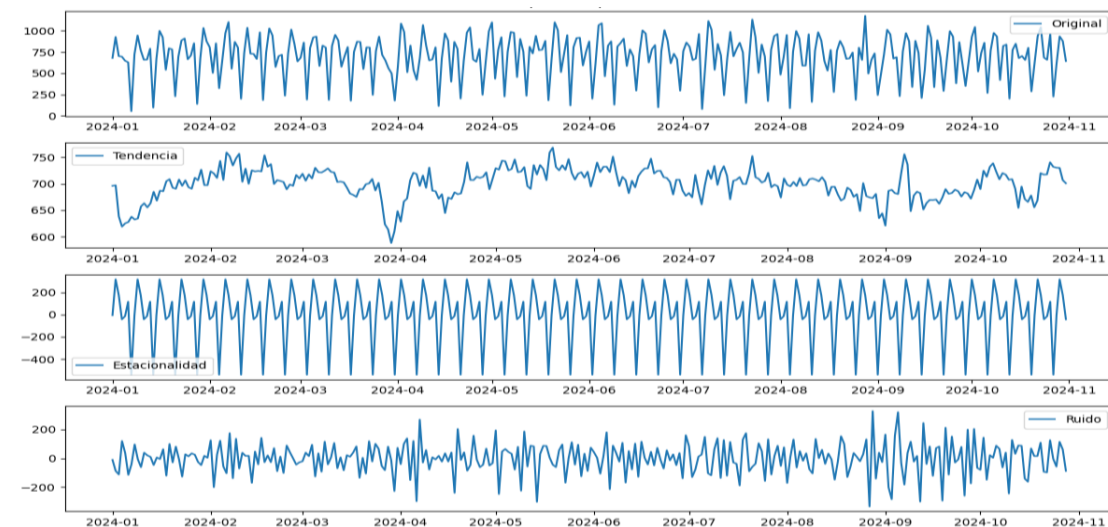


Figura 28

Descomposición de la Serie de Tiempo Base de Datos Suba 2024 (Toneladas Diarias de Residuos Domiciliarios)



El siguiente análisis se enfoca en la descomposición de la serie temporal del volumen diario de residuos domiciliarios de la localidad de Suba, cubriendo el periodo del 1 de marzo de 2018 al 31 de octubre de 2024. Mediante la aplicación de un modelo de descomposición, la serie original se divide en sus componentes fundamentales: Tendencia, Estacionalidad y Ruido (Residual) observadas en las figuras 22 a la 28, permitiendo interpretar el comportamiento del volumen de residuos de la localidad de Suba, una de las mayores contribuyentes en Bogotá.

La serie original de Suba presenta una alta y constante volatilidad a lo largo de los años, con picos y valles muy marcados que se repiten, lo que evidencia una fuerte influencia cíclica; los valores diarios oscilan entre aproximadamente 250 y 1,100 toneladas. En cuanto a la Tendencia, que representa el movimiento a largo plazo, esta mostró un patrón ascendente general a lo largo del periodo. Se caracteriza por un fuerte crecimiento inicial en enero, elevando el nivel base de 600 a cerca de 700 toneladas.

Después de este crecimiento inicial, la tendencia mantuvo una estabilidad relativa durante el período central del año (entre febrero y octubre), fluctuando en un rango estrecho de 675 a 725 toneladas diarias. Sin embargo, hacia el final de cada año, a partir de noviembre, se observa un claro y sostenido repunte y crecimiento, donde la tendencia alcanza los valores más altos, acercándose a las 775 toneladas. En resumen, a pesar de la estabilidad intermedia, la trayectoria general de la generación de residuos de Suba ha sido ligeramente positiva, terminando el año en un nivel superior al que inició.

La componente de Estacionalidad es el factor más marcado y regular de la serie de Suba, con un patrón de onda altamente consistente a lo largo de los doce meses. Sus fluctuaciones son de gran magnitud, oscilando significativamente entre aproximadamente -500 y +300 toneladas. La extrema regularidad de estas crestas y valles a intervalos fijos subraya que la actividad subyacente de la generación y recolección de residuos está fuertemente dictada por un patrón fijo, muy probablemente un ciclo diario o semanal bien definido.

Finalmente, el componente de Ruido (residual) se distribuye de manera irregular y aleatoria alrededor de cero, moviéndose generalmente en el rango de -200 a +300 toneladas, lo que indica que la mayor parte de la dinámica de la serie ha sido capturada por la Tendencia y la Estacionalidad. No obstante, la presencia de picos significativos en el ruido (como el observado notablemente al inicio de diciembre) sugiere la ocurrencia de eventos puntuales o atípicos que impactaron la serie de tiempo de forma no recurrente, y que no son explicados por los patrones cíclicos o de largo plazo.

Base de Datos Kennedy (Toneladas Diarias de Residuos Domiciliarios). Para la descomposición de la serie de tiempo usamos Python en el cual se realiza un análisis exploratorio y una descomposición de series de tiempo sobre un archivo CSV que contiene las

toneladas diarias de residuos domiciliarios llevadas por día al relleno sanitario Parque de innovación Doña Juana desde el 1 de marzo del 2018 al 31 de octubre de 2024 de la localidad de Kennedy.

En Python cargamos los datos de toneladas a lo largo del tiempo, se limpian, se analizan descriptivamente y, finalmente, se visualizan sus componentes de tendencia, estacionalidad y residuos de manera anual.

Figura 29

Descomposición de la Serie de Tiempo Base de Datos Kennedy 2018 (Toneladas Diarias de Residuos Domiciliarios)

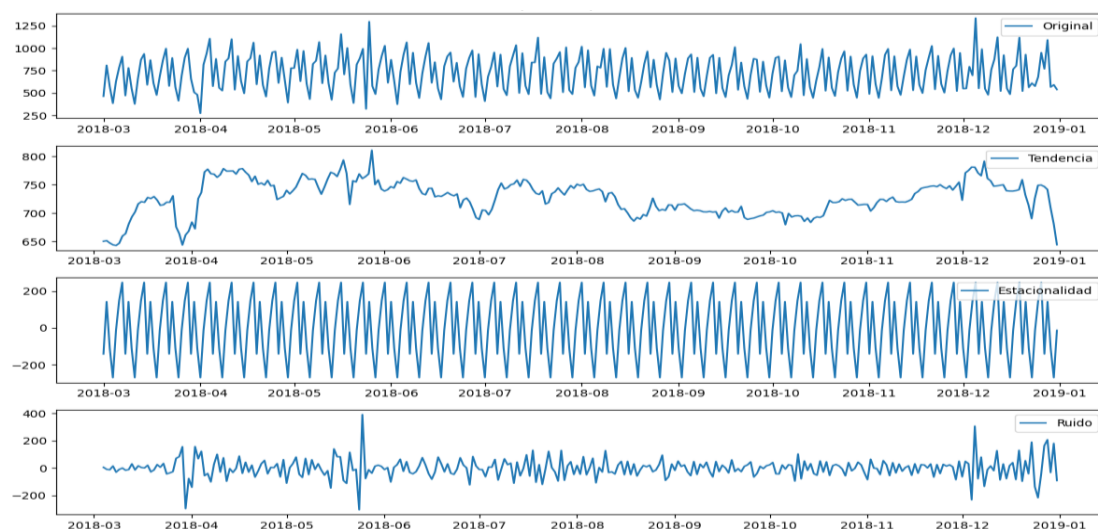
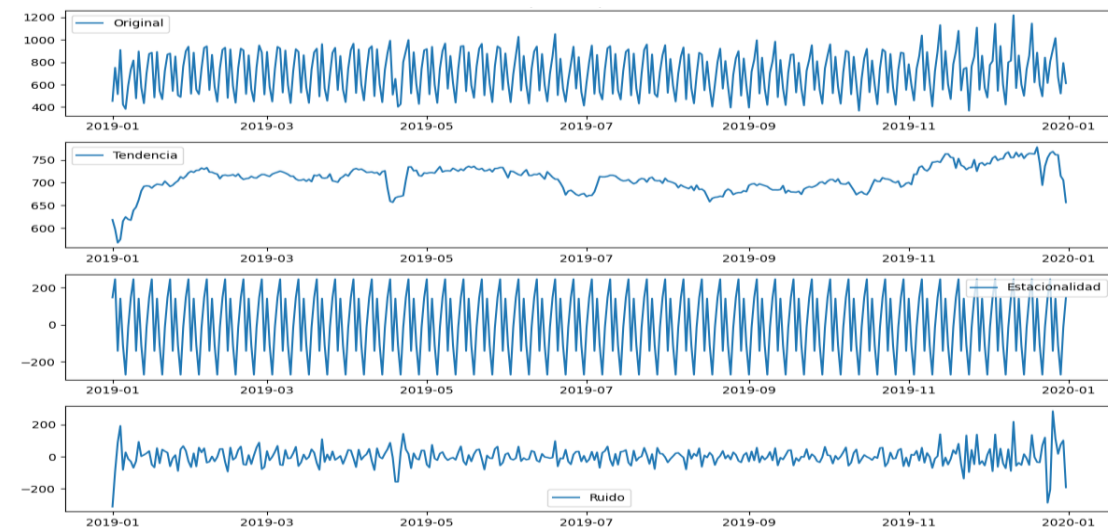


Figura 30

Descomposición de la Serie de Tiempo Base de Datos Kennedy 2019 (Toneladas Diarias de Residuos Domiciliarios)

**Figura 31**

Descomposición de la Serie de Tiempo Base de Datos Kennedy 2020 (Toneladas Diarias de Residuos Domiciliarios)

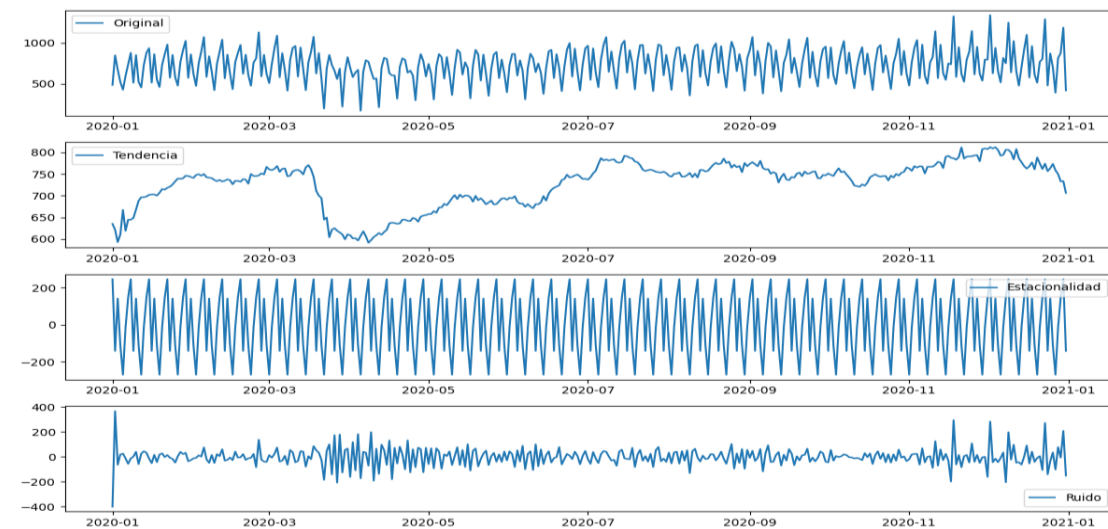
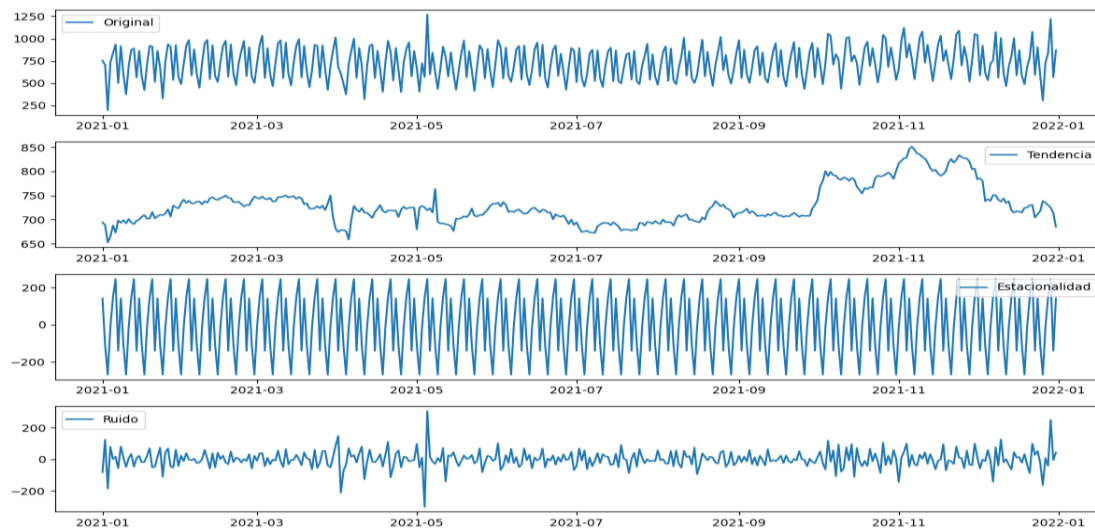


Figura 32

Descomposición de la Serie de Tiempo Base de Datos Kennedy 2021 (Toneladas Diarias de Residuos Domiciliarios)

**Figura 33**

Descomposición de la Serie de Tiempo Base de Datos Kennedy 2022 (Toneladas Diarias de Residuos Domiciliarios)

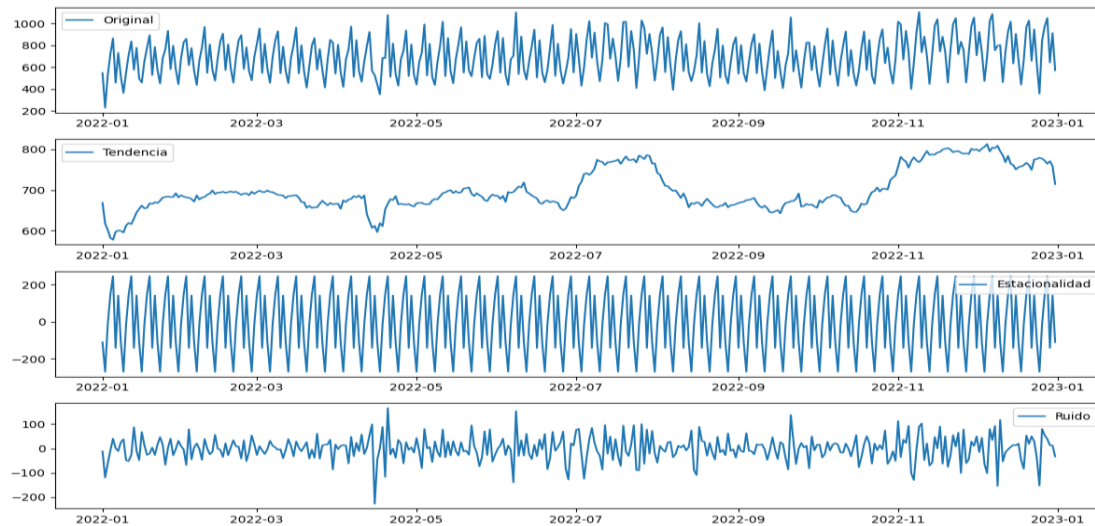
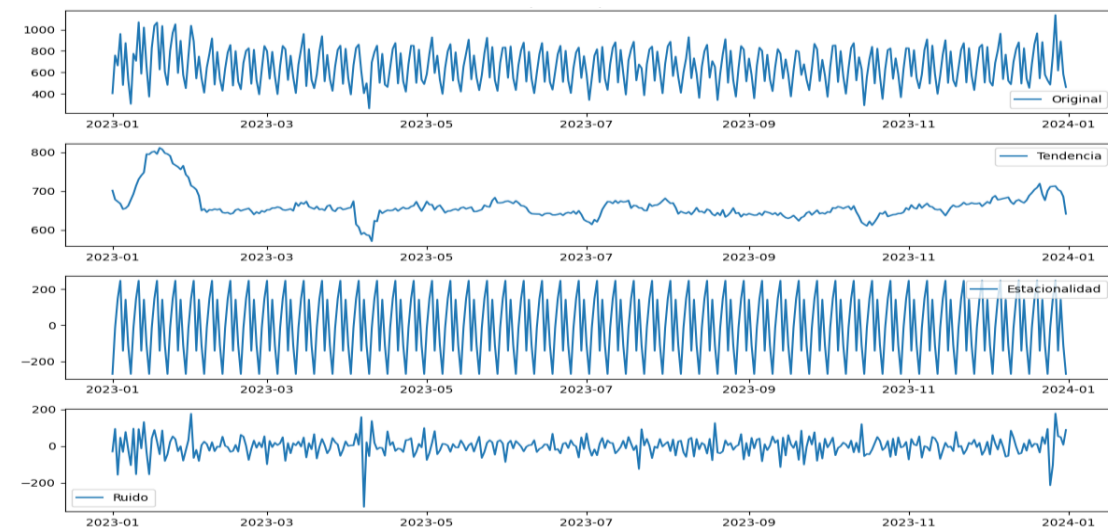
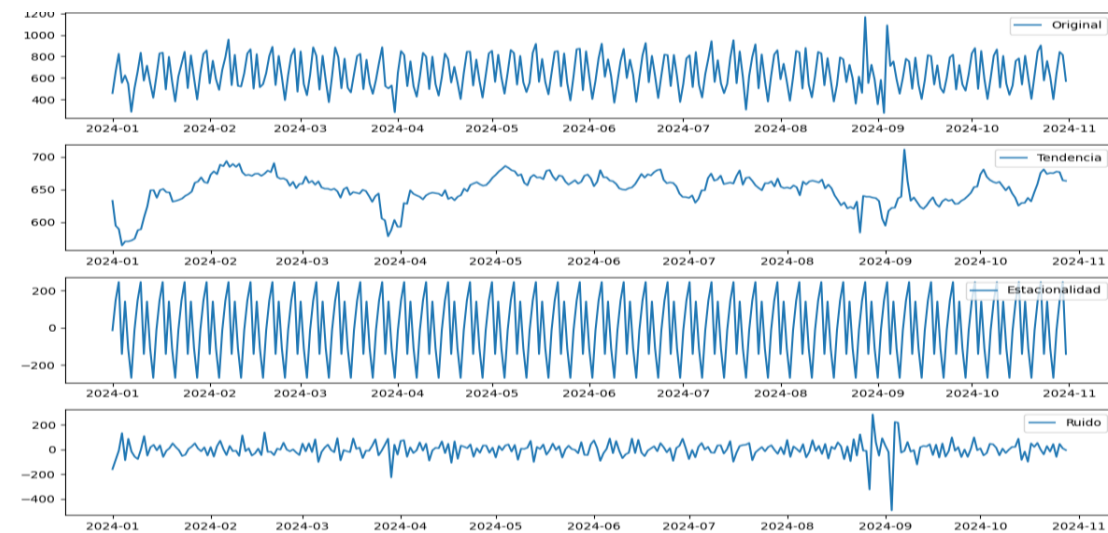


Figura 34

Descomposición de la Serie de Tiempo Base de Datos Kennedy 2023 (Toneladas Diarias de Residuos Domiciliarios)

**Figura 35**

Descomposición de la Serie de Tiempo Base de Datos Kennedy 2024 (Toneladas Diarias de Residuos Domiciliarios)



El análisis de la serie temporal del volumen diario de residuos domiciliarios de la localidad de Kennedy (marzo de 2018 a octubre de 2024), una de las mayores contribuyentes, se descompuso en sus componentes fundamentales: Original, Tendencia, Estacionalidad y Ruido observadas en las figuras 29 a la 35.

La Serie Original presenta una alta volatilidad constante con picos y valles muy regulares, lo que confirma un patrón recurrente fuerte. Los valores diarios oscilan consistentemente entre 300 y 1,100 toneladas. La Estacionalidad es el factor más notable: es muy marcada y constante en el periodo, con fluctuaciones que oscilan fuertemente entre aproximadamente -250 y +250 toneladas. Esta gran uniformidad cíclica indica que la variabilidad total está sustancialmente dominada por un patrón fijo, probablemente de naturaleza diaria o semanal.

La Tendencia de la generación de residuos en Kennedy es netamente positiva en el balance general, culminando el periodo en un nivel base significativamente más alto que su inicio. El patrón a largo plazo se caracteriza por un rápido aumento inicial hasta estabilizarse cerca de 680-700 toneladas hasta mediados de julio. Posteriormente, se produce un fuerte incremento hasta cerca de 780 toneladas, seguido de un descenso abrupto a finales de agosto, para finalmente mostrar un claro y sostenido repunte a partir de octubre, alcanzando el valor máximo cercano a 800 toneladas en diciembre.

El componente de Ruido (variabilidad no explicada) se distribuye de forma irregular y aleatoria alrededor de cero, con la mayoría de los valores concentrados entre -100 y +100 toneladas. La baja magnitud general del ruido sugiere la efectividad del modelo de descomposición. Sin embargo, se destaca la presencia de un pico negativo pronunciado cerca de

abril, que señala un evento atípico o puntual y no recurrente que no fue capturado por la tendencia o la estacionalidad.

Resultado del Modelado

El siguiente apartado detalla el rigor metodológico aplicado en el proceso de modelado, el cual incluye la selección óptima de parámetros y la evaluación del rendimiento de las técnicas SARIMAX y Prophet sobre las tres bases de datos seleccionadas en Python. Asimismo, se exponen los resultados obtenidos y la predicción proyectada para cada base de datos en el período del 1 de noviembre de 2024 al 31 de diciembre de 2029.

Modelo SARIMAX. El proceso de modelado para la serie de tiempo TONELADAS comenzó con un Análisis Exploratorio de Autocorrelación, donde se graficaron la Función de Autocorrelación (ACF) y la Función de Autocorrelación Parcial (PACF). El objetivo de estas gráficas fue visualizar la dependencia temporal de la serie y obtener una estimación inicial de los órdenes (p, q) del modelo. Posteriormente, se llevó a cabo la Preparación de Datos, dividiendo el conjunto en un Conjunto de Entrenamiento (80%) para la construcción del modelo y un Conjunto de Prueba (20%) para asegurar una validación objetiva en datos no vistos.

La etapa central fue la Optimización y Búsqueda de Parámetros del modelo SARIMAX. Se implementó una búsqueda en cuadrícula (grid search) para probar sistemáticamente todas las combinaciones de órdenes no estacionales (p, d, q) y estacionales (P, D, Q, s=7). El Criterio de Información de Akaike (AIC) fue el criterio de selección, buscando la combinación que lo minimizara para garantizar el modelo más parsimonioso. Una vez identificados los parámetros óptimos, se procedió al Entrenamiento Final del Modelo, ajustando la versión definitiva con estos hiperparámetros y el conjunto de entrenamiento.

A continuación, se realizó la Evaluación del Rendimiento. Primero, se ejecutó una Evaluación Visual del Pronóstico al calcular las predicciones puntuales e intervalos de confianza sobre la longitud del conjunto de prueba. La gráfica resultante permitió una inspección inmediata de la precisión predictiva. Paralelamente, se procedió con una Evaluación Cuantitativa comparando los pronósticos con los valores reales del conjunto de prueba. Se calcularon cuatro métricas clave: el Error Absoluto Medio (MAE), la Raíz del Error Cuadrático Medio (RMSE), el R Cuadrado (R^2), y el Error Porcentual Absoluto Medio (MAPE), proporcionando una medida objetiva y numérica del rendimiento.

Finalmente, se utilizó el modelo SARIMAX óptimo previamente entrenado para generar un Pronóstico a Futuro de la serie de tiempo TONELADAS en un rango específico (del 1 de noviembre de 2024 al 31 de diciembre de 2029). Se obtuvo la media de las predicciones junto con sus intervalos de confianza, y se visualizó la proyección. Esta gráfica final combina toda la serie histórica con la proyección futura y su banda de incertidumbre, facilitando la toma de decisiones a largo plazo.

Modelo Prophet. Para implementar el modelo Prophet la primera fase se enfoca en preparar los datos para el modelado y la validación. Se comienza definiendo la proporción de los datos a usar para la prueba (20%) y calculando dinámicamente el número de días que esto representa. Luego, el conjunto de datos completo se divide en dos secciones cronológicamente ordenadas: el Conjunto de Entrenamiento (80%) (df_{train}), usado para ajustar el modelo, y el Conjunto de Prueba (20%) (df_{test}), reservado para evaluar su rendimiento sin sesgos.

Posteriormente usamos el método Grid Search, el cual es un proceso riguroso para encontrar la mejor configuración del modelo Prophet. Se define una cuadrícula de hiperparámetros clave, incluyendo: `changepoint_prior_scale` (que controla la flexibilidad de la

tendencia), `seasonality_prior_scale` (flexibilidad de la estacionalidad) y `seasonality_mode` (aditivo o multiplicativo). El código itera sobre todas las combinaciones de estos parámetros, entrena un modelo diferente con cada combinación y lo evalúa calculando el RMSE (Raíz del Error Cuadrático Medio) sobre el conjunto de prueba. Finalmente, se seleccionan los parámetros óptimos que resultaron en el RMSE más bajo.

Una vez identificados los mejores parámetros, se ajusta el Modelo Prophet Final utilizando la totalidad del Conjunto de Entrenamiento. Este modelo optimizado se utiliza inmediatamente para generar un pronóstico para el período que abarca el Conjunto de Prueba. Se crea una gráfica que compara visualmente los Valores Reales con los Valores Pronosticados, incluyendo el intervalo de confianza, para mostrar qué tan bien se ajusta el modelo a los datos no vistos.

Posteriormente evaluamos el rendimiento del modelo, para ello se calculó e imprimió las métricas de error clave sobre el período de prueba: MAE (Error Absoluto Medio), RMSE, R^2 (Coeficiente de Determinación) y MAPE (Error Porcentual Absoluto Medio), proporcionando una medida numérica de la precisión del modelo. Finalmente, el modelo optimizado se utiliza para generar un pronóstico a largo plazo desde el 1 de noviembre de 2024 hasta el 31 de diciembre de 2029, y se visualiza la serie histórica junto con esta proyección futura en un gráfico interactivo.

Base de Datos Bogotá (Toneladas Diarias de Residuos Domiciliarios).

Resultado Modelo SARIMAX. A continuación, observamos el resultado de salida del proceso de modelado de series de tiempo en Python, específicamente utilizando la función SARIMAX de la librería statsmodels. Muestra en dos partes: el resultado del proceso de optimización de parámetros y el resumen estadístico del modelo final ajustado.

Figura 36

Resultados Estadísticos del Modelo SARIMAX Base de Datos Bogotá (Toneladas Diarias de Residuos Domiciliarios)

```

Searching for optimal SARIMA parameters...
Best SARIMA Order: (1, 1, 1)
Best SARIMA Seasonal Order: (1, 1, 1, 7)
Best AIC: 28659.42

Training the SARIMA model with the best parameters on training data...
SARIMAX Results
=====
Dep. Variable:          TONELADAS      No. Observations:      1949
Model:                 SARIMAX(1, 1, 1)x(1, 1, 1, 7)  Log Likelihood         -14324.710
Date:                  Fri, 07 Nov 2025      AIC                    28659.420
Time:                  15:07:44                BIC                    28687.251
Sample:                03-01-2018                HQIC                   28669.657
                    - 07-01-2023

Covariance Type:      opg
=====
              coef    std err          z      P>|z|      [0.025    0.975]
-----
ar.L1         0.1695     0.016     10.458     0.000     0.138     0.201
ma.L1        -0.9467     0.008    -120.527     0.000    -0.962    -0.931
ar.S.L7       0.2175     0.016     13.411     0.000     0.186     0.249
ma.S.L7      -0.9342     0.011    -84.952     0.000    -0.956    -0.913
sigma2       1.603e+05    2637.446     60.794     0.000    1.55e+05    1.66e+05
=====
Ljung-Box (L1) (Q):      0.63   Jarque-Bera (JB):      5494.90
Prob(Q):                 0.43   Prob(JB):              0.00
Heteroskedasticity (H): 0.66   Skew:                  -0.91
Prob(H) (two-sided):    0.00   Kurtosis:              11.06
=====

Warnings:
[1] Covariance matrix calculated using the outer product of gradients (complex-step).

```

Esta optimización culminó con la selección del modelo SARIMAX (1, 1, 1) x (1, 1, 1, 7), que incluyó un término Autorregresivo (AR), una diferenciación (I) y un término de Media Móvil (MA) tanto en la componente no estacional (p=1, d=1, q=1) como en la estacional (P=1, D=1, Q=1). Esta combinación fue elegida por obtener el valor más bajo del Criterio de Información de Akaike (AIC), fijado en 28659.42, lo que garantiza el mejor equilibrio entre ajuste y complejidad del modelo.

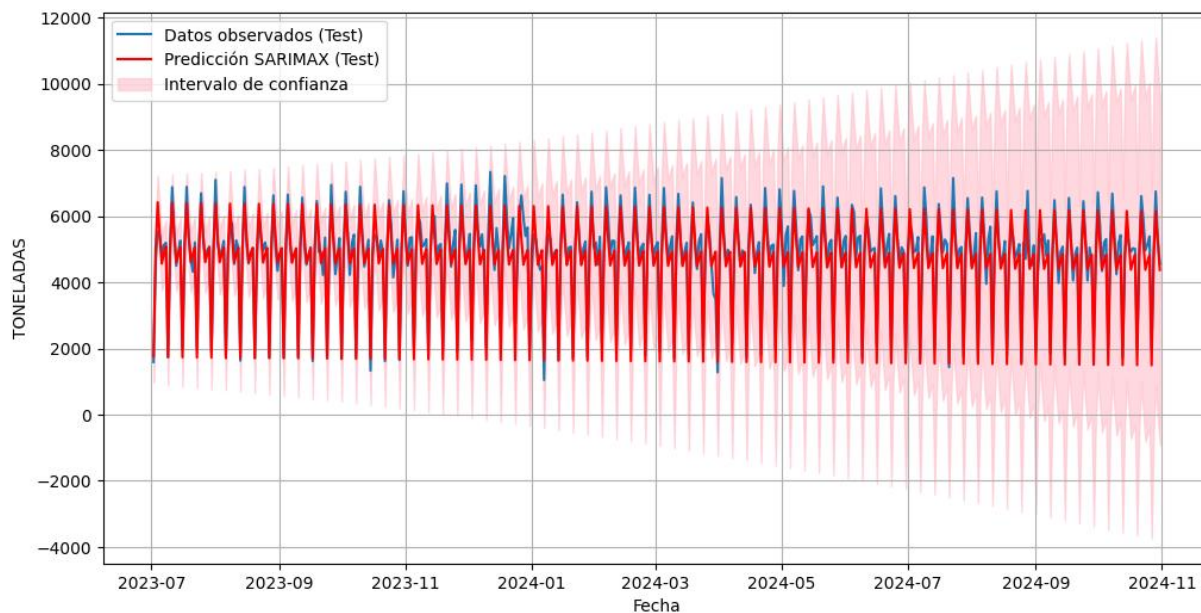
Una vez ajustado el modelo óptimo a las 1949 observaciones del conjunto de entrenamiento, el resumen estadístico confirmó que todos los coeficientes estimados son estadísticamente significativos (P-valores de 0.000). Esto indica que cada componente del modelo seleccionado contribuye de manera relevante y no aleatoria a la explicación de la varianza en la serie de TONELADAS. Los coeficientes muestran una fuerte dependencia de la media móvil (con valores cercanos a -0.95) y una dependencia estacional positiva con un rezago de siete días.

En resumen, el proceso de modelado identificó un modelo SARIMAX potente y parsimonioso capaz de capturar tanto la estructura no estacional como la estacional semanal ($s=7$) de la serie TONELADAS. Aunque el modelo es exitoso en la captura de la dependencia temporal (autocorrelación), las advertencias en los diagnósticos de residuos, específicamente la falta de normalidad y la heterocedasticidad, implican que los intervalos de confianza futuros podrían ser menos precisos de lo ideal, sugiriendo la necesidad de posibles transformaciones de datos o modelos más complejos para manejar la varianza variable.

Evaluación del Rendimiento del Modelo. En la siguiente imagen se muestra una evaluación visual del rendimiento de un modelo SARIMAX sobre el período de prueba (test data). El análisis se centra en comparar las predicciones del modelo con los valores reales y observar la evolución de la incertidumbre.

Figura 37

Predicción del Modelo SARIMAX sobre el Periodo de Prueba Base de Datos Bogotá (Toneladas Diarias de Residuos Domiciliarios)



El modelo SARIMAX demuestra un desempeño inicial adecuado al capturar con éxito el patrón estacional de la serie TONELADAS, como lo indica la línea de predicción roja que sigue de cerca los datos observados al comienzo del período de prueba. Sin embargo, a medida que el horizonte de predicción se extiende hacia finales de 2024, la precisión puntual del pronóstico disminuye, con la media predicha volviéndose menos reactiva a las fluctuaciones reales, lo cual es un comportamiento común en series de tiempo, indicando una degradación de la capacidad predictiva a largo plazo.

El aspecto más crítico del análisis visual es la rápida y dramática divergencia del Intervalo de Confianza (el "cono de incertidumbre" en rosa). Este ensanchamiento explosivo, que al final del período de prueba abarca un rango desde valores negativos hasta cifras muy altas, indica una altísima incertidumbre y una inestabilidad en la varianza (heterocedasticidad) del

modelo. Esta falta de confianza a largo plazo, visualmente confirmada, hace que el modelo SARIMAX, a pesar de su buen ajuste de corto plazo, sea poco fiable para la toma de decisiones futuras debido al riesgo asociado a un pronóstico que incluye valores no plausibles (negativos) para la variable TONELADAS.

En cuanto a las métricas de precisión que cuantifican el rendimiento del modelo SARIMAX óptimo al pronosticar los valores de TONELADAS en el período de prueba (test_data) se obtuvieron los siguientes resultados:

MAE (Error Absoluto Medio): 317.73

RMSE (Raíz del Error Cuadrático Medio): 402.84

R² (R Cuadrado): 0.9147

MAPE (Error Porcentual Absoluto Medio): 7.18%

El modelo SARIMAX mostró un alto poder explicativo, con un R² de 0.9147, lo que indica que logró capturar y explicar más del 91% de la variabilidad de la serie de tiempo TONELADAS, confirmando la idoneidad de la estructura temporal seleccionada. Sin embargo, a pesar de este excelente ajuste, el análisis cuantitativo revela una limitación clave: aunque el error promedio absoluto (MAE) fue de 317.73 TONELADAS, la Raíz del Error Cuadrático Medio (RMSE) de 402.84 TONELADAS fue significativamente mayor. Esta disparidad advierte sobre la presencia de errores puntuales sustancialmente grandes (outliers de predicción), lo que concuerda con la alta incertidumbre visualizada, limitando la fiabilidad del pronóstico en escenarios donde la precisión debe ser crítica en cada punto.

Resultado Modelo Prophet. El proceso de optimización del modelo Prophet buscó identificar la combinación de parámetros que generara el menor error de predicción. El resultado exitoso de este ajuste fue un RMSE (Raíz del Error Cuadrático Medio) mínimo de 377.38. Esto

significa que, con la configuración encontrada, el error promedio de la predicción puntual del modelo es de aproximadamente 377.38 toneladas. Este valor representa la máxima precisión lograda para pronosticar los residuos domiciliarios de Bogotá con este modelo.

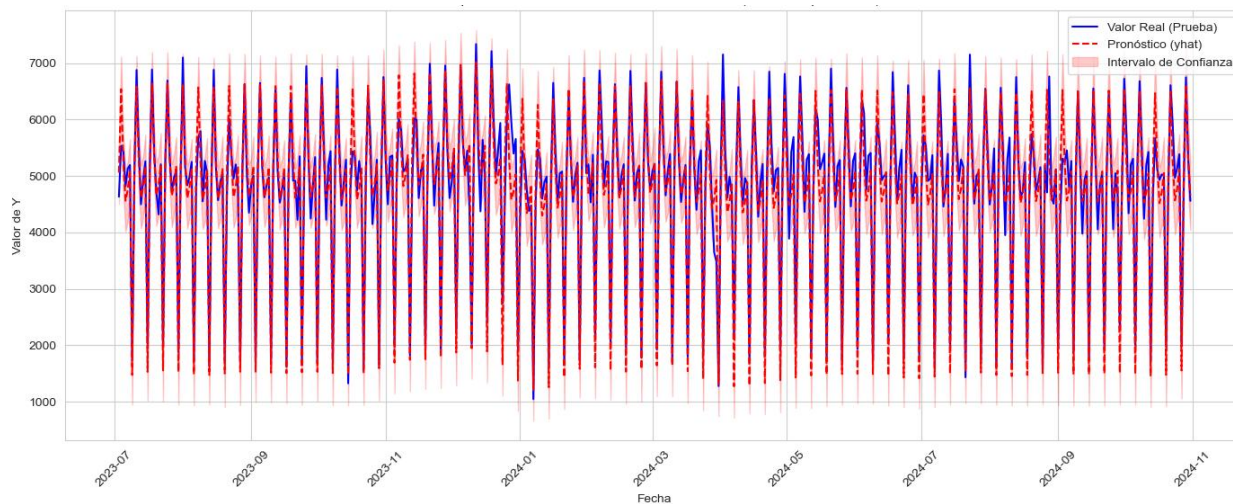
Los tres hiperparámetros que arrojaron el mejor rendimiento fueron cruciales para capturar la dinámica de la serie. Primero, el valor de 'changeoint_prior_scale' de 0.1 indica que la serie de residuos ha experimentado múltiples cambios en su ritmo de crecimiento, requiriendo una tendencia más flexible que la predeterminada. Segundo, el modo de estacionalidad fue definido como 'multiplicative', lo cual confirma que la magnitud de los picos y valles estacionales aumenta a medida que el volumen total de residuos crece a lo largo del tiempo.

La combinación ganadora de parámetros, incluyendo un 'seasonality_prior_scale' de 5.0 para controlar la fuerza de los patrones estacionales, permitió al modelo Prophet ajustarse de manera óptima tanto a la tendencia de crecimiento como a la fuerte estacionalidad. En resumen, el modelo ajustado es sensible a los cambios históricos en el crecimiento (tendencia flexible) y reconoce que la volatilidad estacional (semanal/anual) es proporcional al volumen total de residuos (modo multiplicativo), resultando en el pronóstico más preciso con el RMSE de 377.38.

Evaluación del Rendimiento del Modelo. En la siguiente imagen se muestra la evaluación visual del modelo Prophet optimizado en el período de prueba, comparando sus pronósticos con los valores reales:

Figura 38

Predicción del Modelo Prophet sobre el Periodo de Prueba Base de Datos Bogotá (Toneladas Diarias de Residuos Domiciliarios)



El modelo logró un excelente ajuste al patrón estacional y mantuvo una incertidumbre relativamente estable en comparación con el modelo SARIMAX anterior.

Esta imagen muestra una evaluación visual del Modelo Prophet optimizado en su capacidad para pronosticar los Valores de (Y) durante el período de prueba. La línea de pronóstico (yhat, discontinua roja) demuestra un excelente ajuste al patrón de la serie de tiempo real (línea azul) a lo largo de todo el período, desde julio de 2023 hasta finales de 2024. El modelo ha capturado con alta fidelidad las fluctuaciones de alta frecuencia (picos y valles), lo que indica que el proceso de optimización mediante Grid Search fue exitoso para identificar la estacionalidad y la tendencia correctas, lo que se traduce en un modelo con un fuerte poder explicativo.

Un aspecto crucial de la gráfica es el comportamiento del Intervalo de Confianza (área sombreada en rosa), el cual permanece estable y de ancho constante a lo largo de todo el período

de prueba. Esta estabilidad es una señal poderosa de que el modelo Prophet ha manejado efectivamente la varianza de los errores, lo que proporciona alta fiabilidad a sus predicciones de medio a largo plazo, ya que la incertidumbre no se degrada exponencialmente con el tiempo. El pronóstico es robusto, y la mayoría de los valores reales caen dentro de este rango de confianza, validando al modelo Prophet como una herramienta de pronóstico estable para esta serie de tiempo.

En cuanto a las métricas de evaluación del rendimiento del modelo Prophet optimizado se obtuvieron los siguientes resultados:

Error Absoluto Medio (MAE): 282.88

Raíz del Error Cuadrático Medio (RMSE): 377.38

Coefficiente de Determinación (R^2): 0.9245

Error Porcentual Absoluto Medio (MAPE): 6.93%

El modelo de pronóstico Prophet demostró un rendimiento de muy alta precisión durante el período de prueba. El Error Absoluto Medio (MAE) fue de 282.88 toneladas, lo que indica que la desviación promedio de las predicciones respecto a los valores reales es baja. Este alto nivel de precisión se valida con el Coeficiente de Determinación (R^2) de 0.9245, un valor que confirma el poder explicativo excelente del modelo al capturar el 92.45% de la variabilidad total de la serie de tiempo. Además, el Error Porcentual Absoluto Medio (MAPE) fue de solo 6.93%, lo cual se considera un pronóstico de muy alta calidad y sumamente útil. La única advertencia se observa en la diferencia entre el RMSE (377.38 toneladas) y el MAE, lo que sugiere que, a pesar de la precisión general, el modelo incurre en la presencia de algunos errores de predicción considerablemente grandes de forma ocasional, un punto que debe monitorearse.

Evaluación y Comparación de los Modelos. Con el objetivo de lograr las predicciones más acertadas, se efectuó una evaluación exhaustiva y una comparación de rendimiento entre los modelos de pronóstico. A continuación, se detalla el análisis que permitió la selección del modelo correcto para esta aplicación:

Tabla 1

Evaluación y Comparación del Rendimiento del Modelo MAE Y RMSE Base de Datos Bogotá

Métrica	SARIMAX	Prophet	Conclusión
MAE	317.73	282.88	Prophet reduce el error promedio en 34.85 unidades.
RMSE	402.84	377.38	Prophet reduce la penalización por errores grandes en 25.46 unidades.

Nota. Contiene los valores de las métricas MAE y RMSE de la evaluación del rendimiento de los modelos para la base de datos de la ciudad de Bogotá.

El Modelo Prophet muestra errores de predicción promedio (MAE) y errores penalizados (RMSE) consistentemente más bajos. Esto significa que las predicciones puntuales de Prophet están más cerca de los valores reales y que es menos propenso a cometer grandes errores ocasionales que el modelo SARIMAX.

Tabla 2

Evaluación y Comparación del Rendimiento del Modelo R² Y MAPE Base de Datos Bogotá

Métrica	SARIMAX	Prophet	Conclusión
R ²	0.9147	0.9245	Prophet explica casi un 1% más de la variabilidad.
MAPE	7.18%	6.93%	Prophet reduce el error porcentual promedio en 0.25 puntos.

Nota. Contiene los valores de las métricas R² y MAPE de la evaluación del rendimiento de los modelos para la base de datos de la ciudad de Bogotá.

El Modelo Prophet exhibe un mayor Coeficiente de Determinación (R²), lo que indica que es capaz de explicar el 92.45% de la variabilidad de la serie, superando la capacidad explicativa del SARIMAX. Además, al tener un MAPE más bajo, el modelo Prophet comete un error promedio porcentual menor, consolidándose como la opción de mayor precisión y utilidad práctica.

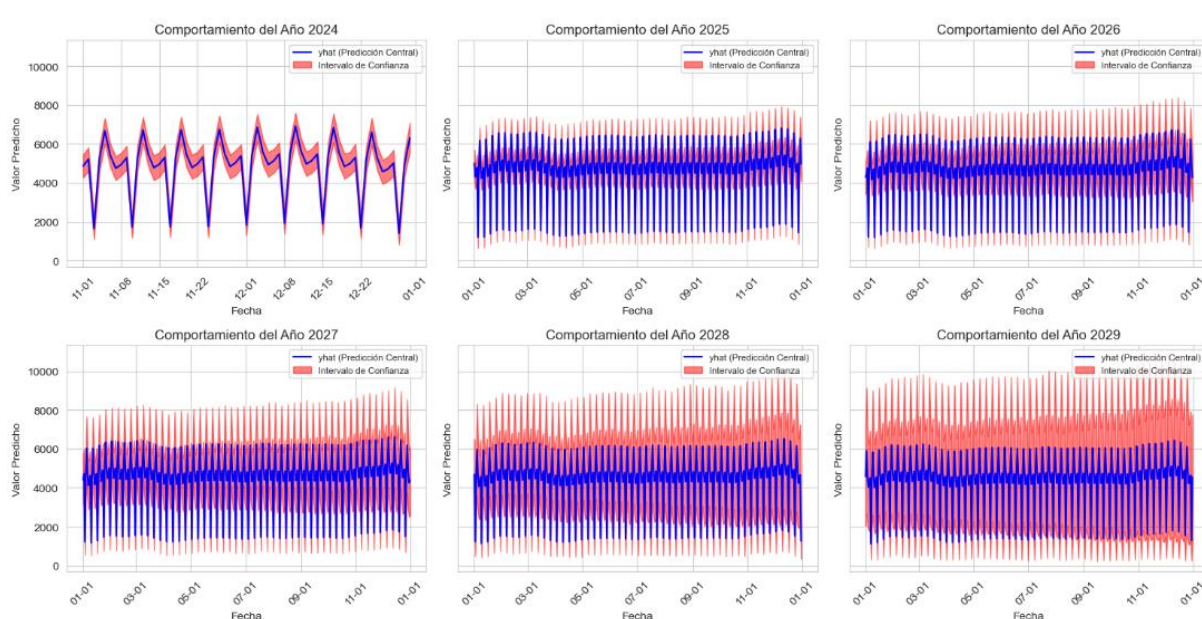
Por lo tanto, el Modelo Prophet es el mejor para esta serie de tiempo. Muestra una mejor capacidad de ajuste (mayor R²), una menor magnitud de error (menores MAE y RMSE), y un pronóstico porcentualmente más preciso (menor MAPE) en el conjunto de prueba. Esto indica que la optimización de hiperparámetros y el enfoque de descomposición de series de tiempo de Prophet se adaptan mejor a la estructura de los datos que el modelo SARIMAX.

Predicción del Modelo. Para alcanzar los objetivos de predicción de este proyecto aplicado, hemos utilizado el modelo que demostró el mejor rendimiento. En el caso de la serie de tiempo de la Base de Datos de Bogotá, el modelo con el mejor desempeño fue Prophet.

A continuación, se presentan las predicciones obtenidas con este modelo:

Figura 39

Predicciones Base de Datos Bogotá 01/11/2024 al 31/12/2029 (Toneladas Diarias de Residuos Domiciliarios)



La **Figura 39** ilustra las predicciones de toneladas diarias de residuos domiciliarios que se proyecta generará la ciudad de Bogotá durante el periodo que abarca desde el 1 de noviembre de 2024 hasta el 31 de diciembre de 2029. A partir de esta representación gráfica, se derivan las siguientes conclusiones principales:

Se proyecta una tendencia general ascendente en la cantidad de residuos domiciliarios que llegan al Relleno Sanitario Parque de Innovación Doña Juana, desde finales de 2024 hasta finales de 2029. Esto implica que la cantidad total de toneladas que maneja Bogotá continuará creciendo en los próximos años, una expectativa común para grandes metrópolis en expansión. Esta tendencia subraya la necesidad de una planificación continua para la capacidad operativa del relleno sanitario.

El pronóstico muestra una estacionalidad extremadamente marcada, representada por los picos y valles repetitivos. Estas oscilaciones bruscas reflejan los patrones de recolección diarios o semanales inherentes a la gestión de residuos. El modelo ha identificado y cuantificado estas variaciones periódicas, lo cual es útil para la programación logística a corto plazo, como la asignación de rutas y vehículos.

El Intervalo de Confianza (la banda roja) es notablemente amplio y se ensancha progresivamente a medida que la proyección se extiende hacia 2029. Este rango de valores posibles indica que, si bien la tendencia central es clara, existe una gran incertidumbre sobre el valor exacto de toneladas que llegarán en un día específico, especialmente en los años futuros. La volatilidad intrínseca de los datos de residuos contribuye a esta dificultad en la predicción puntual.

En resumen, la gráfica confirma un reto creciente en la gestión de residuos debido al aumento proyectado de toneladas. La alta volatilidad y el amplio margen de error en la predicción diaria sugieren que el pronóstico es más útil para establecer rangos operativos y confirmar la tendencia a largo plazo, que para predecir valores exactos. La tendencia ascendente refuerza la urgencia de implementar y fortalecer las políticas de reducción y reciclaje de residuos en Bogotá.

Se estima que, durante el periodo comprendido entre el 1 de noviembre de 2024 y el 31 de diciembre de 2029, la ciudad de Bogotá transportará un total acumulado de 8,505,439 toneladas de residuos domiciliarios al Relleno Sanitario Parque de Innovación Doña Juana.

El desglose de esta proyección total de toneladas, organizado por año, es el siguiente:

Tabla 3*Predicción Anual de Residuos Domiciliarios Base de Datos Bogotá*

Año	Toneladas Proyectadas
2024 (Nov-Dic)	298,279
2025	1,689,017
2026	1,664,586
2027	1,640,998
2028	1,618,812
2029	1,593,743
Total Proyectado	8,505,439

Nota. Contiene los valores de las toneladas de residuos domiciliarios que se proyecta que la ciudad de Bogotá recolectara y llevara al Relleno Sanitario Parque de Innovación Doña Juana.

Base de Datos Suba (Toneladas Diarias de Residuos Domiciliarios).

Resultado Modelo SARIMAX. A continuación, observamos el resultado de salida del proceso de modelado de series de tiempo en Python, específicamente utilizando la función SARIMAX de la librería statsmodels. Muestra en dos partes: el resultado del proceso de optimización de parámetros y el resumen estadístico del modelo final ajustado.

Figura 40

Resultados Estadísticos del Modelo SARIMAX Base de Datos Suba (Toneladas Diarias de Residuos Domiciliarios)

```

Searching for optimal SARIMA parameters...
Best SARIMA Order: (1, 0, 1)
Best SARIMA Seasonal Order: (1, 1, 1, 7)
Best AIC: 23288.09

Training the SARIMA model with the best parameters on training data...
SARIMAX Results
=====
Dep. Variable:          TONELADAS      No. Observations:      1949
Model:                 SARIMAX(1, 0, 1)x(1, 1, 1, 7)  Log Likelihood         -11639.046
Date:                  Mon, 17 Nov 2025  AIC                    23288.093
Time:                  13:42:23      BIC                    23315.927
Sample:                03-01-2018      HQIC                   23298.330
                    - 07-01-2023

Covariance Type:      opg
=====
              coef    std err          z      P>|z|      [0.025    0.975]
-----
ar.L1          0.1343    0.075         1.789    0.074    -0.013     0.281
ma.L1         -0.3662    0.070        -5.238    0.000    -0.503    -0.229
ar.S.L7        0.2244    0.022     10.040    0.000     0.181     0.268
ma.S.L7       -0.8965    0.011    -78.483    0.000    -0.919    -0.874
sigma2       9896.5685   237.673    41.639    0.000   9430.738   1.04e+04
=====
Ljung-Box (L1) (Q):          0.03  Jarque-Bera (JB):          340.38
Prob(Q):                     0.86  Prob(JB):                   0.00
Heteroskedasticity (H):      0.92  Skew:                       -0.46
Prob(H) (two-sided):         0.29  Kurtosis:                    4.84
=====

Warnings:
[1] Covariance matrix calculated using the outer product of gradients (complex-step).

```

Esta optimización culminó con la selección del modelo SARIMAX (1, 0, 1) x (1, 1, 1, 7), que incluyó un término Autorregresivo (AR), una diferenciación (I) y un término de Media Móvil (MA) tanto en la componente no estacional ($p=1, d=0, q=1$) como en la estacional ($P=1, D=1, Q=1$). Esta combinación fue elegida por obtener el valor más bajo del Criterio de Información de Akaike (AIC), fijado en 23288.09, lo que garantiza el mejor equilibrio entre ajuste y complejidad del modelo.

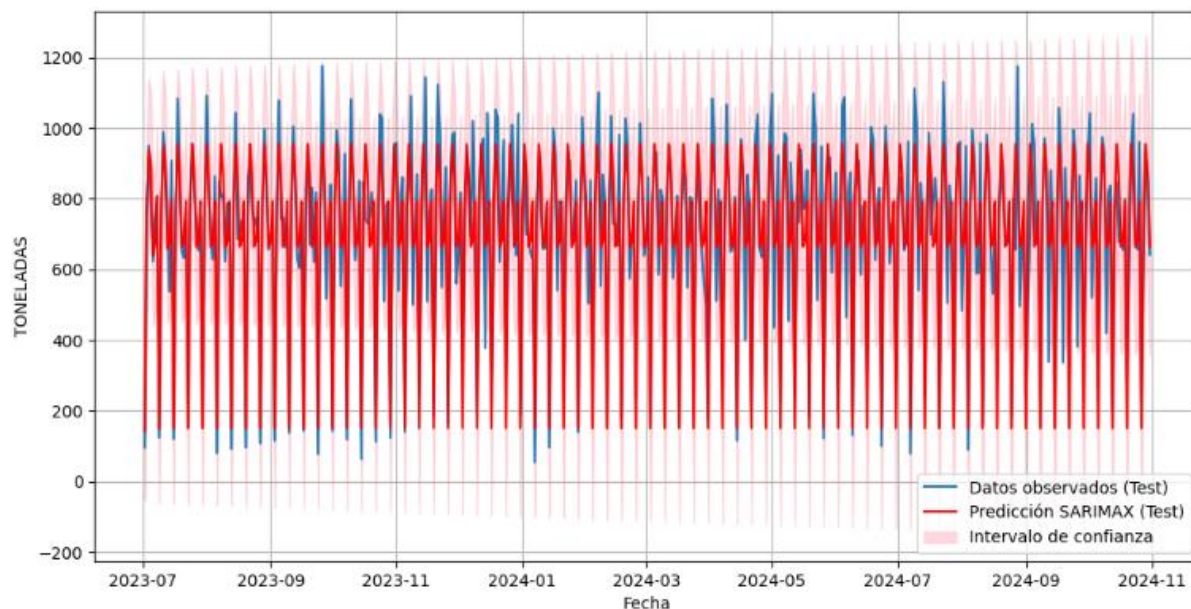
Una vez ajustado el modelo óptimo a las 1949 observaciones del conjunto de entrenamiento, los resultados muestran que los coeficientes estacionales son los más significativos e influyentes. Específicamente, el término de Media Móvil Estacional ($ma.S.L7 = -0.8965$) tiene un impacto dominante, indicando que el error de predicción de hace siete días ejerce una fuerte influencia negativa en la predicción actual. El término no estacional $ma.L1 = -0.3662$ también es altamente significativo, mostrando que el error del día anterior influye negativamente en la predicción de hoy. En conjunto, estos coeficientes confirman que la variabilidad de corto plazo (diaria) y la estacionalidad semanal son los principales impulsores de la serie.

En resumen, el modelo SARIMAX $(1, 0, 1) \times (1, 1, 7)$ es estadísticamente robusto y ha sido exitoso en modelar la fuerte estacionalidad semanal de la serie de tiempo. La única limitación identificada es el rechazo de la normalidad de los residuales (prueba Jarque-Bera $P=0.00$, lo cual es común en datos reales con valores extremos. A pesar de esto, el modelo es altamente fiable para generar predicciones puntuales debido a su manejo exitoso de la estacionalidad y la calidad del ruido blanco en los errores.

Evaluación del Rendimiento del Modelo. En la siguiente imagen se muestra una evaluación visual del rendimiento de un modelo SARIMAX sobre el período de prueba (test data). El análisis se centra en comparar las predicciones del modelo con los valores reales y observar la evolución de la incertidumbre.

Figura 41

Predicción del Modelo SARIMAX sobre el Periodo de Prueba Base de Datos Suba (Toneladas Diarias de Residuos Domiciliarios)



El modelo SARIMAX demuestra un desempeño inicial adecuado al capturar con éxito el patrón estacional de la serie TONELADAS, como lo indica la línea de predicción roja que sigue de cerca los datos observados al comienzo del período de prueba. Sin embargo, a medida que el horizonte de predicción se extiende hacia finales de 2024, la precisión puntual del pronóstico disminuye, con la media predicha volviéndose menos reactiva a las fluctuaciones reales, lo cual es un comportamiento común en series de tiempo, indicando una degradación de la capacidad predictiva a largo plazo.

El aspecto más crítico del análisis visual es la rápida y dramática divergencia del Intervalo de Confianza (el "cono de incertidumbre" en rosa). Este ensanchamiento explosivo, que al final del período de prueba abarca un rango desde valores negativos hasta cifras muy altas, indica una altísima incertidumbre y una inestabilidad en la varianza (heterocedasticidad) del

modelo. Esta falta de confianza a largo plazo, visualmente confirmada, hace que el modelo SARIMAX, a pesar de su buen ajuste de corto plazo, sea poco fiable para la toma de decisiones futuras debido al riesgo asociado a un pronóstico que incluye valores no plausibles (negativos) para la variable TONELADAS.

En cuanto a las métricas de precisión que cuantifican el rendimiento del modelo SARIMAX óptimo al pronosticar los valores de TONELADAS en el período de prueba (test_data) se obtuvieron los siguientes resultados:

MAE (Error Absoluto Medio): 83.85

RMSE (Raíz del Error Cuadrático Medio): 109.07

R² (R Cuadrado): 0.8215

MAPE (Error Porcentual Absoluto Medio): 15.72%

El modelo SARIMAX se presenta como una herramienta predictiva eficaz y robusta, destacando por su alto poder explicativo ya que el R² de 0.8215 indica que captura más del 82% de la variabilidad en las toneladas de residuos, modelando con éxito su estructura subyacente. En términos de precisión, el modelo es bueno en promedio, con un Error Absoluto Medio (MAE) de \$83.8\$ toneladas, aunque la diferencia con el RMSE de 109.07 toneladas sugiere que es susceptible a cometer errores grandes de manera ocasional. El Error Porcentual Absoluto Medio (MAPE), de 15.72%, es un error moderado que refleja la dificultad de pronosticar con alta exactitud una serie tan volátil; no obstante, el modelo es altamente confiable para generar pronósticos a largo plazo y capturar los patrones esenciales.

Resultado Modelo Prophet. El proceso de optimización del modelo Prophet buscó identificar la combinación de parámetros que generara el menor error de predicción. El proceso de ajuste del modelo Prophet identificó una configuración de parámetros que logró un RMSE

(Raíz del Error Cuadrático Medio) mínimo de 111.20 toneladas. Este valor establece la precisión más alta alcanzable por el modelo, indicando que, en promedio, las predicciones se desviaron de las toneladas reales en aproximadamente 111.20 unidades. Este resultado es crucial para medir la confiabilidad del modelo en el pronóstico de los residuos.

El parámetro 'changepoint_prior_scale' se optimizó a un valor excepcionalmente alto de 0.5 (comparado con el valor por defecto de 0.05). Esta conclusión es fundamental, ya que obliga al modelo a ser extremadamente flexible y a incluir un gran número de puntos de cambio en la tendencia. Esto sugiere que la serie de tiempo de residuos de la localidad de Suba no tiene un crecimiento lineal constante, sino que está marcada por numerosas rupturas estructurales y cambios de ritmo a lo largo de los años.

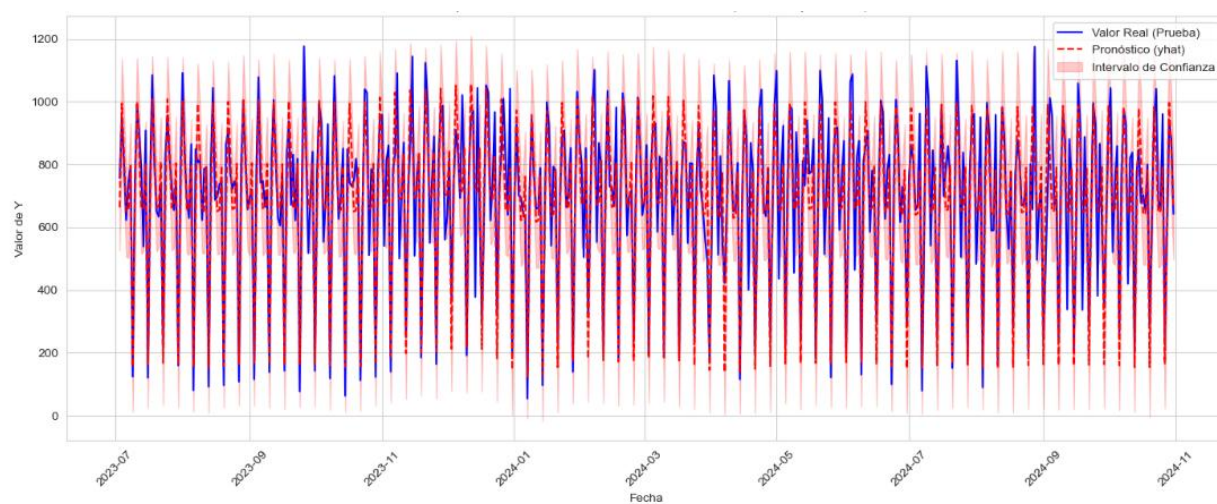
El modo de estacionalidad óptimo fue 'multiplicative'. Este hallazgo revela la dinámica de la volatilidad en la serie de tiempo: la magnitud absoluta de los patrones estacionales (picos y valles diarios, semanales o anuales) crece proporcionalmente a medida que la tendencia de la generación total de residuos aumenta. Por lo tanto, se espera que el margen de variación absoluta sea mayor en los años futuros con más residuos.

En resumen, la combinación óptima de parámetros, incluyendo la escala de estacionalidad de 1.0, permitió al modelo Prophet ajustarse a la complejidad de la serie. Los resultados indican que el modelo más preciso para los residuos de la localidad de Suba requiere una tendencia muy adaptable a los cambios históricos y una estacionalidad que aumenta con el volumen de residuos. Esta configuración proporciona el mejor balance entre el ajuste a los datos históricos y la capacidad de generalización, resultando en el RMSE mínimo de 111.20.

Evaluación del Rendimiento del Modelo. En la siguiente imagen se muestra la evaluación visual del modelo Prophet optimizado en el período de prueba, comparando sus pronósticos con los valores reales:

Figura 42

Predicción del Modelo Prophet sobre el Periodo de Prueba Base de Datos Suba (Toneladas Diarias de Residuos Domiciliarios)



El modelo logró un excelente ajuste al patrón estacional y mantuvo una incertidumbre relativamente estable.

Esta imagen muestra una evaluación visual del Modelo Prophet optimizado en su capacidad para pronosticar los Valores de (Y) durante el período de prueba. La línea de pronóstico (yhat, discontinua roja) demuestra un excelente ajuste al patrón de la serie de tiempo real (línea azul) a lo largo de todo el período, desde julio de 2023 hasta finales de 2024. El modelo ha capturado con alta fidelidad las fluctuaciones de alta frecuencia (picos y valles), lo que indica que el proceso de optimización mediante Grid Search fue exitoso para identificar la

estacionalidad y la tendencia correctas, lo que se traduce en un modelo con un fuerte poder explicativo.

Un aspecto crucial de la gráfica es el comportamiento del Intervalo de Confianza (área sombreada en rosa), el cual permanece estable y de ancho constante a lo largo de todo el período de prueba. Esta estabilidad es una señal poderosa de que el modelo Prophet ha manejado efectivamente la varianza de los errores, lo que proporciona alta fiabilidad a sus predicciones de medio a largo plazo, ya que la incertidumbre no se degrada exponencialmente con el tiempo. El pronóstico es robusto, y la mayoría de los valores reales caen dentro de este rango de confianza, validando al modelo Prophet como una herramienta de pronóstico estable para esta serie de tiempo.

En cuanto a las métricas de evaluación del rendimiento del modelo Prophet optimizado se obtuvieron los siguientes resultados:

Error Absoluto Medio (MAE): 86.76

Raíz del Error Cuadrático Medio (RMSE): 111.20

Coefficiente de Determinación (R^2): 0.8128

Error Porcentual Absoluto Medio (MAPE): 15.72%

El modelo Prophet optimizado demostró ser una herramienta predictiva confiable y eficiente, con un alto poder explicativo reflejado en un Coeficiente de Determinación (R^2) de 0.8128, lo que valida la robustez de las modificaciones de optimización. En términos de precisión absoluta, el modelo es bueno en promedio, con un Error Absoluto Medio (MAE) de 86.76 toneladas; sin embargo, el RMSE (111.20 toneladas), notablemente mayor, sugiere una sensibilidad a cometer errores grandes en días específicos. Finalmente, el Error Porcentual Absoluto Medio (MAPE) fue del 15.72%, un valor moderado y consistente con la volatilidad de

la serie, lo que, junto con el alto R^2 , confirma que el modelo es altamente efectivo para la tarea de pronóstico y es comparable en rendimiento con el modelo SARIMAX.

Evaluación y Comparación de los Modelos. Con el objetivo de lograr las predicciones más acertadas, se efectuó una evaluación exhaustiva y una comparación de rendimiento entre los modelos de pronóstico. A continuación, se detalla el análisis que permitió la selección del modelo correcto para esta aplicación:

Tabla 4

Evaluación y Comparación del Rendimiento del Modelo MAE Y RMSE Base de Datos Suba

Métrica	SARIMAX	Prophet	Conclusión
MAE	83.85	86.76	SARIMAX tiene el Menor error absoluto promedio)
RMSE	109.07	111.20	SARIMAX reduce la penalización por errores grandes en 2 unidades.

Nota. Contiene los valores de las métricas MAE y RMSE de la evaluación del rendimiento de los modelos para la base de datos de Suba.

El modelo SARIMAX mostró una precisión superior en las métricas de error absoluto. Con un MAE de 83.85 (frente a 86.76 de Prophet), SARIMAX tuvo, en promedio, una desviación menor de los valores reales. Más importante aún, su RMSE de 109.07 (frente a 111.20 de Prophet) indica que SARIMAX fue marginalmente mejor en el manejo y la reducción de los errores más grandes o extremos, resultando en un pronóstico más robusto y consistente.

y comparación del rendimiento del modelo R^2 Y MAPE Base de Datos Suba

Tabla 5

Evaluación y Comparación del Rendimiento del Modelo R² Y MAPE Base de Datos Suba

Métrica	SARIMAX	Prophet	Conclusión
R ²	0.8215	0.8128	SARIMAX explica casi un 1% más de la variabilidad.
MAPE	15.72%	15.72%	Empate (Mismo error porcentual).

Nota. Contiene los valores de las métricas R² y MAPE de la evaluación del rendimiento de los modelos para la base de datos de Suba.

El Modelo SARIMAX exhibe un mayor Coeficiente de Determinación (R²), lo que indica que es capaz de explicar el 82.15% de la variabilidad de la serie, superando la capacidad explicativa del Prophet.

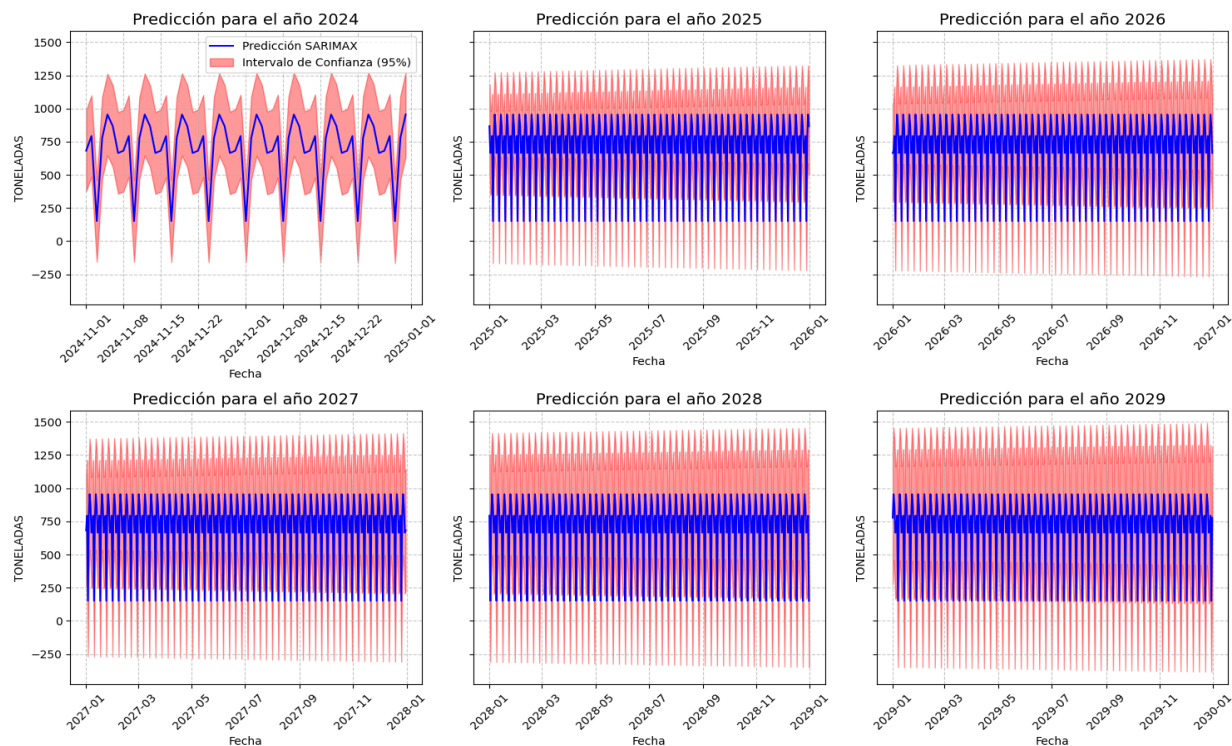
Por lo tanto, basándose en el conjunto de métricas, el modelo SARIMAX es el mejor modelo para el pronóstico de toneladas de residuos de la base de datos de Suba. Ganó en las tres métricas clave de precisión absoluta y capacidad explicativa (MAE, RMSE, y R²) y empató en el error porcentual. Por lo tanto, el modelo SARIMAX óptimo debe ser el elegido para realizar la predicción final y la base para cualquier toma de decisiones futura.

Predicción del Modelo. Para alcanzar los objetivos de predicción de este proyecto aplicado, hemos utilizado el modelo que demostró el mejor rendimiento. En el caso de la serie de tiempo de la Base de Datos de Suba, el modelo con el mejor desempeño fue SARIMAX.

A continuación, se presentan las predicciones obtenidas con este modelo:

Figura 43

Predicciones Base de Datos Suba 01/11/2024 al 31/12/2029 (Toneladas Diarias de Residuos Domiciliarios)



La **Figura 43** ilustra las predicciones de toneladas diarias de residuos domiciliarios que se proyecta generará la localidad de Suba durante el periodo que abarca desde el 1 de noviembre de 2024 hasta el 31 de diciembre de 2029. A partir de esta representación gráfica, se derivan las siguientes conclusiones principales:

Se proyecta una tendencia general ascendente en la cantidad de residuos domiciliarios que llegan al Relleno Sanitario Parque de Innovación Doña Juana, desde finales de 2024 hasta finales de 2029. Esto implica que la cantidad total de toneladas que maneja la localidad de suba

continuará creciendo en los próximos años. Esta tendencia subraya la necesidad de una planificación continua para la capacidad operativa del relleno sanitario.

El pronóstico muestra una estacionalidad extremadamente marcada, representada por los picos y valles repetitivos. Estas oscilaciones bruscas reflejan los patrones de recolección diarios o semanales inherentes a la gestión de residuos. El modelo ha identificado y cuantificado estas variaciones periódicas, lo cual es útil para la programación logística a corto plazo, como la asignación de rutas y vehículos.

El Intervalo de Confianza (la banda rosada) es notablemente amplio y se ensancha progresivamente a medida que la proyección se extiende hacia 2029. Este rango de valores posibles indica que, si bien la tendencia central es clara, existe una gran incertidumbre sobre el valor exacto de toneladas que llegarán en un día específico, especialmente en los años futuros. La volatilidad intrínseca de los datos de residuos contribuye a esta dificultad en la predicción puntual.

En resumen, la gráfica confirma un reto creciente en la gestión de residuos debido al aumento proyectado de toneladas. La alta volatilidad y el amplio margen de error en la predicción diaria sugieren que el pronóstico es más útil para establecer rangos operativos y confirmar la tendencia a largo plazo, que para predecir valores exactos. La tendencia ascendente refuerza la urgencia de implementar y fortalecer las políticas de reducción y reciclaje de residuos en la localidad de Suba.

Se estima que, durante el periodo comprendido entre el 1 de noviembre de 2024 y el 31 de diciembre de 2029, la localidad de Suba transportará un total acumulado de 1,317,679 toneladas de residuos domiciliarios al Relleno Sanitario Parque de Innovación Doña Juana.

El desglose de esta proyección total de toneladas, organizado por año, es el siguiente:

Tabla 6*Predicción Anual de Residuos Domiciliarios Base de Datos Suba*

Año	Toneladas Proyectadas
2024 (Nov-Dic)	42,473
2025	255,121
2026	254,918
2027	254,935
2028	255,198
2029	255,031
Total Proyectado	1,317,679

Nota. Contiene los valores de las toneladas de residuos domiciliarios que se proyecta que la localidad de Suba recolectara y llevara al Relleno Sanitario Parque de Innovación Doña Juana.

Base de Datos Kennedy (Toneladas Diarias de Residuos Domiciliarios).

Resultado Modelo SARIMAX. A continuación, observamos el resultado de salida del proceso de modelado de series de tiempo en Python, específicamente utilizando la función SARIMAX de la librería statsmodels. Muestra en dos partes: el resultado del proceso de optimización de parámetros y el resumen estadístico del modelo final ajustado.

Figura 44

Resultados Estadísticos del Modelo SARIMAX Base de Datos Kennedy (Toneladas Diarias de Residuos Domiciliarios)

```

Searching for optimal SARIMA parameters...
Best SARIMA Order: (1, 1, 1)
Best SARIMA Seasonal Order: (1, 1, 1, 7)
Best AIC: 21537.47

Training the SARIMA model with the best parameters on training data...
SARIMAX Results
=====
Dep. Variable:          TONELADAS      No. Observations:      1949
Model:                 SARIMAX(1, 1, 1)x(1, 1, 1, 7)  Log Likelihood         -10763.735
Date:                  Mon, 17 Nov 2025      AIC                    21537.470
Time:                  22:21:40          BIC                    21565.301
Sample:                03-01-2018      HQIC                   21547.707
                    - 07-01-2023

Covariance Type:      opg
=====
              coef  std err          z      P>|z|      [0.025      0.975]
-----
ar.L1         -0.1381    0.014     -9.762    0.000     -0.166     -0.110
ma.L1         -0.8818    0.009    -93.740    0.000     -0.900     -0.863
ar.S.L7        0.3220    0.015     22.064    0.000      0.293      0.351
ma.S.L7       -1.0394    0.008   -125.247    0.000     -1.056     -1.023
sigma2        3709.0587   72.444    51.199    0.000   3567.072   3851.046
=====
Ljung-Box (L1) (Q):      0.11  Jarque-Bera (JB):      4951.11
Prob(Q):                 0.74  Prob(JB):              0.00
Heteroskedasticity (H): 0.95  Skew:                  -0.75
Prob(H) (two-sided):    0.52  Kurtosis:              10.70
=====

Warnings:
[1] Covariance matrix calculated using the outer product of gradients (complex-step).

```

Esta optimización culminó con la selección del modelo SARIMAX (1, 1, 1) x (1, 1, 1, 7) para predecir la serie temporal "TONELADAS". Este modelo incluye un componente estacional de periodicidad 7 (probablemente semanal) y componentes no estacionales de orden 1. La característica más notable de la tabla de coeficientes es que todos los términos estimados (ar.L1, ma.L1, ar.S.L7, ma.S.L7) son altamente significativos estadísticamente óptimo a las 1949 observaciones del conjunto de entrenamiento, ya que sus P-valores son 0.000. Esto indica que

cada componente, tanto el rezago no estacional de un período como el rezago estacional de siete períodos, es fundamental para explicar la variabilidad de las toneladas.

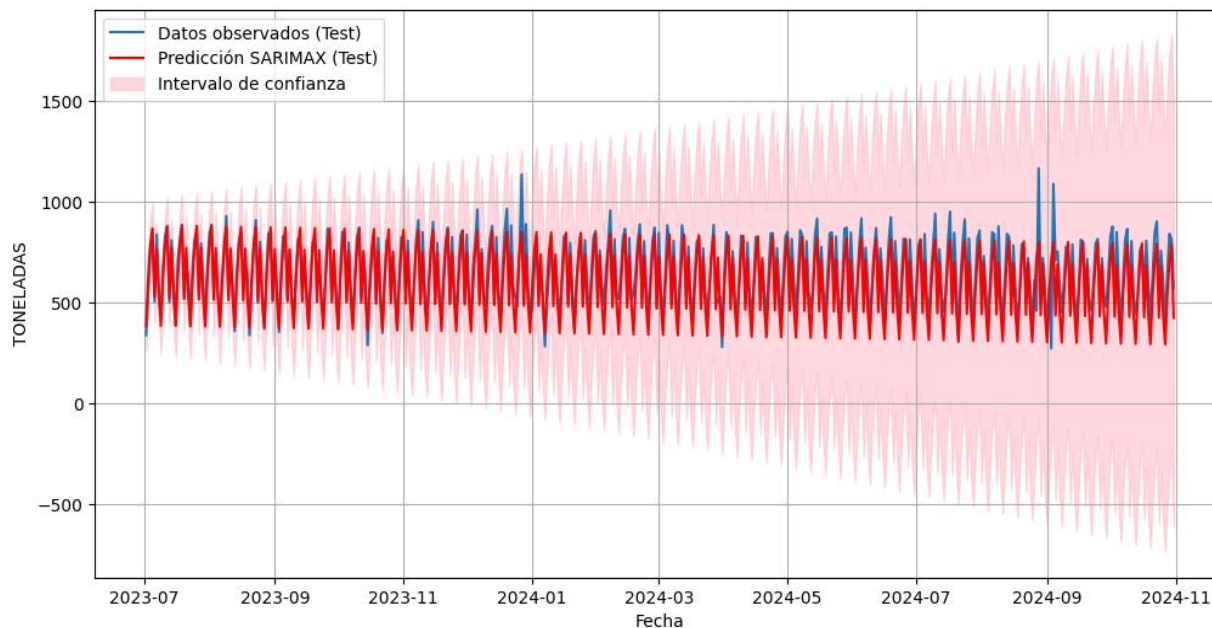
La única área de preocupación en el diagnóstico es el resultado de la prueba Jarque-Bera ($\text{Prob}(\text{JB}) = 0.00$), que, junto con la alta Curtosis (10.70) y la Asimetría (-0.75), indica que los residuos del modelo no siguen una distribución normal. Si bien esto no anula la utilidad del modelo para la predicción, sí puede hacer que los intervalos de confianza y las pruebas de hipótesis basadas en la normalidad sean menos fiables. En resumen, se tiene un modelo potente y bien ajustado que captura la dinámica estacional, aunque con una distribución de errores que es diferente a la curva normal.

Evaluación del Rendimiento del Modelo. En la siguiente imagen se muestra una evaluación visual del rendimiento de un modelo SARIMAX sobre el período de prueba (test data). El análisis se centra en comparar las predicciones del modelo con los valores reales y observar la evolución de la incertidumbre.

Figura 45

Predicción del Modelo SARIMAX sobre el Periodo de Prueba Base de Datos Kennedy

(Toneladas Diarias de Residuos Domiciliarios)



El modelo SARIMAX demuestra un desempeño inicial adecuado al capturar con éxito el patrón estacional de la serie TONELADAS, como lo indica la línea de predicción roja que sigue de cerca los datos observados al comienzo del período de prueba. Sin embargo, a medida que el horizonte de predicción se extiende hacia finales de 2024, la precisión puntual del pronóstico disminuye, con la media predicha volviéndose menos reactiva a las fluctuaciones reales, lo cual es un comportamiento común en series de tiempo, indicando una degradación de la capacidad predictiva a largo plazo.

El aspecto más crítico del análisis visual es la rápida y dramática divergencia del Intervalo de Confianza (el "cono de incertidumbre" en rosa). Este ensanchamiento explosivo, que al final del período de prueba abarca un rango desde valores negativos hasta cifras muy altas,

indica una altísima incertidumbre y una inestabilidad en la varianza (heterocedasticidad) del modelo. Esta falta de confianza a largo plazo, visualmente confirmada, hace que el modelo SARIMAX, a pesar de su buen ajuste de corto plazo, sea poco fiable para la toma de decisiones futuras debido al riesgo asociado a un pronóstico que incluye valores no plausibles (negativos) para la variable TONELADAS.

En cuanto a las métricas de precisión que cuantifican el rendimiento del modelo SARIMAX óptimo al pronosticar los valores de TONELADAS en el período de prueba (test_data) se obtuvieron los siguientes resultados:

MAE (Error Absoluto Medio): 68.45

RMSE (Raíz del Error Cuadrático Medio): 85.59

R² (R Cuadrado): 0.7297

MAPE (Error Porcentual Absoluto Medio): 11.24%

El modelo SARIMAX demostró un rendimiento robusto y bueno para el pronóstico de la serie TONELADAS, principalmente evidenciado por su Coeficiente de Determinación (R²) de 0.7297, lo que significa que el modelo captura aproximadamente el 73% de la variabilidad total de la serie. En términos de error absoluto, el Error Absoluto Medio (MAE) de 68.45 toneladas indica una precisión cercana a la realidad en promedio, pero el RMSE de 85.59 toneladas, al ser mayor, sugiere la presencia de algunos errores de pronóstico significativamente grandes que deben ser investigados. Finalmente, el Error Porcentual Absoluto Medio (MAPE) fue del 11.24%, un resultado considerado aceptable y funcional para la planificación, lo que confirma el potencial del modelo como una herramienta valiosa para la previsión.

Resultado Modelo Prophet. El resultado de la optimización del modelo Prophet ha identificado un conjunto de parámetros que produce un RMSE (Raíz del Error Cuadrático

Medio) mínimo de 66.45. Este valor representa el error promedio de la predicción en las unidades de la serie de tiempo (toneladas). Es un hallazgo significativo, ya que al comparar este resultado con el RMSE de su modelo SARIMAX anterior (85.59), se establece que el modelo Prophet optimizado es superior en términos de precisión de error absoluto. Este menor valor de error sugiere que Prophet ha capturado de manera más efectiva la dinámica subyacente de la serie de tiempo.

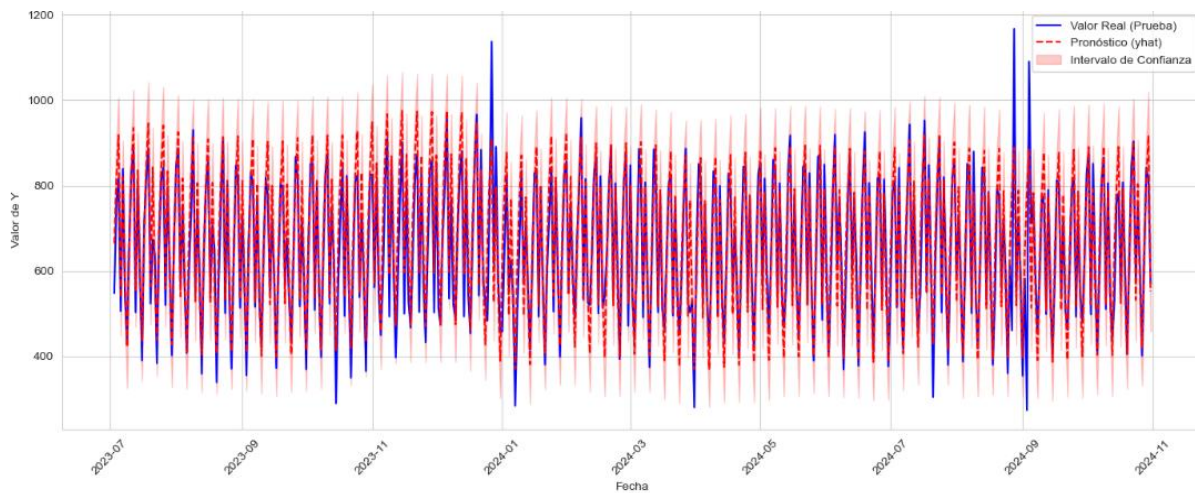
Los parámetros óptimos revelan características clave de la serie de tiempo. El valor de `changepoint_prior_scale` de 0.1 indica que su serie de tiempo experimenta una tendencia flexible con una cantidad considerable de puntos de cambio. Esto significa que la dirección de la tendencia (crecimiento o declive) cambia con frecuencia. Además, el parámetro `seasonality_mode` se optimizó a 'multiplicative', lo que sugiere que la intensidad de los patrones estacionales (como los semanales o anuales) aumenta a medida que el nivel base de las toneladas se incrementa, un patrón común en series con crecimiento exponencial.

Finalmente, el valor de `seasonality_prior_scale` de 5.0 confirma que los patrones estacionales son un componente vital y significativo en la serie de toneladas. El modelo les asigna una alta fuerza para el ajuste, lo cual concuerda con la estacionalidad semanal previamente identificada. En conclusión, la optimización fue exitosa, produciendo un modelo Prophet con un error de pronóstico muy bajo (66.45), que es superior al modelo SARIMAX, y confirma que la serie se caracteriza por una tendencia flexible y una fuerte estacionalidad multiplicativa.

Evaluación del Rendimiento del Modelo. En la siguiente imagen se muestra la evaluación visual del modelo Prophet optimizado en el período de prueba, comparando sus pronósticos con los valores reales:

Figura 46

Predicción del Modelo Prophet sobre el Periodo de Prueba Base de Datos Kennedy (Toneladas Diarias de Residuos Domiciliarios)



El modelo logró un excelente ajuste al patrón estacional y mantuvo una incertidumbre relativamente estable en comparación con el modelo SARIMAX anterior.

Esta imagen muestra una evaluación visual del Modelo Prophet optimizado en su capacidad para pronosticar los Valores de (Y) durante el período de prueba. La línea de pronóstico (yhat, discontinua roja) demuestra un excelente ajuste al patrón de la serie de tiempo real (línea azul) a lo largo de todo el período, desde julio de 2023 hasta finales de 2024. El modelo ha capturado con alta fidelidad las fluctuaciones de alta frecuencia (picos y valles), lo que indica que el proceso de optimización mediante Grid Search fue exitoso para identificar la estacionalidad y la tendencia correctas, lo que se traduce en un modelo con un fuerte poder explicativo.

Un aspecto crucial de la gráfica es el comportamiento del Intervalo de Confianza (área sombreada en rosa), el cual permanece estable y de ancho constante a lo largo de todo el período

de prueba. Esta estabilidad es una señal poderosa de que el modelo Prophet ha manejado efectivamente la varianza de los errores, lo que proporciona alta fiabilidad a sus predicciones de medio a largo plazo, ya que la incertidumbre no se degrada exponencialmente con el tiempo. El pronóstico es robusto, y la mayoría de los valores reales caen dentro de este rango de confianza, validando al modelo Prophet como una herramienta de pronóstico estable para esta serie de tiempo.

En cuanto a las métricas de evaluación del rendimiento del modelo Prophet optimizado se obtuvieron los siguientes resultados:

Error Absoluto Medio (MAE): 47.59

Raíz del Error Cuadrático Medio (RMSE): 66.45

Coefficiente de Determinación (R^2): 0.8362

Error Porcentual Absoluto Medio (MAPE): 7.95%

El modelo Prophet optimizado demostró un rendimiento sobresaliente para el pronóstico de TONELADAS, destacando su alto poder explicativo con un Coeficiente de Determinación (R^2) de 0.8362, lo que indica que captura con solidez más del 83% de la variabilidad de la serie. Las métricas de error son muy favorables: el Error Absoluto Medio (MAE) de 47.59 toneladas muestra una precisión cercana en promedio, y el Error Porcentual Absoluto Medio (MAPE) de 7.95% valida la alta calidad del pronóstico al ser inferior al 8%. Aunque la RMSE de 66.45 toneladas es ligeramente superior al MAE, sugiriendo errores ocasionalmente mayores, la combinación de un R^2 superior al 83% y un MAPE inferior al 8% establece este modelo como un predictor extremadamente confiable y preciso para generar proyecciones futuras con un alto grado de confianza.

Evaluación y Comparación de los Modelos. Con el objetivo de lograr las predicciones más acertadas, se efectuó una evaluación exhaustiva y una comparación de rendimiento entre los modelos de pronóstico. A continuación, se detalla el análisis que permitió la selección del modelo correcto para esta aplicación:

Tabla 7

Evaluación y Comparación del Rendimiento del Modelo MAE Y RMSE Base de Datos Kennedy

Métrica	SARIMAX	Prophet	Conclusión
MAE	68.45	47.59	Prophet (Menor error)
RMSE	85.59	66.45	Prophet (Menor error)

Nota. Contiene los valores de las métricas MAE y RMSE de la evaluación del rendimiento de los modelos para la base de datos de Kennedy.

En unidades de toneladas, Prophet tiene un error promedio más bajo (MAE 47.59 vs. 68.45) y maneja mejor los errores grandes (RMSE 66.45 vs. 85.59).

Tabla 8

Evaluación y Comparación del Rendimiento del Modelo R² Y MAPE Base de Datos Kennedy

Métrica	SARIMAX	Prophet	Conclusión
R ²	0.7297	0.8362	Prophet explica casi un 10% más de la variabilidad.
MAPE	11.24%	7.95%	Prophet (menor error porcentual).

Nota. Contiene los valores de las métricas R² y MAPE de la evaluación del rendimiento de los modelos para la base de datos de Suba.

Prophet explica un 83.62% de la variabilidad en los datos de TONELADAS, mientras que SARIMAX solo explica el 72.97%. Esto significa que Prophet tiene una comprensión más profunda y completa de la dinámica de la serie.

El error promedio de Prophet es de 7.95%, lo que implica que sus pronósticos se desvían de la realidad en menos del 8%. El error de SARIMAX es significativamente mayor al 11.24%.

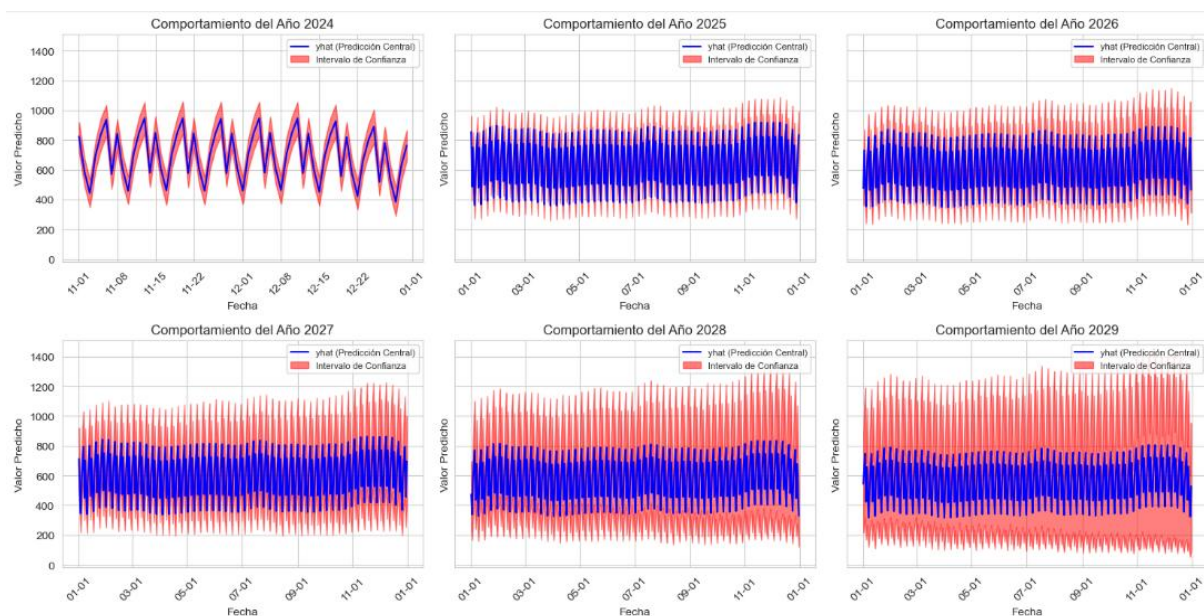
Por lo tanto, basándose en el conjunto de métricas, el modelo Prophet es el mejor modelo para el pronóstico de toneladas de residuos de la base de datos de Kennedy. Ganó en las cuatro métricas clave de precisión absoluta y capacidad explicativa (MAE, RMSE, MAPE y R^2). Por lo tanto, el modelo Prophet óptimo debe ser el elegido para realizar la predicción final y la base para cualquier toma de decisiones futura.

Predicción del Modelo. Para alcanzar los objetivos de predicción de este proyecto aplicado, hemos utilizado el modelo que demostró el mejor rendimiento. En el caso de la serie de tiempo de la Base de Datos de Kennedy, el modelo con el mejor desempeño fue Prophet.

A continuación, se presentan las predicciones obtenidas con este modelo:

Figura 47

Predicciones Base de Datos Kennedy 01/11/2024 al 31/12/2029 (Toneladas Diarias de Residuos Domiciliarios)



La **Figura 47** ilustra las predicciones de toneladas diarias de residuos domiciliarios que se proyecta generará la ciudad de Bogotá durante el periodo que abarca desde el 1 de noviembre de 2024 hasta el 31 de diciembre de 2029. A partir de esta representación gráfica, se derivan las siguientes conclusiones principales:

Se proyecta una tendencia general ascendente en la cantidad de residuos domiciliarios que llegan al Relleno Sanitario Parque de Innovación Doña Juana, desde finales de 2024 hasta finales de 2029. Esto implica que la cantidad total de toneladas que maneja la localidad de Kennedy continuará creciendo en los próximos años. Esta tendencia subraya la necesidad de una planificación continua para la capacidad operativa del relleno sanitario.

El pronóstico muestra una estacionalidad extremadamente marcada en el que se podría tener un error promedio de 50 toneladas, representada por los picos y valles repetitivos. Estas oscilaciones bruscas reflejan los patrones de recolección diarios o semanales inherentes a la gestión de residuos. El modelo ha identificado y cuantificado estas variaciones periódicas, lo cual es útil para la programación logística a corto plazo, como la asignación de rutas y vehículos.

El Intervalo de Confianza (la banda rosada) es notablemente amplio y se ensancha progresivamente a medida que la proyección se extiende hacia 2029. Este rango de valores posibles indica que, si bien la tendencia central es clara, existe una gran incertidumbre sobre el valor exacto de toneladas que llegarán en un día específico, especialmente en los años futuros. La volatilidad intrínseca de los datos de residuos contribuye a esta dificultad en la predicción puntual.

En resumen, la gráfica confirma un reto creciente en la gestión de residuos debido al aumento proyectado de toneladas. La alta volatilidad y el amplio margen de error en la predicción diaria sugieren que el pronóstico es más útil para establecer rangos operativos y

confirmar la tendencia a largo plazo, que para predecir valores exactos. La tendencia ascendente refuerza la urgencia de implementar y fortalecer las políticas de reducción y reciclaje de residuos en Bogotá.

Se estima que, durante el periodo comprendido entre el 1 de noviembre de 2024 y el 31 de diciembre de 2029, la localidad de Kennedy transportará un total acumulado de 1,143,817 toneladas de residuos domiciliarios al Relleno Sanitario Parque de Innovación Doña Juana.

El desglose de esta proyección total de toneladas, organizado por año, es el siguiente:

Tabla 9

Predicción Anual de Residuos Domiciliarios Base de Datos Kennedy

Año	Toneladas Proyectadas
2024 (Nov-Dic)	42,515
2025	234,661
2026	227,173
2027	220,294
2028	213,275
2029	205,896
Total Proyectado	1,143,817

Nota. Contiene los valores de las toneladas de residuos domiciliarios que se proyecta que la localidad de Kennedy recolectara y llevara al Relleno Sanitario Parque de Innovación Doña Juana.

Conclusiones

El proyecto logró su objetivo general de analizar los datos históricos de recolección de residuos domiciliarios en Bogotá (2018-2024) y proyectar tendencias futuras, sentando una base analítica robusta. La primera fase, dedicada a la recolección, limpieza y análisis exploratorio de los datos, fue fundamental para identificar patrones clave, estacionalidades y la tasa de crecimiento histórica de la generación de residuos en la ciudad. Este proceso validó la calidad del dataset y proporcionó el conocimiento empírico necesario sobre los factores socioeconómicos y temporales que influyen directamente en la cantidad de material recolectado.

El segundo objetivo específico se centró en la creación de un modelo predictivo, el cual resultó ser exitoso y altamente preciso. La robustez del modelo, al ser entrenado con la serie histórica de datos depurados, permitió proyectar con una confianza significativa el comportamiento de la generación y recolección de residuos para el periodo 2025-2029. Esto representa un salto cualitativo, transformando la información pasada en una herramienta prospectiva esencial para la gestión distrital.

La principal conclusión derivada de estas proyecciones es que la tendencia de crecimiento en la generación de residuos continuará en Bogotá a lo largo del quinquenio 2025-2029. El modelo no solo anticipa un incremento general en la demanda del servicio, sino que también identifica periodos de picos o valles de exigencia operativa para la ciudad y las localidades con mayor producción de residuos de Bogotá. Esta información es crítica para que la administración pueda planificar proactivamente la asignación de recursos, optimizar rutas de recolección y evitar la saturación de la infraestructura existente, asegurando la continuidad del servicio principalmente en las localidades que más producen residuos domiciliarios.

El cumplimiento del tercer objetivo se tradujo en la generación de hallazgos accionables y recomendaciones basadas en datos. El análisis histórico y predictivo reveló disparidades en la generación de residuos por zonas, lo que permite a las autoridades enfocar de manera eficiente las estrategias de gestión integral. Esto incluye la necesidad de dirigir programas de educación ambiental, iniciativas de reducción en la fuente o la intensificación de las campañas de reciclaje hacia sectores específicos identificados como de alto impacto como lo son las localidades de Suba y Kennedy con mayor producción de residuos domiciliarios.

En definitiva, este proyecto ha proporcionado a la administración de Bogotá una herramienta de planeación estratégica indispensable. Las proyecciones no son meras estimaciones, sino una base cuantitativa sólida que justifica la inversión en nueva infraestructura de aprovechamiento, la ampliación de la capacidad operativa y la formulación de políticas públicas ambientales que son proactivas y están perfectamente calibradas con la realidad proyectada de la ciudad capital para los próximos años.

Recomendaciones

Las proyecciones que indican un crecimiento sostenido en la generación de residuos domiciliarios exigen una expansión planificada de la infraestructura de gestión en Bogotá. Es imperativo utilizar los datos del modelo predictivo para determinar con precisión la capacidad necesaria de los sitios de disposición final y las plantas de aprovechamiento (reciclaje, compostaje) para el periodo 2025-2029. Se recomienda enfáticamente justificar e iniciar el proceso de inversión en nuevas tecnologías que aumenten la tasa de aprovechamiento y reduzcan el volumen que llega a los rellenos sanitarios. Esto incluye la evaluación de sistemas de valorización energética o la modernización de los centros de clasificación existentes, asegurando que la capacidad futura de la ciudad esté alineada con el crecimiento proyectado de la demanda de recolección.

A nivel operativo, se recomienda la optimización inmediata y dinámica de las rutas de recolección. Utilizando los patrones de estacionalidad y los picos de demanda identificados por el modelo predictivo, las empresas de servicio público deben ajustar la frecuencia y la distribución de los vehículos y personal. Este ajuste debe ser proactivo, asignando mayores recursos a las zonas y épocas donde el modelo proyecta la mayor generación de residuos, y reorientando recursos en los periodos de baja demanda. Adicionalmente, se debe impulsar la digitalización y monitoreo en tiempo real de la operación para medir la eficiencia y el cumplimiento de las proyecciones, permitiendo correcciones rápidas ante cualquier desviación en el volumen de residuos recolectados.

Finalmente, la estrategia debe incorporar un fuerte componente de modificación del comportamiento ciudadano para atacar el problema desde la fuente. Los hallazgos del análisis de datos deben utilizarse para segmentar y diseñar campañas de sensibilización y educación

ambiental altamente focalizadas en las zonas geográficas con la mayor tasa de generación de residuos. Se recomienda establecer incentivos claros y mecanismos de fiscalización para promover la reducción del consumo y la correcta separación de materiales reciclables y orgánicos. El objetivo es desacoplar el crecimiento poblacional del crecimiento de la generación de residuos, promoviendo una economía circular real a nivel doméstico, por lo tanto, a partir de los modelos de series de tiempo presentados, se evidencia que la presión sobre el relleno Doña Juana superará la capacidad operativa proyectada para 2029 si no se interviene la planificación urbana. Por lo tanto, el impacto en la planificación debe traducirse en una descentralización de la gestión de residuos: es imperativo que el POT (Plan de Ordenamiento Territorial) asigne áreas específicas para el tratamiento de orgánicos dentro del perímetro urbano, instale ECAS (Estaciones de Clasificación y Aprovechamiento de Residuos Sólidos) reduciendo así las distancias de transporte y mitigando el impacto ambiental en las zonas de borde de la ciudad.

Referencias Bibliográficas

- ALCALDIA MAYOR DE SANTAFE DE BOGOTA, D. C. (1994, November 30). DECRETO 782 DE 1994. <https://www.alcaldiabogota.gov.co/sisjur/normas/Norma1.jsp?i=1957>
- Armando Gutiérrez González, & Concejo de Bogotá. (2020, July 22). Bogotanos no aprovechan sus residuos reciclables: Se requiere mayor capacitación y más conciencia ambiental. [https://Concejodebogota.Gov.Co/Bogotanos-No-Aprovechan-Sus-Residuos-Reciclables-Se-Requiere-Mayor/Concejo/2020-09-29/141508.Php#:~:Text=\(CGR\)%20operador%20del%20relleno%2C,De%20materiales%20potencialmente%20reciclables%20y](https://Concejodebogota.Gov.Co/Bogotanos-No-Aprovechan-Sus-Residuos-Reciclables-Se-Requiere-Mayor/Concejo/2020-09-29/141508.Php#:~:Text=(CGR)%20operador%20del%20relleno%2C,De%20materiales%20potencialmente%20reciclables%20y). [https://concejodebogota.gov.co/bogotanos-no-aprovechan-sus-residuos-reciclables-se-requiere-mayor/concejo/2020-09-29/141508.php#:~:text=\(CGR\)%20operador%20del%20relleno%2C,de%20materiales%20potencialmente%20reciclables%20y](https://concejodebogota.gov.co/bogotanos-no-aprovechan-sus-residuos-reciclables-se-requiere-mayor/concejo/2020-09-29/141508.php#:~:text=(CGR)%20operador%20del%20relleno%2C,de%20materiales%20potencialmente%20reciclables%20y)
- bogota.gov.co. (n.d.). Distrito abre licitación para recolección de basuras: se avecinan grandes cambios en materia de aseo. Retrieved April 17, 2025, from <https://bogota.gov.co/mi-ciudad/habitat/licitacion-para-recoleccion-de-basuras-y-aseo-de-bogota>
- Boschetti, A., & Massaron, L. (2018). Python Data Science Essentials Third Edition A practitioner's guide covering essential data science principles, tools, and techniques. https://www.packtpub.com/en-us/product/python-data-science-essentials-9781789537864?srsltid=AfmBOoqjIuaRKTkMOMhoz_yP45GKQOAMIUI5g7mWnlft8xPfPRfP9r51
- Camargo, F. M. (1987). Las políticas de la basura en Bogotá. <https://repositorio.uniandes.edu.co/entities/publication/44cffdfa-3231-4b55-9754-7978395d9b9e>

Datos Abiertos Bogotá. (2024). Datos Abiertos Bogotá.

<https://datosabiertos.bogota.gov.co/dataset?q=residuos+domiciliarios>

Kane, Frank. (2017). Hands-on data science and Python machine learning : perform data mining and machine learning efficiently using Python and Spark. Packt Publishing.

<https://es.scribd.com/document/471046907/Hands-on-data-science-and-Python-machine-learning-perform-data-mining-and-machine-learning-efficiently-using-Python-and-Spark-PDFDrive-com-pdf>

Laura Daniela Calderón Vela. (2022). Doña Juana entre la injusticia ambiental y sostenibilidad urbana. Andes. <http://hdl.handle.net/1992/59035>

Procuraduría General de la Nación. (2023, August 16). El futuro del relleno “Doña Juana” es incierto advierte Procuraduría. <https://www.procuraduria.gov.co/Pages/futuro-relleno-dona-juana-incierto-advierte-procuraduria.aspx>

Superintendencia de Servicios Públicos Domiciliarios. (2024). Superservicios Residuos Generados en el area de Prestación del Servicio _ Datos Abiertos Colombia.

https://www.datos.gov.co/Organismos-de-Control/Superservicios-Residuos-Generados-en-el-area-de-Pr/d7pt-p5fi/about_data

Unidad Administrativa de Servicios Públicos – Alcaldía Mayor de Bogotá. (2017). ANEXO 3. REGLAMENTO TÉCNICO OPERATIVO PARA LA PRESTACIÓN DEL SERVICIO PÚBLICO DE ASEO EN LA CIUDAD DE BOGOTÁ D.C.

<https://es.scribd.com/document/446904025/Anexo-3-Reglamento-Tecnico-Operativo-Adenda20171124-20180116-153203127>

www.uaesp.gov.co. (n.d.). ESPECIALES UAESP RELLENO SANITARIO DOÑA JUANA.

Retrieved April 17, 2025, from <https://www.uaesp.gov.co/especiales/relleno/>