

## **Análisis de tendencia de vehículos híbridos y eléctricos en Colombia**

Juan Felipe Campo Trochez

José Domingo Pino Díaz Granados

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

2024

## Resumen

El cambio climático ha generado gran preocupación a nivel mundial, esto ha generado la implementación de medidas que permitan el uso de tecnologías que empleen energías alternas que reemplacen los motores de combustión interna los cuales generan una cantidad significativa de gases de efecto invernadero. Actualmente, la venta de vehículos eléctricos (EV) en el mercado colombiano ha presentado un incremento destacado, sin embargo, este valor sigue siendo inferior a la cantidad de venta de vehículos que funcionan con combustibles fósiles. Esta tendencia global requiere un análisis detallado para comprender sus causas y proponer soluciones.

El impacto de la contaminación ambiental en Colombia ha llevado a que algunas ciudades densamente pobladas adopten medidas de mitigación, como la restricción de la movilidad mediante el "pico y placa" o el día sin carro y moto, evidenciando la significativa repercusión en la calidad de vida de las personas. El abordaje del problema mediante técnicas de *machine learning* es explorado desde diferentes ángulos, desde lo tecnológico hasta lo funcional y económico.

Aunque se han analizado barreras que restringen la adopción masiva de EV en otros países, esta investigación propone un análisis de los factores locales en Colombia, con el objetivo de identificar causas y proponer estrategias de mitigación.

En este marco, el objetivo general pretende diseñar un modelo predictivo usando técnicas de *machine learning* que permita pronosticar la venta de vehículos eléctricos en Colombia así mismo, describir los factores asociados con la baja tendencia de ventas de EV en Colombia, identificar causas.

**Palabras clave:** Vehículos eléctricos, aprendizaje automatizado, series de tiempo, pronósticos, machine learning.

## Abstract

Climate change has generated great concern worldwide, which has led to the implementation of measures that allow the use of technologies that use alternative energies to replace internal combustion engines, which generate a significant amount of greenhouse gases. Currently, the sale of electric vehicles (EV) in the Colombian market has shown a significant increase, however, this value is still lower than the number of sales of vehicles that run on fossil fuels. This global trend requires a detailed analysis to understand its causes and propose solutions.

The impact of environmental pollution in Colombia has led some densely populated cities to adopt mitigation measures, such as restricting mobility through "pico y placa" or the day without cars and motorcycles, evidencing the significant impact on people's quality of life. The approach to the problem through machine learning techniques is explored from different angles, from the technological to the functional and economic. Although barriers restricting mass adoption of EV have been analyzed in other countries, this research proposes an analysis of local factors in Colombia, with the aim of identifying causes and proposing mitigation strategies.

Within this framework, the general objective is to design a predictive model using machine learning techniques that allows forecasting the sale of electric vehicles in Colombia, as well as to describe the factors associated with the low trend of EV sales in Colombia, and to identify causes.

**Keywords:** Electric Vehicles, machine learning, time series, forecasting.

## Tabla de Contenido

Introducción .....	9
Justificación .....	11
Objetivos.....	13
Planteamiento del Problema .....	14
Marco Conceptual y Teórico.....	16
Metodología .....	23
Análisis Exploratorio de Datos .....	24
Tipos de Vehículos .....	29
Marcas.....	34
BMW.....	35
BYD .....	40
Renault .....	44
Departamentos .....	49
Bogotá.....	53
Antioquia.....	60
Cundinamarca .....	65
Predicción de Ventas a 2030 .....	72
Proyección con Prophet.....	74
Proyección Ventas 2023 .....	76
Conclusiones.....	79
Recomendaciones .....	82
Referencias.....	83

## Lista de Tablas

<b>Tabla 1</b> <i>Cantidad de Vehículos Según la Clase</i> .....	29
<b>Tabla 2</b> <i>Ventas por marca y Tipo de Vehículo</i> .....	35
<b>Tabla 3</b> <i>Comparación de la Métrica R2 Obtenida con los Modelos Predictivos</i> .....	48
<b>Tabla 4</b> <i>Vehículos Matriculados en Antioquia, Bogotá D. C. y el Resto del País</i> .....	49
<b>Tabla 5</b> <i>Cantidad de Vehículos Eléctricos e Híbridos por Departamento</i> .....	50
<b>Tabla 6</b> <i>Resultados de Optimización Modelos de Regresión Bogotá</i> .....	55
<b>Tabla 7</b> <i>Resultados de Optimización de Modelos de Regresión para Bogotá</i> .....	56
<b>Tabla 8</b> <i>Valores Óptimos de Prophet con Datos Quincenales</i> .....	57
<b>Tabla 9</b> <i>Métricas de los Modelos</i> .....	59
<b>Tabla 10</b> <i>Resultados de Optimización de Modelos de Regresión para Antioquia (<math>R^2</math>)</i> .....	62
<b>Tabla 11</b> <i>Resultados de Optimización de Modelos de Regresión para Antioquia (AIC)</i> .....	62
<b>Tabla 12</b> <i>Valores Óptimos para Prophet con Datos Mensuales de Antioquia</i> .....	64
<b>Tabla 13</b> <i>Métricas de los Modelos para Antioquia</i> .....	65
<b>Tabla 14</b> <i>Resultados de Optimización de Modelos de Regresión para Cundinamarca (<math>R^2</math>)</i> .....	67
<b>Tabla 15</b> <i>Resultados de Optimización de Modelos de Regresión para Cundinamarca (AIC)</i> .....	68
<b>Tabla 16</b> <i>Valores Óptimos Prophet para Datos Mensuales de Cundinamarca</i> .....	70

## Lista de Figuras

<b>Figura 1</b> <i>Metodología</i> .....	23
<b>Figura 2</b> <i>Matriz de Completitud de Datos</i> .....	24
<b>Figura 3</b> <i>Tipos de Clases de Vehículos Eléctricos a Nivel Nacional</i> .....	25
<b>Figura 4</b> <i>Matriz de Correlación entre Variables del Dataset</i> .....	26
<b>Figura 5</b> <i>Cantidad de Vehículos Registrados por Año a Nivel Nacional</i> .....	28
<b>Figura 6</b> <i>Venta de EV en los Principales Departamentos de Colombia</i> .....	31
<b>Figura 7</b> <i>Motocicletas Eléctricas Registradas en Colombia</i> .....	32
<b>Figura 8</b> <i>Ventas de las Principales marcas de EV en Colombia</i> .....	34
<b>Figura 9</b> <i>Descomposición de la Serie de Tiempo para BMW</i> .....	36
<b>Figura 10</b> <i>Modelo SE para BMW</i> .....	37
<b>Figura 11</b> <i>Correlaciones sobre la Serie Estacionaria (Original) para BMW</i> .....	38
<b>Figura 12</b> <i>Optimización de Parámetros del Modelo AR para BMW</i> .....	38
<b>Figura 13</b> <i>Modelo ARMA para BMW (1, 0, 1)</i> .....	39
<b>Figura 14</b> <i>Modelo ARIMA para BMW (1, 1, 1)</i> .....	39
<b>Figura 15</b> <i>Descomposición de la Serie de Tiempo para BYD</i> .....	40
<b>Figura 16</b> <i>Datos para BYD</i> .....	41
<b>Figura 17</b> <i>Correlaciones Sobre la Serie no Diferenciada (Original) para BYD</i> .....	42
<b>Figura 18</b> <i>Optimización de Parámetros del Modelo ARMA para BYD</i> .....	42
<b>Figura 19</b> <i>Modelo ARMA para BYD (1, 0, 2)</i> .....	43
<b>Figura 20</b> <i>Modelo ARIMA para BYD (4, 1, 5)</i> .....	43
<b>Figura 21</b> <i>Descomposición de la Serie de Tiempo para Renault</i> .....	44
<b>Figura 22</b> <i>Modelo SE para Renault</i> .....	45

<b>Figura 23</b> <i>Correlaciones de la Serie no Diferenciada (Original) para Renault</i> .....	45
<b>Figura 24</b> <i>Modelo ARMA para Renault (1, 0, 1)</i> .....	46
<b>Figura 25</b> <i>Modelo ARIMA para Renault (1, 1, 1)</i> .....	47
<b>Figura 26</b> <i>Ejemplo de Permutación de Parámetros Obtenida para Modelos AR</i> .....	48
<b>Figura 27</b> <i>Distribución de Vehículos por Departamento</i> .....	51
<b>Figura 28</b> <i>Distribución de Vehículos Eléctricos por Departamento</i> .....	52
<b>Figura 29</b> <i>Ventas de EV en los Principales Departamentos de Colombia</i> .....	53
<b>Figura 30</b> <i>Descomposición de la Serie de Tiempo para Bogotá D. C.</i> .....	54
<b>Figura 31</b> <i>Datos para Bogotá D. C.</i> .....	55
<b>Figura 32</b> <i>Datos para Bogotá D. C. (0, 2, 5)</i> .....	56
<b>Figura 33</b> <i>Resultados del Modelo Prophet para Bogotá (<math>R^2</math>: -0.04)</i> .....	58
<b>Figura 34</b> <i>Descomposición de la Serie de Tiempo de Antioquia</i> .....	61
<b>Figura 35</b> <i>Predicción del Modelo de SE para Antioquia (<math>R^2</math>: -0.37)</i> .....	61
<b>Figura 36</b> <i>Modelo ARIMA para Antioquia (3, 1, 2)</i> .....	63
<b>Figura 37</b> <i>Resultados del Modelo Prophet para Antioquia (<math>R^2</math>: -0.26)</i> .....	64
<b>Figura 38</b> <i>Descomposición de la Serie de Tiempo de EV en Cundinamarca</i> .....	66
<b>Figura 39</b> <i>Predicción del Modelo SE para la venta de EV en Cundinamarca (<math>R^2</math>: -0.49)</i> .....	67
<b>Figura 40</b> <i>Modelo ARIMA (0, 1, 1) para la Serie de Tiempo de Cundinamarca</i> .....	68
<b>Figura 41</b> <i>Modelo ARIMA (5, 1, 5) para la Serie de Tiempo de Cundinamarca</i> .....	69
<b>Figura 42</b> <i>Resultados del Modelo Prophet para Cundinamarca (<math>R^2</math>: -0.55)</i> .....	70
<b>Figura 43</b> <i>Métricas de los Modelos Aplicados a la Serie de Cundinamarca</i> .....	71
<b>Figura 44</b> <i>Descomposición Estacional de la Serie Total EV e híbridos</i> .....	72
<b>Figura 45</b> <i>Proyección de Ventas de EV e Híbridos hasta 2030 Basada en SE</i> .....	73

<b>Figura 46</b> <i>Efecto del Número de Registros Sobre la Predicción</i> .....	74
<b>Figura 47</b> <i>Proyección de Ventas Basada en el Método Prophet</i> .....	75
<b>Figura 48</b> <i>Ventas de Vehículos Eléctricos e Híbridos en Colombia</i> .....	76
<b>Figura 49</b> <i>Proyección de Ventas 2023 con Prophet</i> .....	77
<b>Figura 50</b> <i>Proyección de Ventas 2023 con Variación Hiperparámetros en Prophet</i> .....	78

## Introducción

El deterioro medioambiental como parte de la actividad social representa uno de los desafíos más importantes de la actualidad, promoviendo estrategias de mitigación sostenibles, en particular, aquellas que buscan reducir la emisión de gases de efecto invernadero. Este tema adopta una dimensión crítica en los grandes centros urbanos debido a la gran cantidad de vehículos con motores a combustión que concentran la emisión de estos gases en un espacio reducido, llevando a los gobiernos locales a diseñar estrategias de mitigación, esencialmente enfocadas a la limitar la circulación diaria de vehículos. Es aquí donde el vehículo eléctrico (EV) surge como una alternativa importante para reducir el impacto en la calidad de las ciudades.

Sin embargo, en Colombia, la adopción del EV sigue siendo limitada, a pesar del incremento sostenido desde 2011, según los registros del Ministerio de Transporte. Esta situación motiva la investigación desde varios frentes en aras de favorecer una alternativa de movilidad que en países desarrollados ha demostrado efectos totalmente positivos para el medioambiente.

En ciudades como Medellín y Bogotá D. C. la calidad del aire se ha convertido en un factor de seguimiento prioritario por parte de las autoridades ambientales debido a su lento pero sostenido deterioro. Las restricciones vehiculares solo resultan en un freno temporal a la polución que resulta en un sostenimiento aceptable de los niveles de contaminación del aire, impactando negativamente la movilidad urbana, trayendo consigo efectos psicosociales y económicos que exceden el alcance de este estudio pero que se menciona con el ánimo de llamar la atención sobre la relevancia que el tema reviste.

La adopción masiva del EV también exige estudios de dinámicas de mercado, el precio de la energía en las diferentes regiones del país, la carga adicional sobre la red eléctrica y preferencias del consumidor, entre otras, a fin de contar con insumos que permitan elaborar estrategias multisectoriales contundentes que favorezcan el aumento de la circulación de este tipo de vehículos desplazando (eventualmente) los vehículos con motor de combustión. Se trata sin duda, de un tema vasto e interesante, de alto impacto social y determinante en la calidad de vida de generaciones futuras.

Dado lo anterior, mediante el uso de herramientas de *machine learning*, se propondrá un modelo de serie temporal que pronostique las ventas de EV en Colombia hasta el año 2030, con base en los registros públicos gestionados por el Ministerio de Transporte y disponibles en el sitio web Datos Abiertos Colombia. Con este modelamiento se pretende aportar un punto de vista relevante, con base científica y dentro del alcance de la ciencia de datos, que permita mejorar la comprensión de la fenomenología tratada y que sirva de referente a posteriores investigaciones.

## Justificación

La investigación propuesta se basa en la baja penetración de EV en Colombia, evidenciado en los datos de (Datos Abiertos Colombia, 2022) y en las ventas de estos vehículos durante el primer trimestre de este año (Ramos Robayo & Herrera Romero, 2024a). Colombia enfrenta desafíos significativos en la adopción masiva de EV como los señalados por (Hernández-Ambato et al., 2022a) lo cual plantea interrogantes que justifican una investigación detallada por el impacto positivo y directo que la adopción de EV puede generar el medioambiente de las núcleos urbanos del país, inicialmente, lo cual iría acorde a la creciente conciencia mundial sobre la importancia de la transición hacia formas de movilidad más sostenibles, que ha impulsado tecnologías de movilidad alternativas como los EV, dado que buena parte de la contaminación medioambiental en las ciudades proviene de los gases de efecto invernadero generado por vehículos con motores de combustión interna (Castro Mediavilla et al., 2017).

Esta investigación generará valor teórico sobre los estudios abordados en este tema encontrando causas y generando recomendaciones útiles para la toma decisiones de fondo que promuevan la adopción del EV como una alternativa a la mitigación de los efectos de la contaminación generados, en buena medida, por los medios de transporte terrestre impulsados con motores térmicos.

El proyecto propuesto permitirá una mejor comprensión de la problemática y formulación de soluciones basadas en el análisis de datos disponibles en fuentes públicas, principalmente gubernamentales, que sustentarán las recomendaciones para promover la masificación del EV en el país y derivando en un impacto positivo para el medio ambiente, por cuanto se trata de estimular la adopción de energías limpias cuyo efecto será notorio en centros urbanos, traduciéndose en mejor calidad de vida para las comunidades y de forma colateral, el impacto económico en el país derivado del incremento en las ventas de los EV (Afandizadeh et al., 2023).

Durante el curso del proyecto serán empleadas herramientas de *machine learning* para identificar causas y predecir comportamientos de las ventas y de otras variables que se consideren importantes durante su avance, sirviendo de referente al conocimiento aplicado en el análisis de una problemática local, realista y de impacto ambiental, social, económico y tecnológico. De esta forma el proyecto ofrecerá información útil derivada de la aplicación práctica de la ciencia de datos, mostrando su relevancia en la generación de valor académico.

## Objetivos

Desarrollar un modelo de *machine learning* para el pronóstico de la demanda de vehículos híbridos y eléctricos en Colombia para el año 2024.

Determinar los factores que presentan mayor influencia en la demanda de vehículos híbridos y eléctricos

Identificar el modelo que ofrezca el mejor rendimiento en términos de precisión de pronóstico y capacidad de generalización

Estimar la cantidad de vehículos eléctricos en el mercado colombiano para el año 2030.

## Planteamiento del Problema

En Colombia, el sector transporte genera el alrededor del 12 % de los gases de efecto invernadero, de los cuales el 78 % corresponde a vehículos terrestres (IDEAM et al., 2022), frente a esto, desde el año 2019 el gobierno colombiano ha implementó políticas que permitían incentivar la transición hacia vehículos que usen nuevas energías. La Implementación de estas tecnologías en Colombia, ha presentado un rápido crecimiento, sin embargo, aún es un mercado muy nuevo y las previsiones para cumplir estas metas no son alentadoras.

Esta incertidumbre se puede atribuir a causas como la baja oferta de modelos de vehículos eléctricos que en 2022 el 48.3 % de los autos vendidos eran fabricados por BYD, distribución que cambió en 2023 pasando a ser el 19 % del total de vehículos eléctricos vendidos. Un factor importante para explicar esta baja oferta de vehículos es las políticas de importación las cuales otorgan una cantidad máxima de vehículos a importar cada año a los cuales se les aplica descuento de importación, a pesar de que estos cupos buscan fomentar la producción interna, aún limitan la entrada de estos vehículos al país.

Otro factor importante es los altos precios de venta de estos automóviles lo cual se atribuye al tipo de tecnología y su reciente introducción, estos vehículos presentan precios de venta partiendo de los 60 millones de pesos aproximadamente, lo cual los hace poco asequibles para la población en general.

Finalmente, el bajo desarrollo de la red de carga de vehículos es una preocupación ya que esto reduce la confianza de los colombianos y prefieren seguir comprando vehículos a gasolina debido a la facilidad para encontrar estaciones de combustible, actualmente se sabe que existen alrededor de 200 puntos de carga para vehículos eléctricos y 3.500 equipos de carga de hogares y empresas, los cual es un número que limita este mercado (Vargas, 2023).

El abordaje del problema empleando técnicas de *machine learning* ha sido tratado desde varios puntos de vista, principalmente desde lo tecnológico en aras de comprender el impacto en la red eléctrica, como lo ha expuesto Rojas Quiroga et al. (2019) o Zhao et al. (2023); también se ha revisado desde lo funcional y económico, como el estudio de Bampos et al. (2024), los cuales brindan una base de análisis de la problemática desde la ciencia de datos. También se han examinado los elementos que influyen la decisión de compra de un EV, como el estudio de Giansoldati et al. (2020), pero en gran medida, con foco en el número de estaciones de carga y vida útil de las baterías de los EV, como lo planteado por Mouhy-Ud-Din et al. (2024) y Lo Franco et al. (2023a).

Si bien se han analizado barreras que restringen la masificación de los EV en otros países, como el precio, según lo hallado por Giansoldati et al. (2020) o por Pamidimukkala et al. (2023), o el impacto en la red eléctrica, como lo indica Mouhy-Ud-Din et al. (2024), entre otras, en esta investigación se propone un análisis de los factores locales inherentes al caso Colombiano, identificando sus causas y proponiendo estrategias para su mitigación.

## Marco Conceptual y Teórico

Machine learning. El aprendizaje automático es la habilidad de cambiar según los estímulos externos y en especial, recordando las experiencias previas, el aprendizaje automático surge como alternativa para solucionar problemas que puedan modelarse y traducirse a lenguaje informático, con el propósito de hacer predicciones sobre los valores de entrada. Esta área de la informática se clasifica en tres grandes ramas o paradigmas: el aprendizaje supervisado, el aprendizaje no supervisado y el aprendizaje por refuerzo (Bonaccorso, 2017).

El aprendizaje supervisado abarca varias técnicas como la regresión lineal, árboles de regresión, regresión con soporte vectorial, redes neuronales y métodos de ensamble, entre otros; aplicados en trabajos de modelado y predicción como el realizado por Selvaraj & Vairavasundaram (2023), en el que estiman el estado de carga de las baterías de EV mediante la selección óptima de hiperparámetros de varios modelos de *machine learning*. Aquí también se aplicaron métodos de ensamble, que consiste en la agrupación de varios algoritmos para tomar ventaja de su poder de predicción conjunto y con los que se obtienen medidas de exactitud y precisión superiores a las que se lograrían con dichos algoritmos por separado.

De acuerdo con (Madhavan, 2015), los algoritmos más usados en el aprendizaje supervisado son: árboles de decisión, regresión lineal, regresión logística, PCA y K-Means, siendo los algoritmos de análisis de series temporales parte del aprendizaje supervisado, específicamente de los métodos de regresión. El modelado de series temporales será empleado en este proyecto para la predicción de las ventas de EV.

Árboles de decisión. En el contexto del *machine learning*, se trata del más simple de los modelos predictivos que registra las entradas y salidas de los nodos representándolas gráficamente, basándose en la probabilidad, siendo fácil de visualizar, facilitando la interpretación de patrones ocultos en los datos, sin embargo, puede incurrir en errores cuando la cantidad de datos es elevada, además de demandar recursos informáticos exponencialmente según la cantidad de variables. Puede aplicarse para clasificación de variables categóricas y regresión con variables numéricas (Madhavan, 2015).

Regresión lineal. Es el modelo de predicción más ampliamente extendido y uno de los más simples de implementar, basado en el estudio lineal de la relación aditiva entre variables y que facilita la comprensión de la influencia de cada uno de los factores representados (Madhavan, 2015).

Regresión logística. Si bien, es conocida como una regresión, en realidad se trata de un algoritmo de clasificación que determina la probabilidad de que una variable asuma un valor dicotómico (cero o uno, verdadero o falso, etc.), por lo que requiere que las variables analizadas sean transformadas a una escala de cero a uno (Bonaccorso, 2017).

El aprendizaje no supervisado se aplica en casos en los que los datos se registran sin elementos que los identifiquen como, por ejemplo, un conjunto de datos con coordenadas y una magnitud de naturaleza no especificada; en este caso, el algoritmo de aprendizaje no supervisado se empleará para identificar los patrones de agrupación que siguen los datos e inferir su significado (Bonaccorso, 2017).

Marino & Marufuzzaman (2020) emplearon análisis de componentes principales (PCA) y K-means para proponer una metodología de optimiza el despliegue de estaciones de carga para EV en ciudades latinoamericanas en expansión, encontrando ahorros superiores al 18 % en el clúster energético si se integra la electromovilidad entre los edificios y las estaciones de carga.

Agrupamiento jerárquico. Es un proceso mediante el que se agrupan las instancias de un conjunto de datos identificando patrones o características comunes, y se repite hasta conseguir un grupo de jerarquía máxima que contenga todos los demás grupos identificados, método conocido como aglomerativo (Kyriakides & Margaritis, 2019).

A diferencia de los paradigmas anteriores, en el aprendizaje por refuerzo (*reinforcement learning*) el modelo aprende a tomar decisiones secuenciales para alcanzar un objetivo específico, interactuando con un entorno dinámico y aprendiendo a través de la retroalimentación en forma de recompensas o penalizaciones en el entorno, de acuerdo con Madhavan (2015), quien relaciona los siguientes como los algoritmos más usados en este paradigma: aprendizaje de diferencia temporal y *Q-learning*.

Se han obtenido modelos basados en *Q-learning* y aplicando redes neuronales para el pronóstico de demanda de carga de EV con una exactitud superior al 50 % sobre los pronósticos obtenidos con redes neuronales únicamente (Dabbaghjamesh et al. (2021), como se citó en Bampos et al. (2024)).

Serie temporal. Es una secuencia de observaciones ordenadas cronológicamente a intervalos estrictamente regulares, sobre uno o varios fenómenos, en cuyos casos adopta el nombre de serie univariante o escalar, o serie multivariante o vectorial, respectivamente (Mauricio, 2007) . El empleo del análisis de series temporales es imprescindible cuando se estudian fenómenos de ocurrencia permanente y que son registrados a intervalos regulares, permitiendo obtener componentes de dicha serie que revelan aspectos indetectables a simple vista si se le representa gráficamente.

AR Models. Los modelos autorregresivos son métodos para la predicción de valores de una serie de tiempo con base en registros pasados en lugar de variables independientes, empleando un modelo de regresión. Este método tiene en cuenta la homocedasticidad (la varianza del error es constante) y la estacionalidad (Pal & Prakash, 2017). La aplicación de modelos autorregresivos facilitará el cálculo de proyecciones regionales, al momento de estimar las ventas por los departamentos más representativos en el país.

ARIMA. Promedio móvil autorregresivo integrado (*Autoregressive Integrated Moving Average*) es también conocido como modelo Box-Jenkins, es particularmente útil cuando la serie de tiempo a analizar no es estacionaria. Además, ARIMA emplea la diferenciación iterativa sobre series de tiempo, hasta eliminar el efecto de estacionalidad sobre la serie y realizar una predicción confiable (Pal & Prakash, 2017).

Vehículos eléctricos. De acuerdo con Afandizadeh et al. (2023), Hernández-Ambato et al. (2022b) y Zarazua De Rubens (2019), los vehículos eléctricos (EV) surgen como alternativa a la movilidad particular y de servicio público (principalmente), contribuyendo a la reducción de gases de efecto invernadero y a la descarbonización de la sociedad. Pese a múltiples barreras técnicas (Golsefidí et al., 2023) y esencialmente, de precios (Pamidimukkala et al., 2023), las ventas de los vehículos presentan una tendencia ascendente a nivel global sin embargo se aprecia una ralentización significativa en Colombia (Steer, 2020).

Dos de los mayores retos que experimenta el crecimiento de los EV en Colombia son el alto costo respecto a los vehículos con motores de combustión, y las dificultades de la ampliación de la capacidad de la red eléctrica colombiana, cuya generación a partir de hidroeléctricas representa el 61 % del total generado.

Otro factor crítico es la optimización del suministro eléctrico en horas pico, cuando los EV se encuentren en etapa de carga y la caída en el consumo durante las horas de mayor circulación, lo cual no solo lleva a un planteamiento de nuevas soluciones de infraestructura sino tarifarias con el consecuente impacto en el mercado energético y en el gasto doméstico de los habitantes.

Este fenómeno ha sido abordado por investigadores en varios países debido al impacto directo sobre la decisión de los consumidores al momento de optar por la adquisición de un nuevo vehículo y las repercusiones en los hábitos de consumo, así como en las expectativas económicas de las empresas generadoras de energía (Afandizadeh et al.; Giansoldati et al., 2020; Kamis & Susan Abraham, 2024; Lo Franco et al., 2023; Marino & Marufuzzaman, 2020; Núñez Hernández & Arcos-Vargas, 2018; Ramos Robayo & Herrera Romero, 2024; Rojas Quiroga et al., 2019; Xing et al., 2021).

En el contexto de la predicción de ventas de autos eléctricos, los algoritmos de series temporales ofrecen un enfoque especializado que aprovecha la estructura implícita en los datos, permitiendo identificar patrones, tendencias y estacionalidades, elementos que permiten ahondar en la comprensión del fenómeno estudiado resultando clave en la generación de predicciones claras. En el análisis de series temporales aplicado al pronóstico de carga eléctrica de los EV realizado por Buzna et al. (2019) obtuvo un desempeño 26 % superior al obtenido al aplicar modelos de *random forest*.

Afandizadeh et al. (2023) generaron un modelo basado en modelos convolucionales que permitió apreciar el impacto del precio, las políticas fiscales y las especificaciones de los vehículos en la intención de compra por parte de los clientes en Estados Unidos a partir de registros de ventas de vehículos proporcionados por empresas distribuidoras que aceptaron participar en dicho estudio, y que generó información para que fabricantes y concesionarios conozcan mejor su mercado objetivo y estructuren estrategias de mercadeo efectivas.

Un método similar fue empleado por Ramos-Pérez et al. (2023) para comprobar la capacidad de predicción de los modelos de *machine learning* aplicados al pronóstico de demanda energética y generación fotovoltaica en España, captando de manera efectiva características no lineales y correlaciones temporales a partir de datos históricos reales tomados de bases de datos públicas (planta fotovoltaica experimental ubicada en la Escuela Técnica Superior de Ingeniería en Sevilla).

Marino & Marufuzzaman (2020) estiman ahorros promedio del 18.24 % en costos de energía derivados de la integración de la electromovilidad, demostrando que la construcción de estaciones de carga representará más ahorros que de no construirlas, a partir de su estudio basado en análisis de componentes principales, clasificación K-means, modelo matemático, con base en datos públicos suministrados por el Instituto Nacional de Estadística e Informática (Perú).

El estudio realizado por Golsefidi et al. (2023) muestra la proyección del costo y la demanda de energía para cubrir los requerimientos energéticos que representa el incremento estimado de EV en Dundee (Escocia), a partir de datos públicos suministrados por la ciudad de Dundee, Escocia, con registros que van de octubre de 2017 a septiembre de 2018 y empleando modelos de *machine learning* basados en procesos gaussianos.

Selvaraj & Vairavasundaram (2023) emplearon modelos de *machine learning* basados en regresión lineal, árboles de decisión, regresión de soporte vectorial, métodos de ensamble, procesos de regresión gaussiano y redes neuronales artificiales en su estudio para estimar el estado de carga de las baterías de los EV, midiendo sus desempeños y determinando los hiperparámetros más influyentes en cada aplicación en India.

Por ello, el proyecto de investigación presentado busca explorar los datos recabados y eventualmente, datos adicionales que permitan un modelamiento de tendencia en la adquisición de los EV en el país con los estados actuales de variables tales como la cobertura eléctrica zonal en el país, el ingreso económico medio de los hogares y el precio de los EV propiamente.

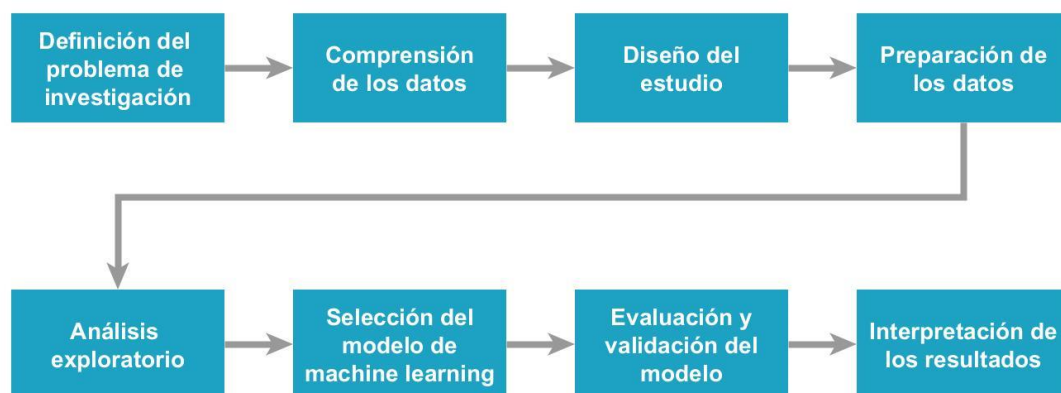
Se espera que las conclusiones emanadas de este proyecto brinden una visión tanto a particulares como a instituciones, para la formulación de medidas favorecedoras de la adquisición de medios de transporte que mitiguen el impacto de los vehículos de combustión interna, principalmente en los centros urbanos.

## Metodología

El proyecto se delinea por los estándares de la investigación aplicada, siguiendo el enfoque mixto, bajo el diseño explicativo secuencial (DEXPLIS), siguiendo los siguientes pasos se explica en González Mares (2019):

**Figura 1**

*Metodología*



Una vez seleccionadas las características más relevantes que puedan influir en las ventas de EV en Colombia, se construyan nuevas características derivadas de las variables disponibles (de ser necesario), y codificar las variables categóricas, se evaluarán varios modelos predictivos basado en el análisis de series temporales, dado que hay evidencia de su desempeño en la literatura técnica recolectada.

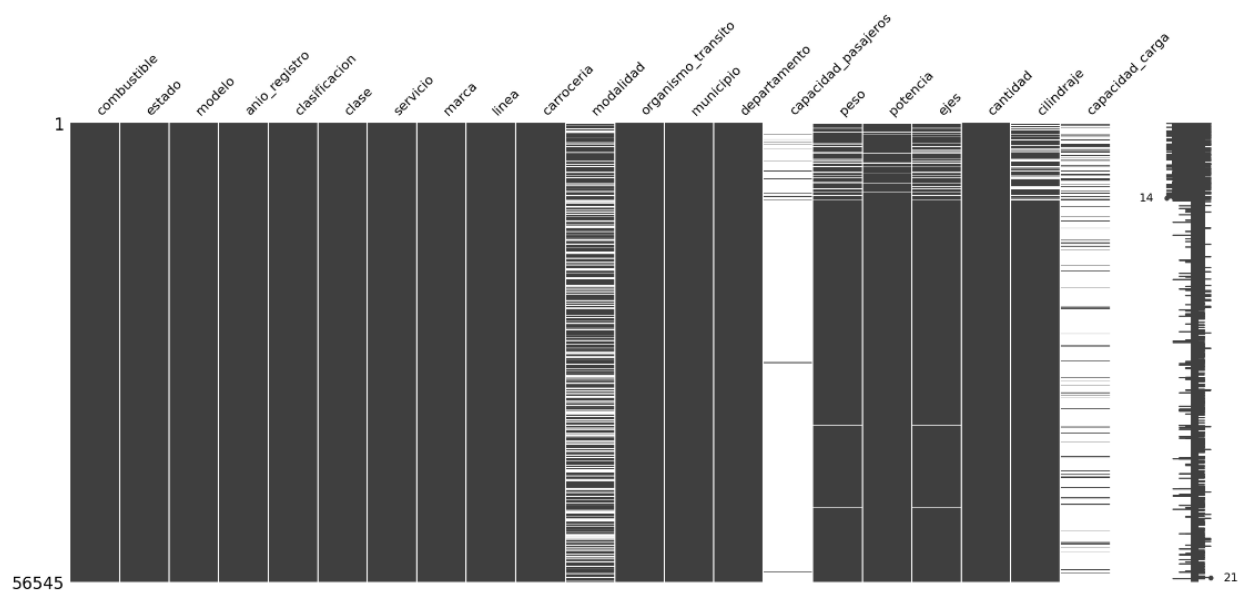
Los resultados generados por el modelo serán analizados, así como las relaciones entre las variables predictoras y la variable objetivo (ventas de EV en Colombia), lo cual derivará en las conclusiones del proyecto aplicado con las que se aspira contribuir a la comprensión del mercado de vehículos eléctricos en el país.

## Análisis Exploratorio de Datos

A partir de la base de datos proporcionada por Datos Abiertos sobre los vehículos eléctricos e híbridos en Colombia, se realizó el análisis exploratorio de los datos para determinar cuál es el estado general de los datos y como es la dinámica de este mercado.

### Figura 2

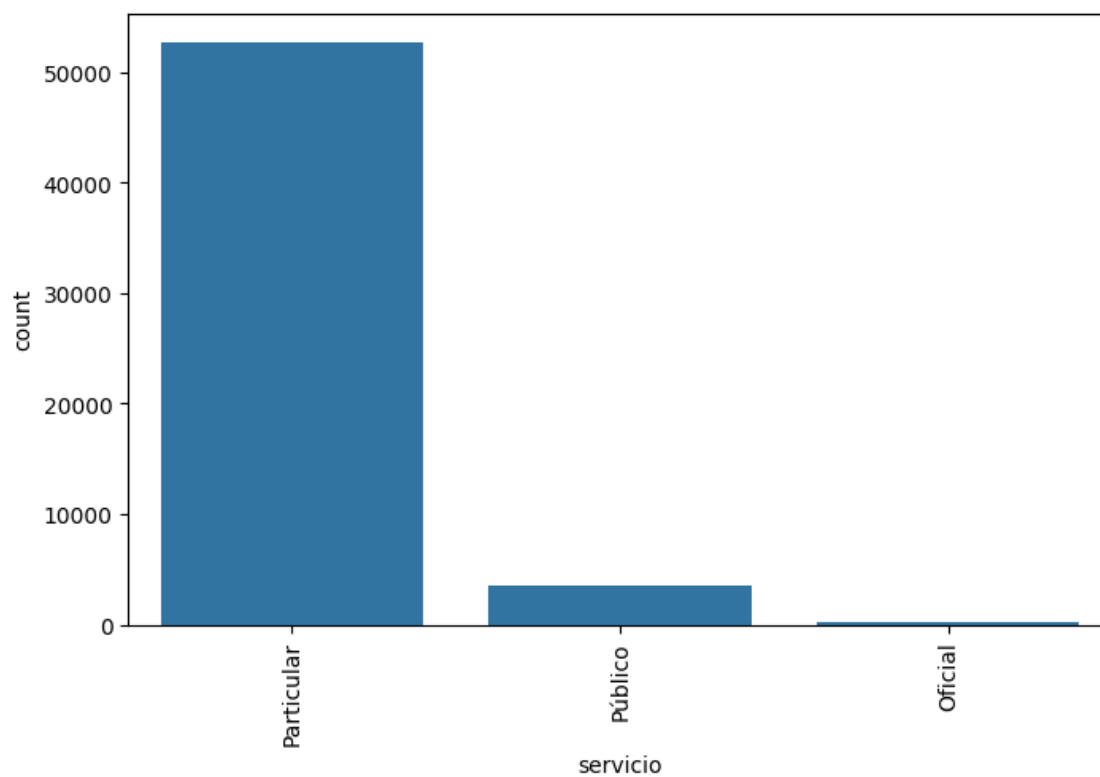
*Matriz de Completitud de Datos*



Podemos observar que el dataset contienen 21 características y un total de 56.545 instancias donde cada una representa un registro de un vehículo ya sea 100% eléctrico el cual se denominará vehículo eléctrico o EV; así como los vehículos híbridos los cuales son del tipo gasolina-eléctrico o Diesel-eléctrico. Dentro del dataset podemos evidenciar que las características de modalidad, capacidad\_pasajeros y capacidad contienen la menor cantidad de datos válidos, tan solo el 3% de los datos de capacidad\_pasajeros son no nulos.

**Figura 3**

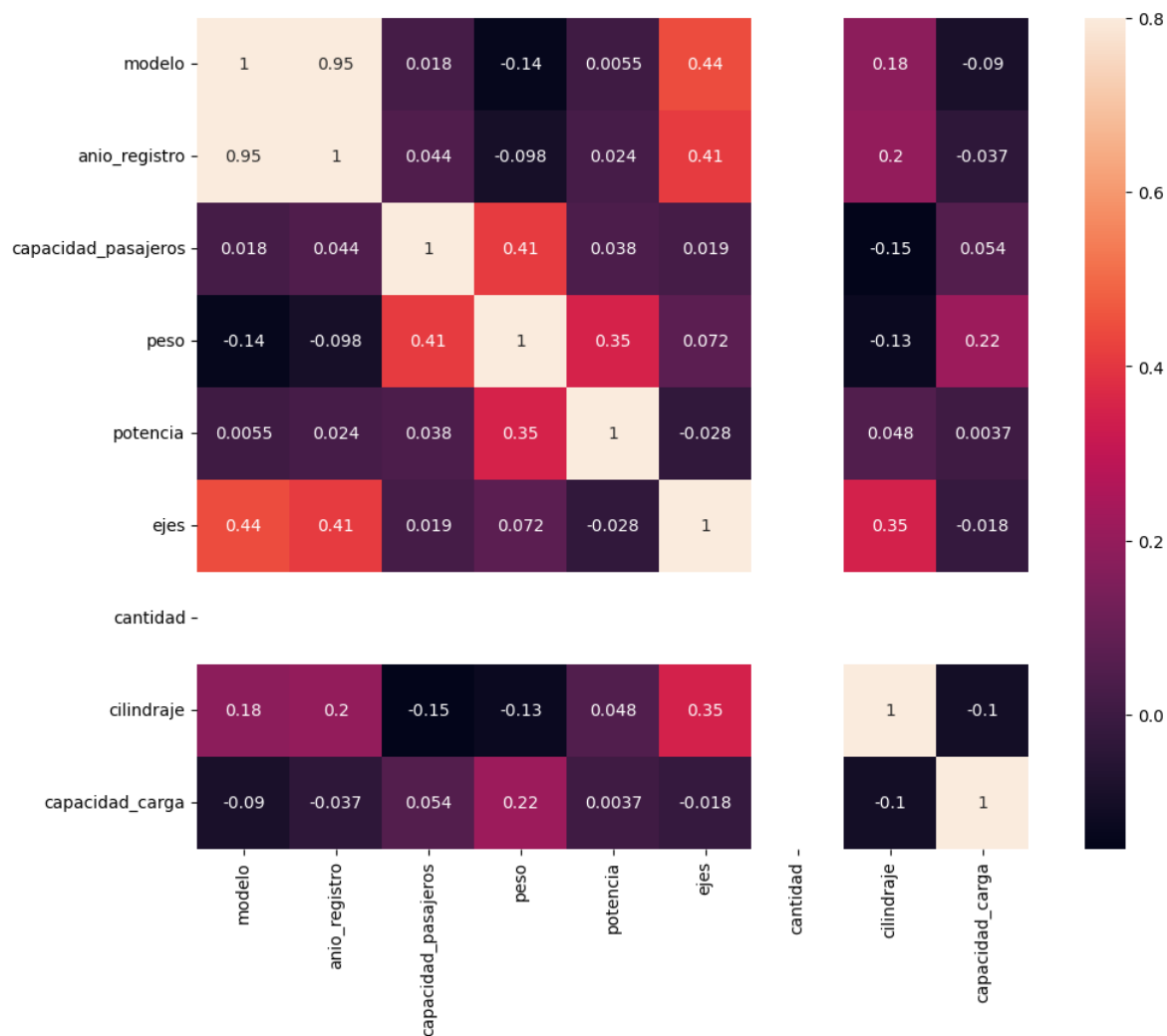
*Tipos de Clases de Vehículos Eléctricos a Nivel Nacional*



El dataset contiene datos de datos de vehículos particulares, públicos y oficiales, donde más del 90% de los registros son vehículos particulares.

**Figura 4**

*Matriz de Correlación entre Variables del Dataset*



Del mapa de correlación de las variables numéricas podemos encontrar las siguientes relaciones:

- Las variables de modelo y año de registro tienen una fuerte correlación (0.95), esto sugiere que los modelos más nuevos son vendidos con relativa rapidez haciendo que sean registrados en los años cercanos a su modelo.
- Las variables peso y capacidad de pasajeros tienen una relación moderada, esto indica que los vehículos para la movilidad de pasajeros tienden a ser más pesados.

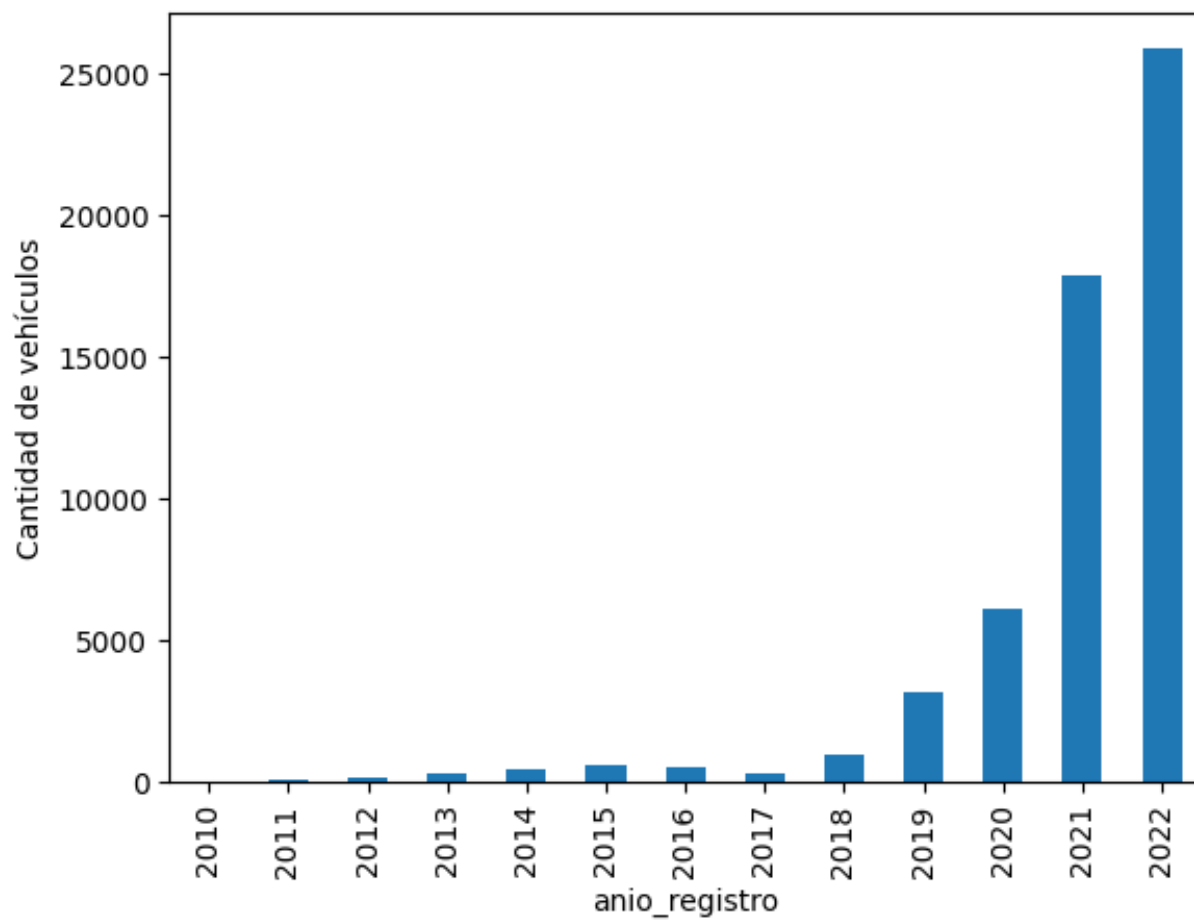
- De igual forma, se evidencia una relación moderada entre la potencia y el peso de los vehículos, indicando que los vehículos más pesados tienen motores que generan mayor potencia.
- La cantidad de ejes tiene relación con el modelo, año de registro y cilindraje, esto sugiere que los vehículos más nuevos tienden a tener mayor cantidad de ejes y a su vez mayor cilindraje.

Con base en el año de registro, podemos evidenciar un incremento significativo de vehículos eléctricos e híbridos a partir del año 2020. Para análisis posteriores se escoge los vehículos eléctricos debido a que estos presentan la menor penetración en el mercado colombiano, se sabe que tan solo el 16 % de los vehículos registrados son 100 % eléctricos, mientras que los híbridos son el 80 % del total.

Debido a la necesidad de entender mejor el comportamiento de los vehículos eléctricos en el mercado, los análisis se dividirán en dos secciones, una para el análisis por marcas donde se hará énfasis en las principales marcas que ofrecen vehículos 100 % eléctricos, y el segundo análisis para cantidad de EV por departamento.

**Figura 5**

*Cantidad de Vehículos Registrados por Año a Nivel Nacional*



## Tipos de Vehículos

A nivel nacional, encontramos que tanto los automóviles como las camionetas son los vehículos con mayor presencia.

**Tabla 1**

*Cantidad de Vehículos Según la Clase*

Clase	Cantidad	Porcentaje	Acumulado
Automóvil	2685	28.04	28.04
Camioneta	1946	20.32	48.37
Bus	1622	16.94	65.31
Motocicleta	1416	14.79	80.10
Cuadriciclo	1016	10.61	90.71
Camión	435	4.54	95.25
Motocarro	240	2.50	97.76
Tricimoto	188	1.96	99.72

Notamos que el 80 % de los vehículos eléctricos vendidos en Colombia están distribuidos en cuatro tipos de vehículos. La clase con mayor cantidad de unidades vendidas son los automóviles (sedanes), con 2,685 unidades, seguidos de las camionetas (SUV) con 1,946 unidades, los buses con 1,622 unidades y las motocicletas con 1,416 unidades. En el caso de los automóviles, se evidencia una leve tendencia positiva en la cantidad de unidades vendidas, con picos a finales de 2020 e inicios de 2022. Para los vehículos tipo camioneta, observamos que las ventas se incrementaron significativamente en el último trimestre de 2021, manteniendo una tendencia positiva a lo largo de 2022 y 2023.

En el caso de los buses, se observa que este tipo de vehículos ha tenido históricamente una baja cantidad de ventas. Sin embargo, hubo dos picos importantes de unidades vendidas en 2021 y 2022, cuando entraron en circulación más de 400 y 500 unidades en diferentes trimestres, respectivamente, convirtiendo a este tipo de vehículos en el de mayor número de unidades vendidas por trimestre.

**Figura 6**

*Venta de EV en los Principales Departamentos de Colombia*

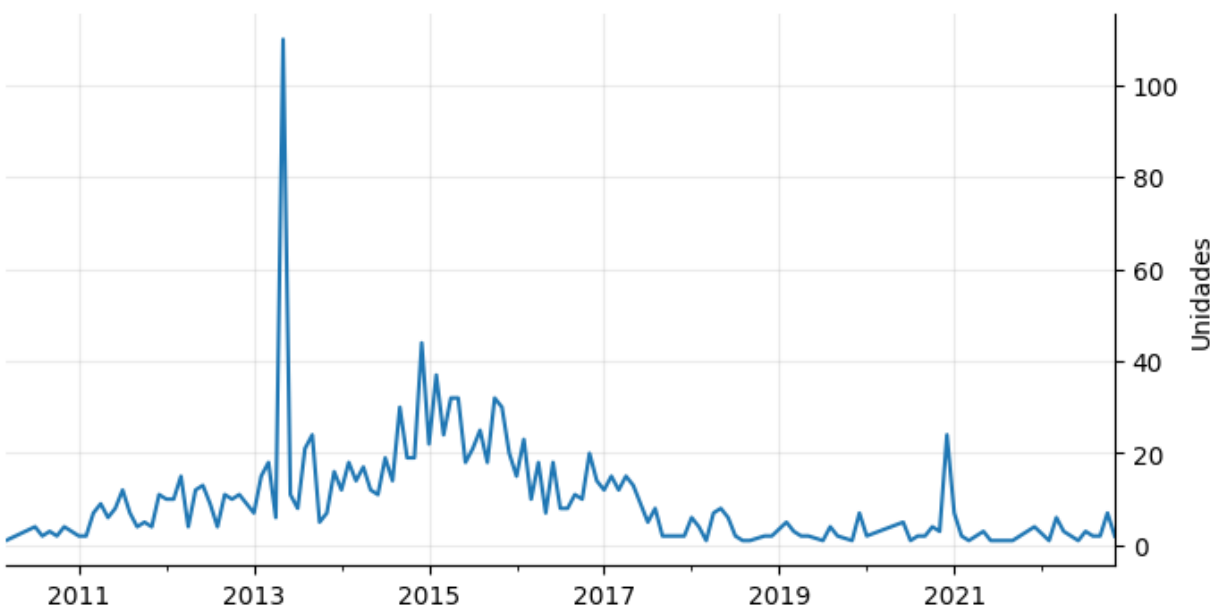


Podemos notar que los picos observados en la cantidad de buses en circulación coinciden con los de la venta de vehículos eléctricos en Bogotá. Esto se debe a que en 2021 el Ministerio de Transporte realizó la compra de 1,589 buses eléctricos, los cuales se distribuirían entre Cali, Medellín y Bogotá. Esta medida fue implementada por el gobierno nacional para fomentar la transición energética. Para esta iniciativa, el gobierno integró inicialmente 483 buses al sistema de transporte público de Bogotá; estos comenzaron a circular en mayo de 2021, y posteriormente, en 2022, se entregaron el resto de los buses eléctricos (Ministerio de transporte, 2021).

Respecto a las motocicletas eléctricas, observamos que presentan una tendencia negativa (**Figura 7**), es decir, la cantidad de unidades vendidas disminuye a medida que avanza el tiempo. El periodo en que más motocicletas se vendieron fue entre 2013 y 2016. Durante este tiempo, encontramos que en 2013 entraron en circulación alrededor de 100 motocicletas eléctricas, debido a un contrato realizado por la Alcaldía de Bogotá para la adquisición de motocicletas eléctricas para la Policía Metropolitana (Luis et al., 2012).

### Figura 7

*Motocicletas Eléctricas Registradas en Colombia*



Otro pico en la cantidad de vehículos eléctricos vendidos en Bogotá ocurrió a finales de 2022, en este caso asociado a la venta de vehículos tipo camioneta eléctrica. En este segmento, la marca BYD lideró el mercado, gracias a la popularidad que obtuvo después de ganar la licitación de buses eléctricos y porque fue una de las primeras marcas en ofrecer vehículos tipo SUV eléctricos en Colombia a precios relativamente bajos frente a la competencia.

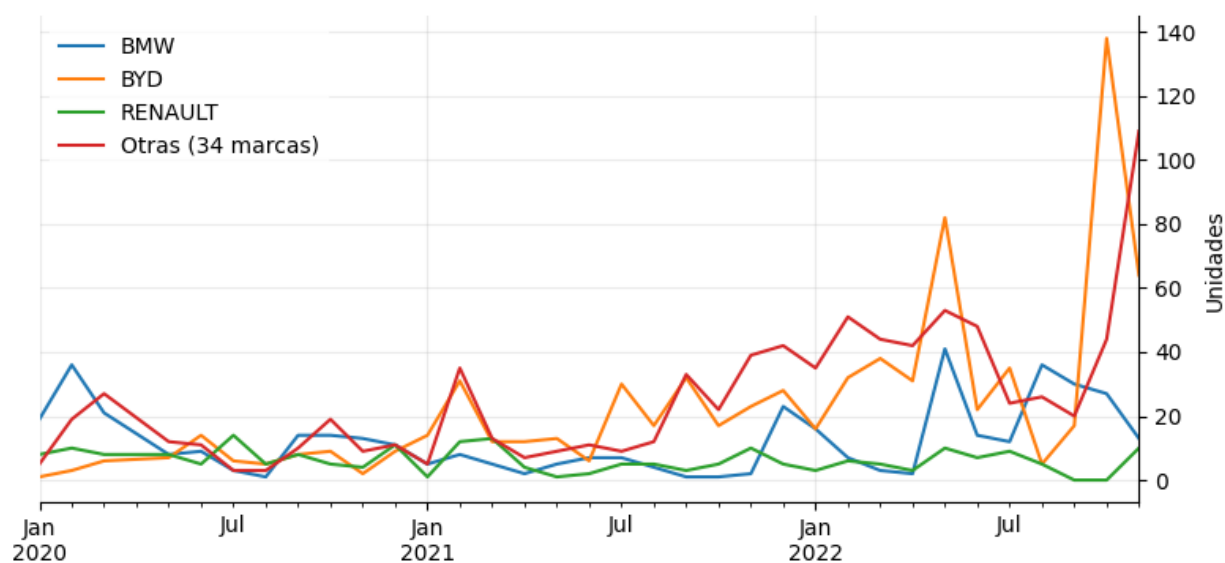
Para el desarrollo de los modelos de predicción de ventas de vehículos eléctricos, el proyecto tomará dos secciones, una para el análisis para las principales marcas de vehículos eléctricos en el país y otra sección para el análisis de venta de vehículos eléctricos por departamento.

## Marcas

Con base en los registros de ventas de EV en Colombia de 2020 a 2022 consolidados por Mintransporte y tomando solo las ventas de automóviles y camionetas de uso particular se generarán los pronósticos basados en el algoritmo de suavización exponencial (SE), con el objetivo de determinar su ajuste respecto a los datos de base.

### Figura 8

*Ventas de las Principales marcas de EV en Colombia*



Durante la última década, el mercado del EV en el país es liderado por tres compañías que comprenden el 62 % de las ventas registradas acumuladas: Renault, BMW y BYD.

**Tabla 2***Ventas por marca y Tipo de Vehículo*

Marca	Automóviles	Camionetas	Total
Renault	179	31	210
BYD	256	529	785
BMW	239	181	420
Otras (34 marcas)	363	499	862

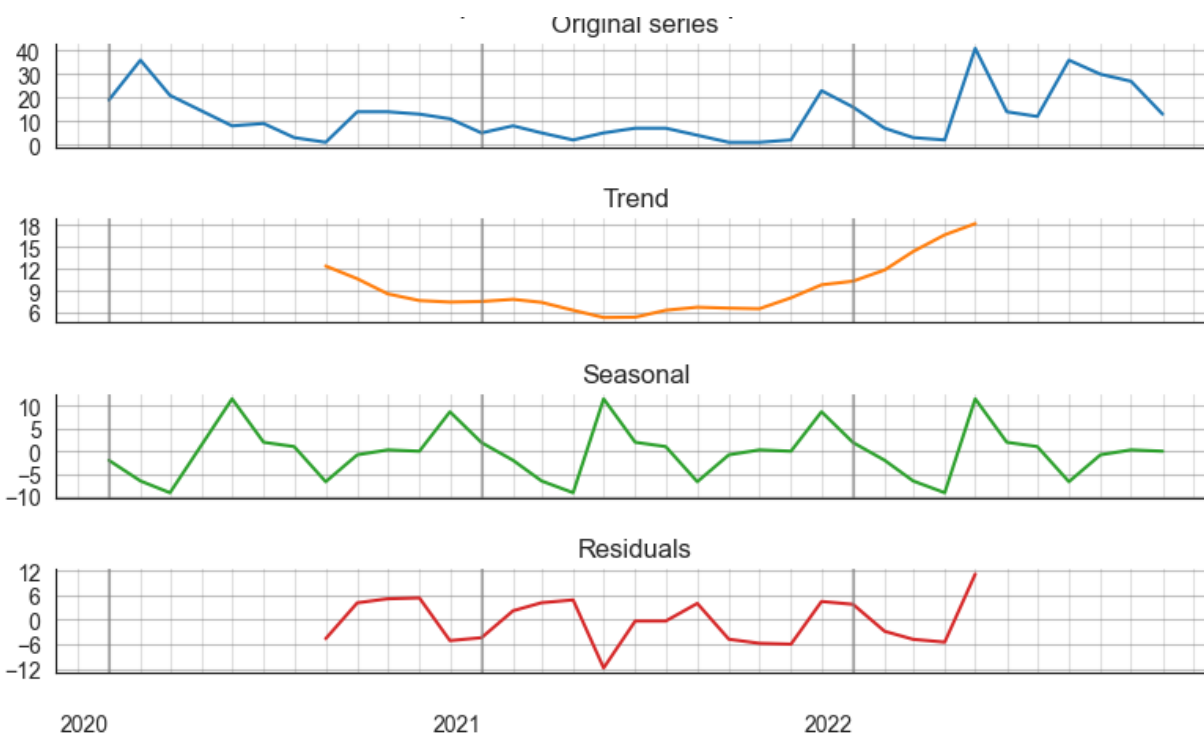
**BMW**

El fabricante alemán BMW se consolidó como líder del mercado de EV, principalmente en la categoría automóviles, destacándose por una oferta de vehículos con excelentes prestaciones, autonomía que supera los 400 km y acabados de lujo, que lo ubican en la gama alta del mercado y enfocado en un nicho muy bien definido (Quintero Abril, 2020).

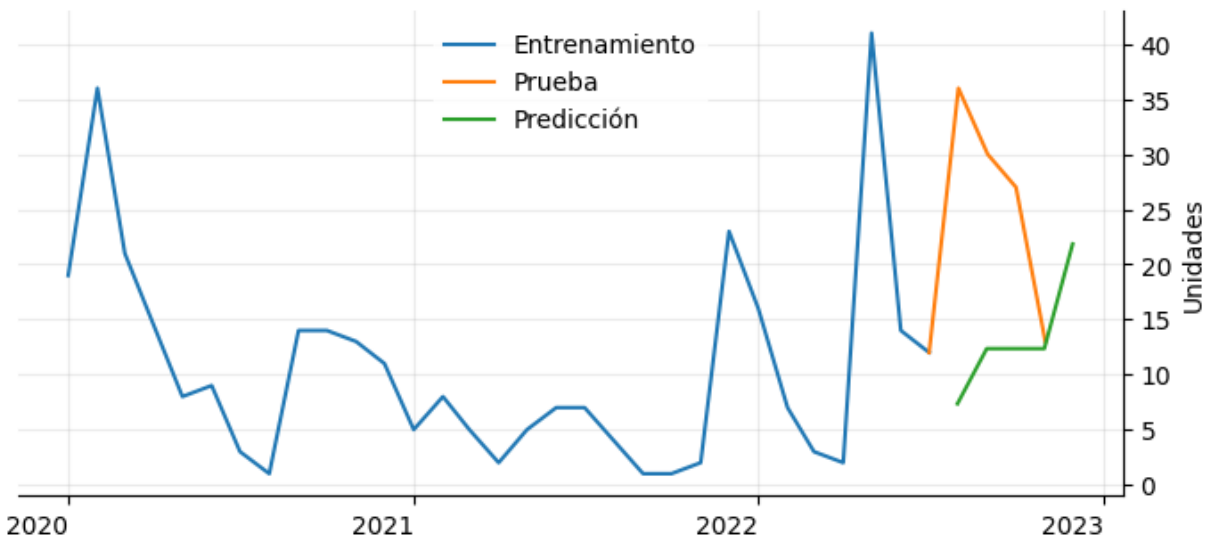
La serie de tiempo que describe las ventas de la marca exhibe dos aumentos importantes: 2019 y 2022. El primero deriva del compromiso con el cuidado medioambiental en los consumidores, como resultado de las campañas enfocadas en la mitigación de los efectos del cambio climático, el segundo se corresponde con la reactivación económica tras la declaración de pandemia en 2020 que trajo consigo un aumento en la demanda global de EV (KrivEVki, 2022).

**Figura 9**

*Descomposición de la Serie de Tiempo para BMW*



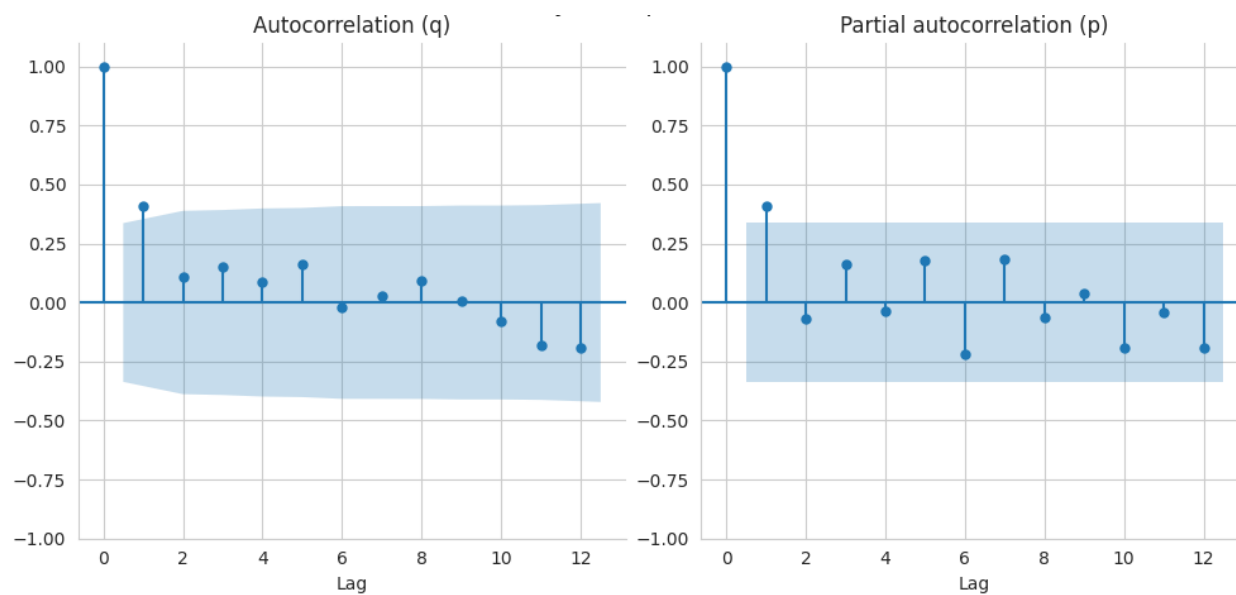
El modelado para las ventas de la marca se realizó a partir de la serie de tiempo original mediante un modelo de SE con valor  $R^2$  de -11.7056, reflejando una disminución en las ventas mensuales de EV.

**Figura 10***Modelo SE para BMW*

El modelo de suavización exponencial aplicado a la serie de tiempo de BMW arroja las siguientes métricas:  $R^2$ : -1.61, MSE: 237.02, RMSE: 15.4, MAE: 13.89.

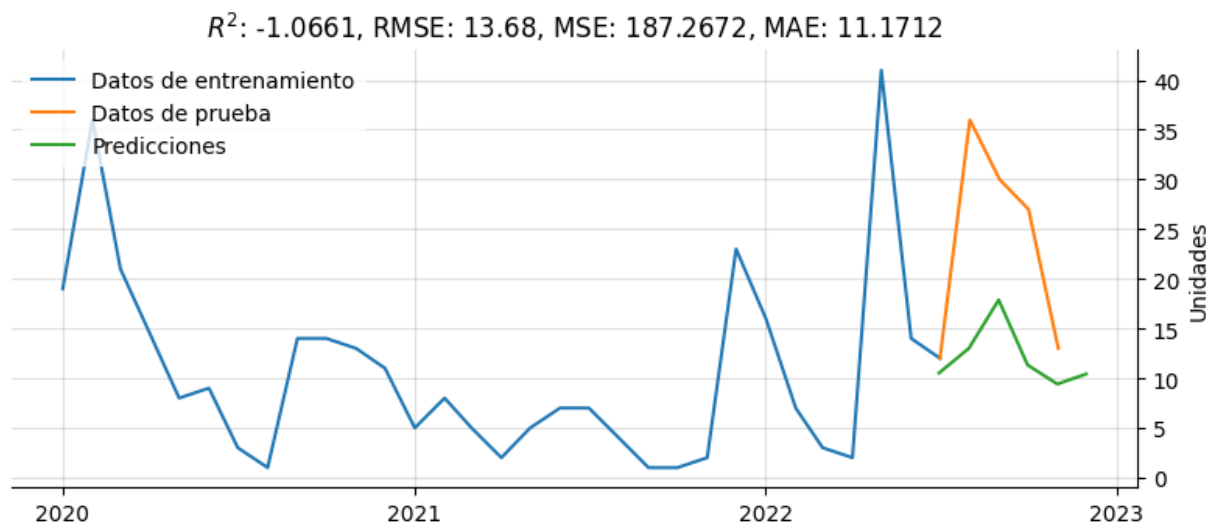
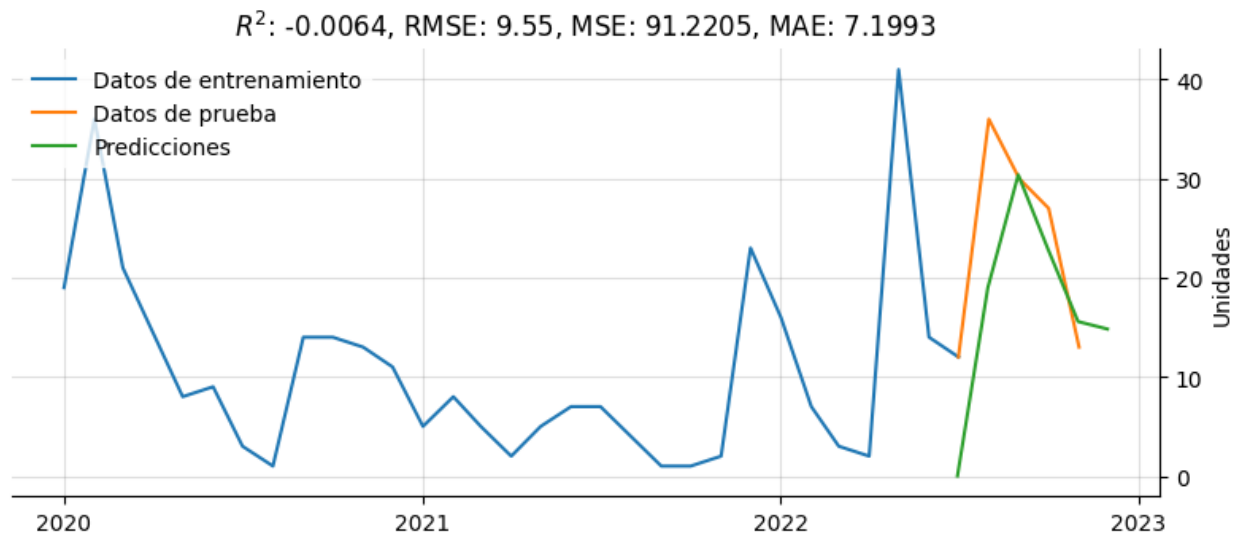
**Figura 11**

*Correlaciones sobre la Serie Estacionaria (Original) para BMW*

**Figura 12**

*Optimización de Parámetros del Modelo AR para BMW*

<b>p</b>	<b>d</b>	<b>q</b>	<b>AIC</b>	<b>R2</b>
<b>4</b>	<b>1</b>	<b>0</b>	<b>256.02</b>	<b>26.33</b>
<b>1</b>	<b>0</b>	<b>0</b>	<b>257.38</b>	<b>18.40</b>
<b>2</b>	<b>0</b>	<b>0</b>	<b>257.53</b>	<b>22.93</b>
<b>5</b>	<b>1</b>	<b>0</b>	<b>257.67</b>	<b>27.20</b>
<b>3</b>	<b>1</b>	<b>0</b>	<b>258.23</b>	<b>16.72</b>

**Figura 13***Modelo ARMA para BMW (1, 0, 1)***Figura 14***Modelo ARIMA para BMW (1, 1, 1)*

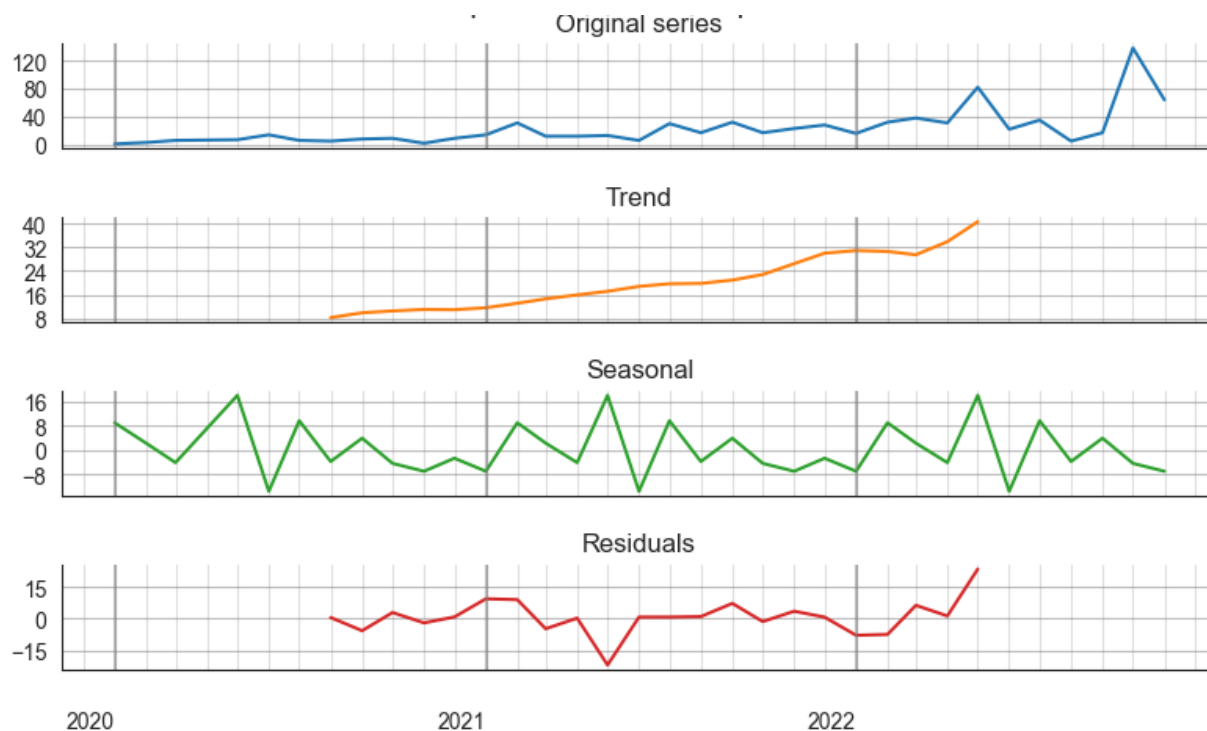
## BYD

BYD es el fabricante de EV más grande de China y se consolida como un competidor fuerte, con presencia global y apalancado por los beneficios arancelarios que la mayoría de los países otorgan a los EV. BYD. Además, ha establecido alianzas con las generadoras Enel y Celsia, para la venta e instalación de cargadores domiciliarios para EV, lo cual ha potenciado sus ventas en el centro del país (Ramos Robayo & Herrera Romero, 2024).

Esta marca es, además, la que lidera el mercado del transporte público eléctrico con una cuota de participación del 40 % a 2022 en Colombia y ofrece tecnologías eficientes en las baterías de los vehículos afrontando uno de los principales factores determinantes en la adquisición de EV: la autonomía (González, 2022).

### Figura 15

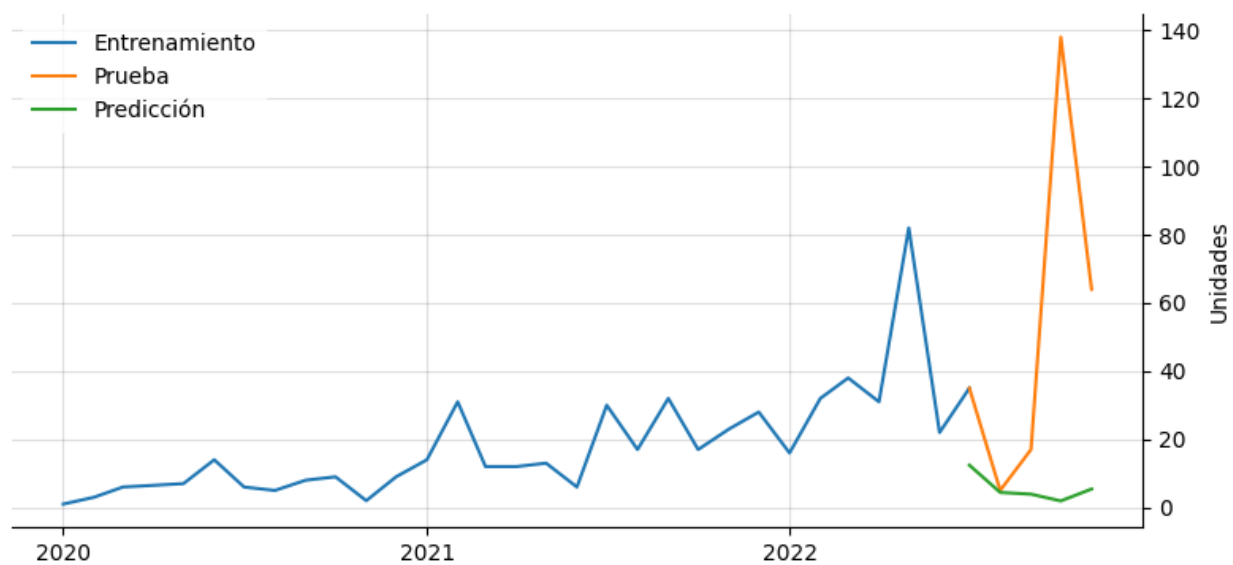
*Descomposición de la Serie de Tiempo para BYD*



Con un  $R^2$  de -1.0100, el modelo indica ventas ligeramente por encima de las 5 unidades mensuales.

### Figura 16

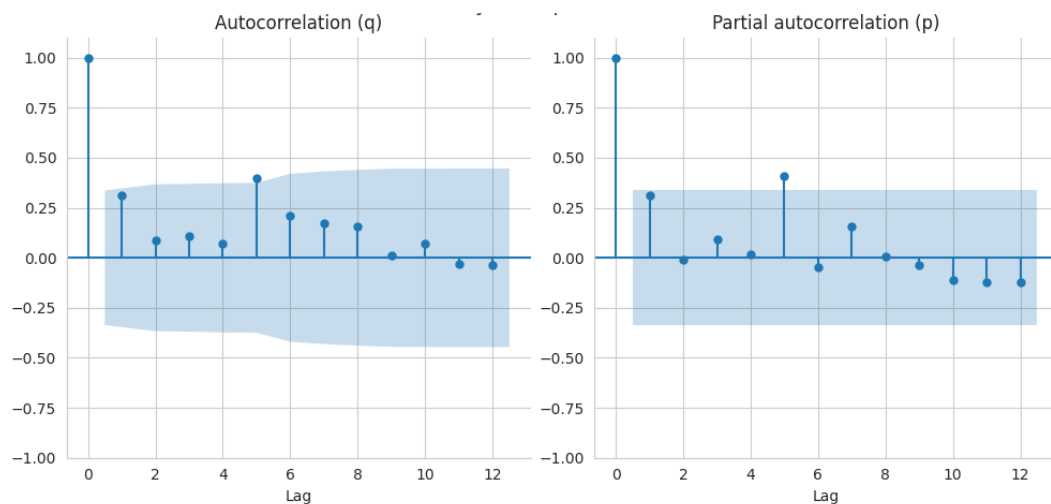
*Datos para BYD*



El modelo registró las siguientes métricas  $R^2$ : -1.61, MSE: 237.02, RMSE: 15.4, MAE: 13.89,

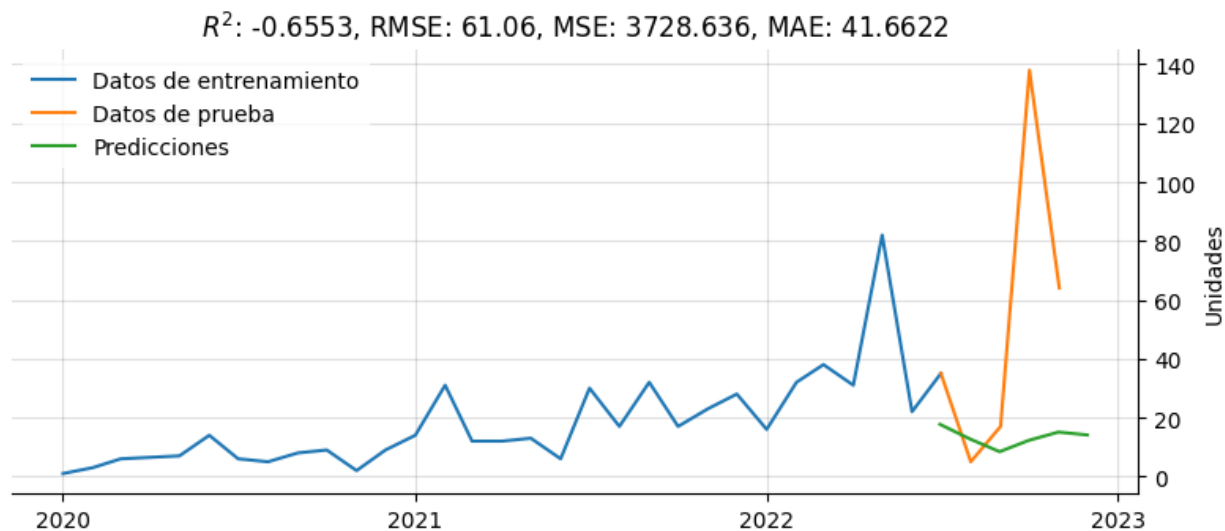
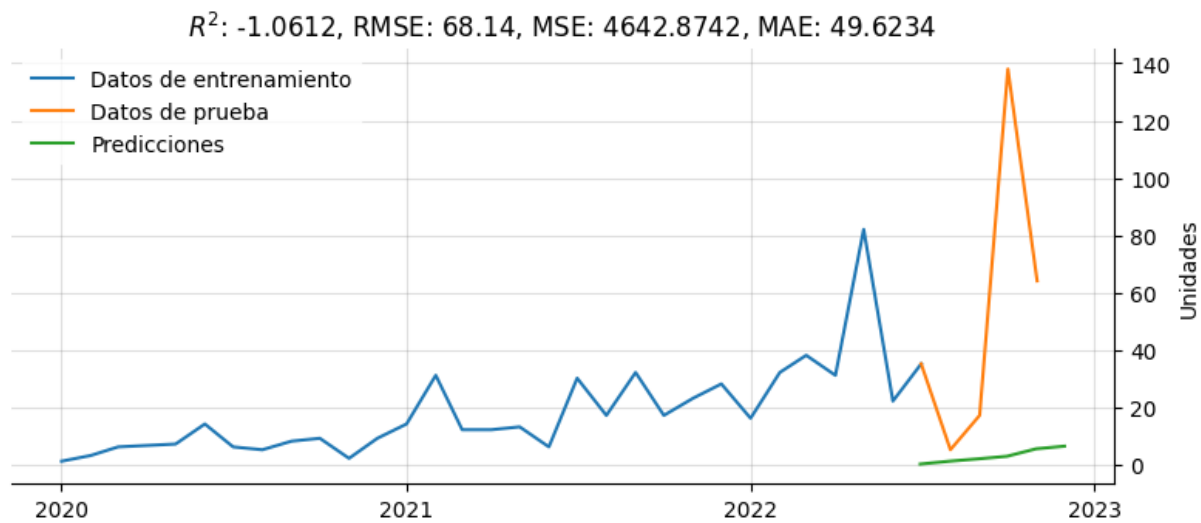
**Figura 17**

*Correlaciones Sobre la Serie no Diferenciada (Original) para BYD*

**Figura 18**

*Optimización de Parámetros del Modelo ARMA para BYD*

<b>p</b>	<b>d</b>	<b>q</b>	<b>AIC</b>	<b>R2</b>
<b>6</b>	<b>1</b>	<b>0</b>	<b>318.98</b>	<b>30.41</b>
<b>3</b>	<b>0</b>	<b>0</b>	<b>319.05</b>	<b>29.94</b>
<b>5</b>	<b>1</b>	<b>0</b>	<b>320.71</b>	<b>17.09</b>
<b>7</b>	<b>1</b>	<b>0</b>	<b>320.93</b>	<b>30.30</b>
<b>4</b>	<b>1</b>	<b>0</b>	<b>321.09</b>	<b>10.39</b>

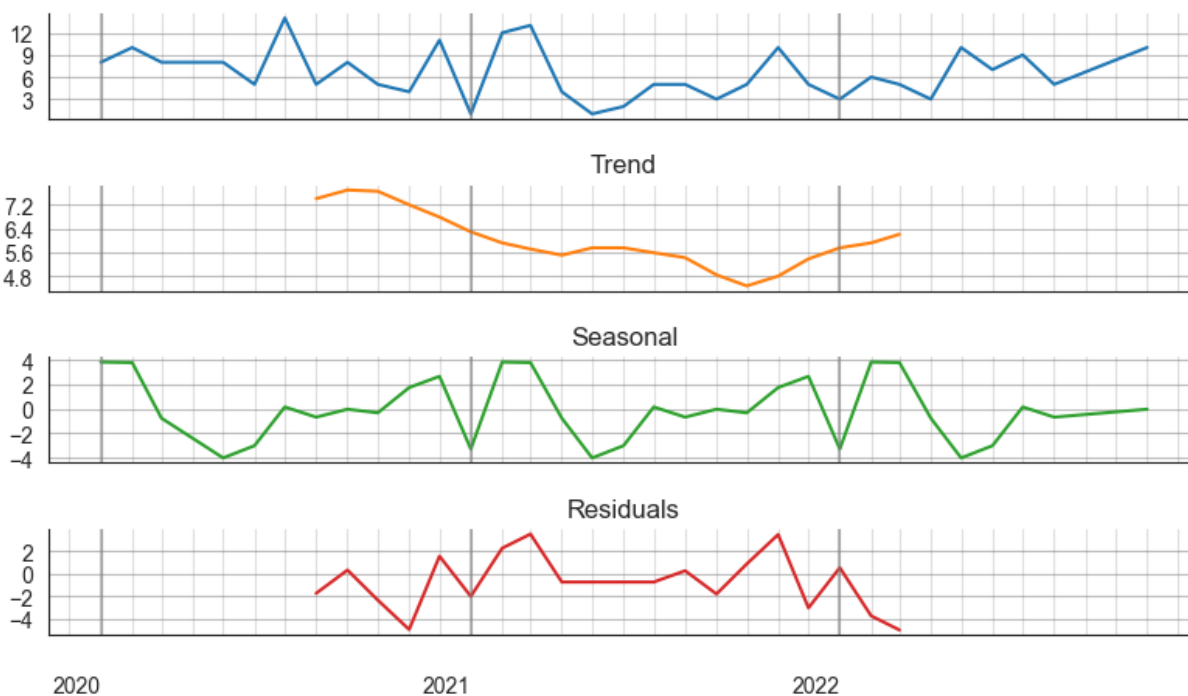
**Figura 19***Modelo ARMA para BYD (1, 0, 2)***Figura 20***Modelo ARIMA para BYD (4, 1, 5)*

## Renault

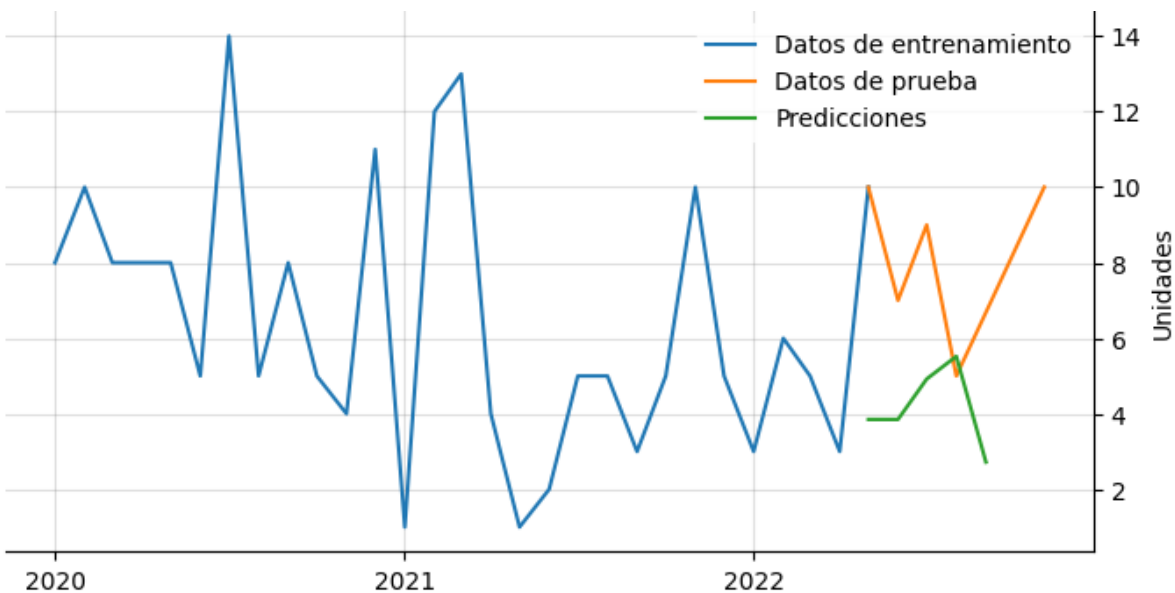
La marca francesa Renault es precursora del EV en Colombia, introduciendo modelos destacados por su precio accesible como el modelo Zoe, y que en los casos de su modelo Kangoo Z. E. correspondían a la versión eléctrica de sus modelos con motor a combustión.

### Figura 21

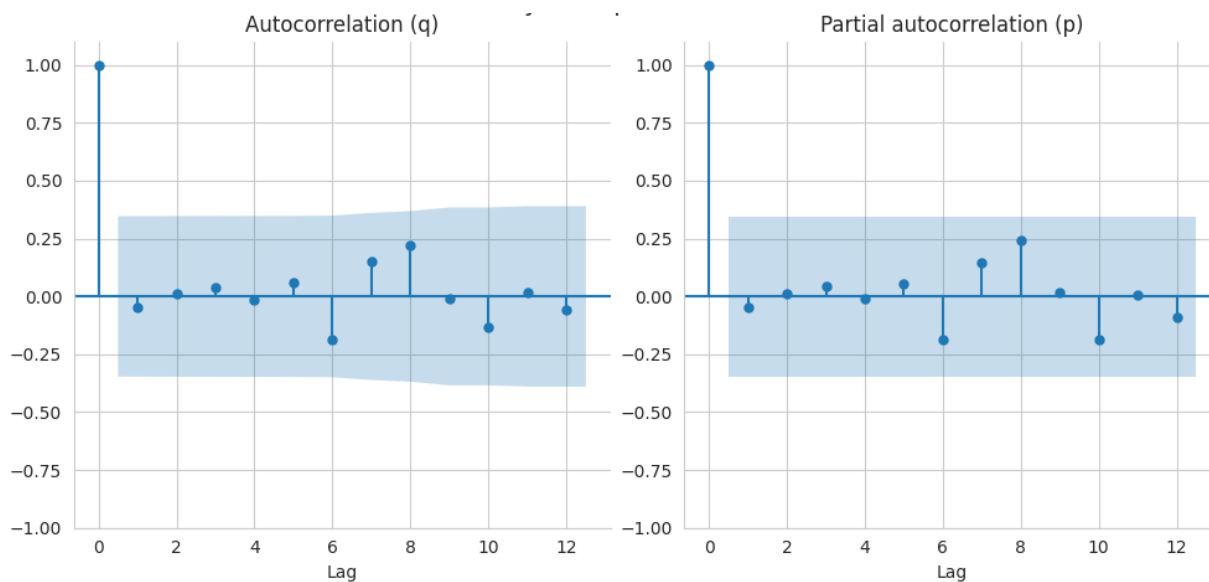
*Descomposición de la Serie de Tiempo para Renault*



El modelo aplicado a la serie (**Figura 22**) obtuvo las siguientes métricas  $R^2$ : -5.26, MSE: 23.53, RMSE: 4.85 y MAE: 4.24; pronosticando ventas que oscilan entre 3 y 5 unidades mensuales.

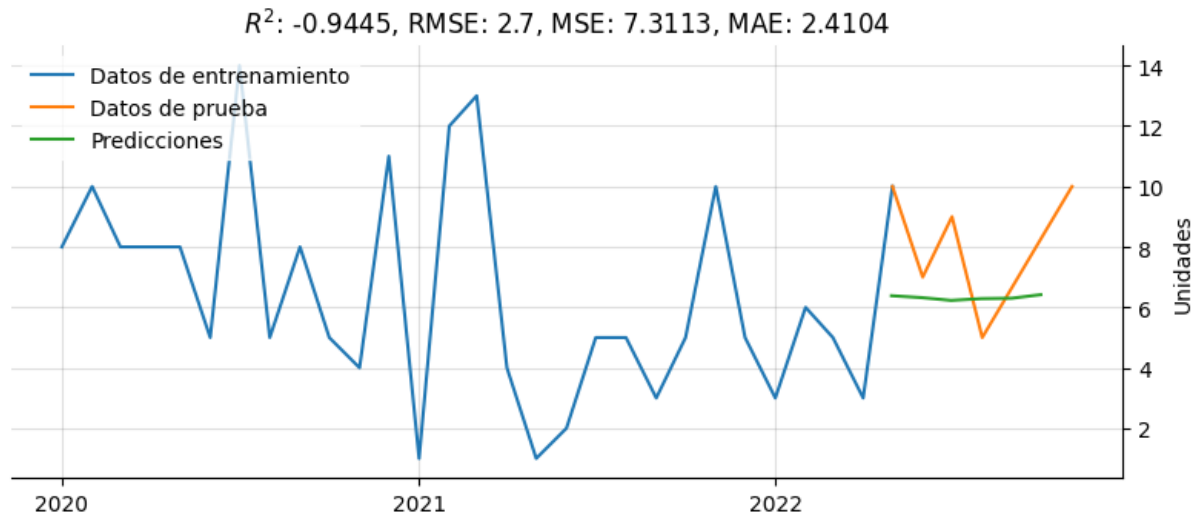
**Figura 22***Modelo SE para Renault*

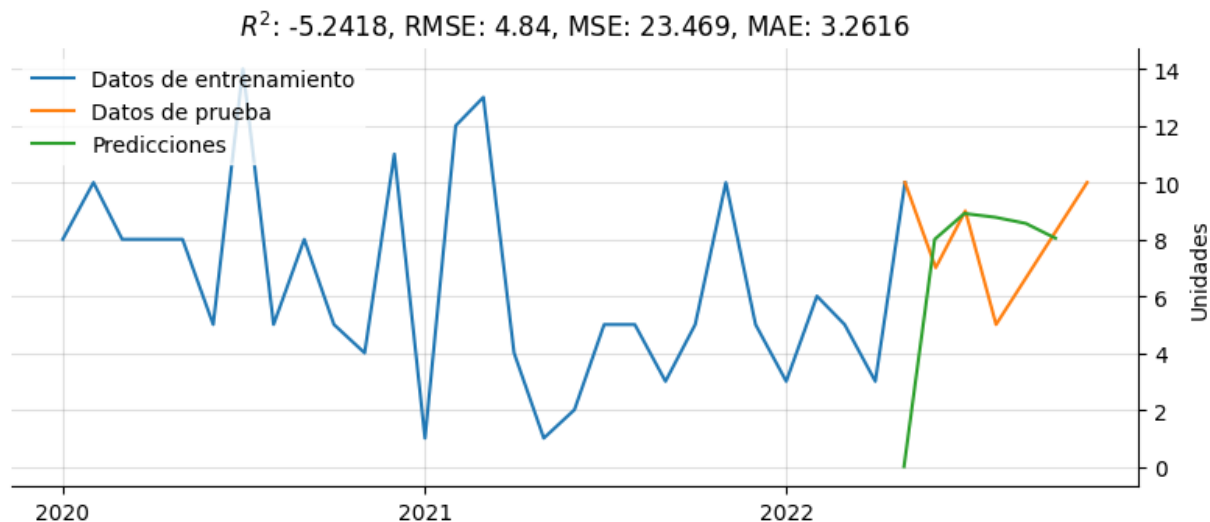
A continuación, se grafica la correlación y la correlación parcial para la serie de tiempo original, con el propósito de determinar la independencia entre los registros.

**Figura 23***Correlaciones de la Serie no Diferenciada (Original) para Renault*

**Figura 24**

*Modelo ARMA para Renault (1, 0, 1)*



**Figura 25***Modelo ARIMA para Renault (1, 1, 1)*

Los modelos preliminares de SE presentan bajo ajuste a los datos, con captación limitada de los fenómenos subyacentes en las series de tiempo, por lo que, tras generar predicciones con modelos autorregresivos (AR), se compara su desempeño observando la métrica  $R^2$  para cada tras su aplicación en las series de tiempo a cada una de las series de tiempo de las marcas.

Para la elección de los parámetros  $p$ ,  $d$  y  $q$  se probaron permutaciones de estas variables con valores asociados al resultado de las gráficas ACF y PACF de cada serie, eligiendo los valores de variables con menor valor AIC.

**Figura 26**

*Ejemplo de Permutación de Parámetros Obtenida para Modelos AR*

	<b>p</b>	<b>d</b>	<b>q</b>	<b>AIC</b>	<b>R2</b>
<b>5</b>	0	1	1	306.21	-8.72
<b>6</b>	0	1	2	307.68	-7.50
<b>13</b>	1	1	1	307.76	-7.68
<b>7</b>	0	1	3	309.27	-6.68
<b>21</b>	2	1	1	309.58	-7.35

En general, el desempeño de los modelos AR capta la variación en los datos mucho mejor que el modelo SE en todas las series de tiempo de las marcas analizadas. En particular, el modelo ARMA presentó mejor ajuste a los datos para las marcas BMW y Renault, el modelo ARIMA se ajustó mejor a la serie de tiempo de la marca BYD (**Tabla 3**).

**Tabla 3**

*Comparación de la Métrica R2 Obtenida con los Modelos Predictivos*

Marca	S. E.	ARMA	ARIMA
BMW	-11.7056	0.4602	-0.3609
BYD	-3.206	0.1886	0.6719
Renault	-0.8379	0.088	-0.4131

## Departamentos

Bogotá D. C. , Antioquia y Cundinamarca lideran la adopción del EV como medio de transporte alternativo de acuerdo con las cifras del Ministerio de Transporte (2022) así como también poseen los mayores números de vehículos registrados (**Figura 8**), superando ampliamente la media nacional por departamento que se ubica en 614.229 vehículos (Datos Abiertos Colombia, 2024) como se ve en la distribución departamental de la **Figura 27**.

**Tabla 4**

Vehículos Matriculados en Antioquia, Bogotá D. C. y el Resto del País

Departamento	Vehículos registrados
Antioquia	3.073.961
Bogotá D. C.	2.693.164
Cundinamarca	2.364.136
Otros	11.524.081

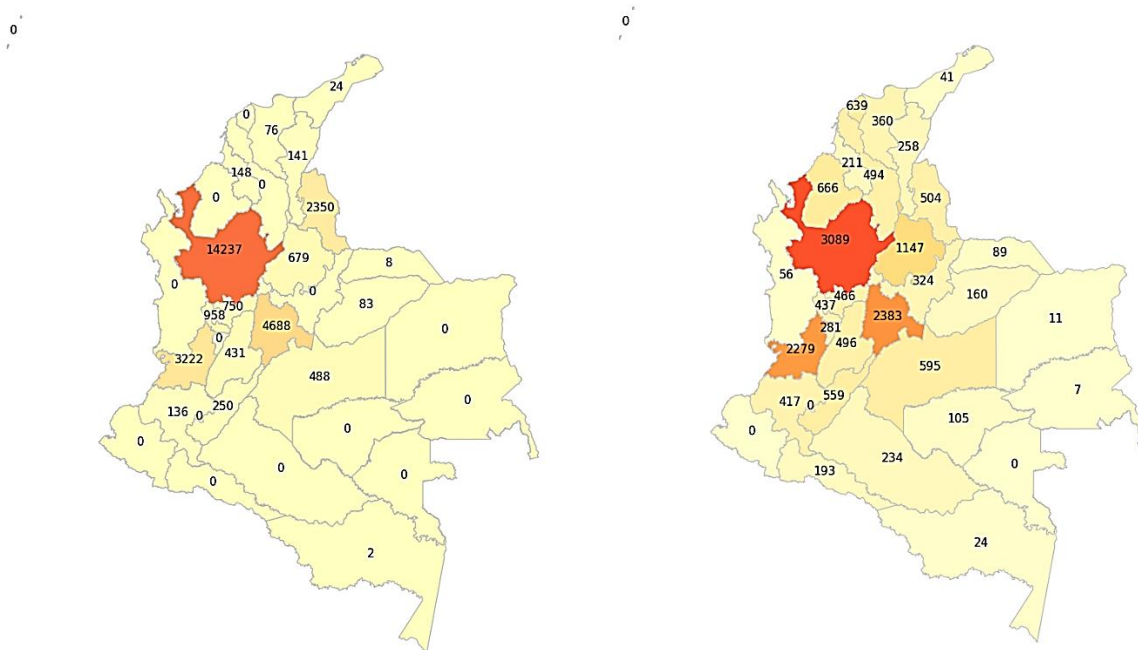
*Nota.* Tomado de Datos Abiertos Colombia, RUNT 2.0 (Datos Abiertos Colombia, 2024)

Esta realidad trae consigo, entre otras, afectaciones a la calidad del aire, generando altos niveles de contaminación. Esto ha llevado a los organismos gubernamentales a tomar decisiones como restringir la movilidad de vehículos de combustión interna según el número de la placa (Ministerio de Ambiente y Desarrollo Sostenible, 2018). Además, se han adoptado medidas como la reducción de la tarifa por kilovatio en las estaciones de carga o descuentos en el pago de impuestos para los vehículos eléctricos (EV).

**Tabla 5***Cantidad de Vehículos Eléctricos e Híbridos por Departamento*

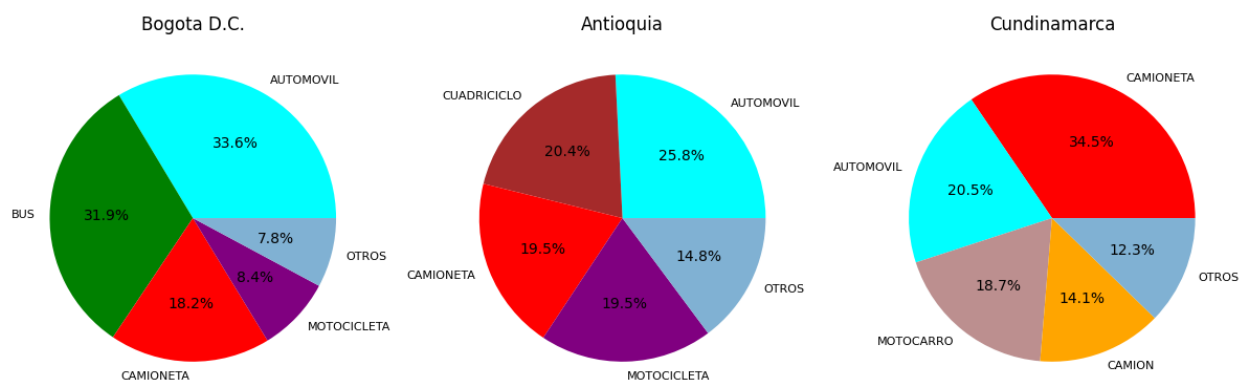
Departamento	Cantidad	Porcentaje	Acumulado
Bogotá D.C.	25507	45.109205	45.109205
Antioquia	14237	25.178177	70.287382
Cundinamarca	4688	8.290742	78.578124
Valle del Cauca	3222	5.698117	84.276240
Norte de Santander	2350	4.155982	88.432222
Risaralda	958	1.694226	90.126448
Atlántico	846	1.496154	91.622601
Caldas	750	1.326377	92.948979
Santander	679	1.200814	94.149792

En la **Tabla 5** observamos como está distribuida la cantidad de vehículos eléctricos en el país. Departamentos como Bogotá, Antioquia y Cundinamarca tienen el 78,5 % del total de los EV del país. Tan solo Bogotá tiene 25.507 vehículos eléctricos dentro de los cuales encontramos la clase particular y pública, y a diferencia de lo presentado con los vehículos de combustión interna en la capital del país, los vehículos eléctricos en Bogotá son casi la mitad del total de EV en Colombia.

**Figura 27***Distribución de Vehículos por Departamento*

*Nota.* Registros de EV acumulados por departamento de 2010 a 2022 (Izquierda), vehículos registrados en el RUNT desde noviembre al 20 de diciembre de 2024 (derecha) tomado de reporte DANE RUNT 2.0 (Datos Abiertos Colombia, 2024).

Al revisar en detalle las clases de vehículos eléctricos más vendidos en las ciudades principales, encontramos que los automóviles y las camionetas se mantienen entre los vehículos más vendidos. Sin embargo, dependiendo del departamento, aparece una tercera clase. Por ejemplo, en Bogotá, los buses ocupan esta posición; en Antioquia, los cuadríciclos son el segundo tipo de transporte con mayor cantidad de unidades en circulación, lo que se debe a la gran popularidad del Renault Twizy en el departamento antioqueño. Finalmente, en Cundinamarca, el tercer vehículo con mayor participación son los motocarros.

**Figura 28***Distribución de Vehículos Eléctricos por Departamento*

De acuerdo con la Unidad de Planeación Minero-Energética (UPME), Bogotá y Antioquia se encuentran entre los departamentos líderes en el desarrollo de infraestructura de recarga. Se han realizado inversiones públicas y privadas para ampliar la red de equipos de suministro de vehículos eléctricos, incluidos los cargadores rápidos, que son fundamentales para la adopción de vehículos eléctricos. La infraestructura es más sólida en estas regiones en comparación con otras, lo que facilita que los residentes posean y operen vehículos eléctricos.

Dado que la cantidad de vehículos de las clases buses y motocicletas se ha visto influenciada por factores externos al mercado de vehículos eléctricos, derivados de políticas gubernamentales que buscaban acelerar la transición hacia este tipo de tecnologías, según lo expuesto anteriormente, para el desarrollo de los modelos de predicción, se hará énfasis únicamente en los vehículos de las clases automóviles y camionetas, que representan el 48.3 % de los vehículos eléctricos.

**Figura 29**

*Ventas de EV en los Principales Departamentos de Colombia*

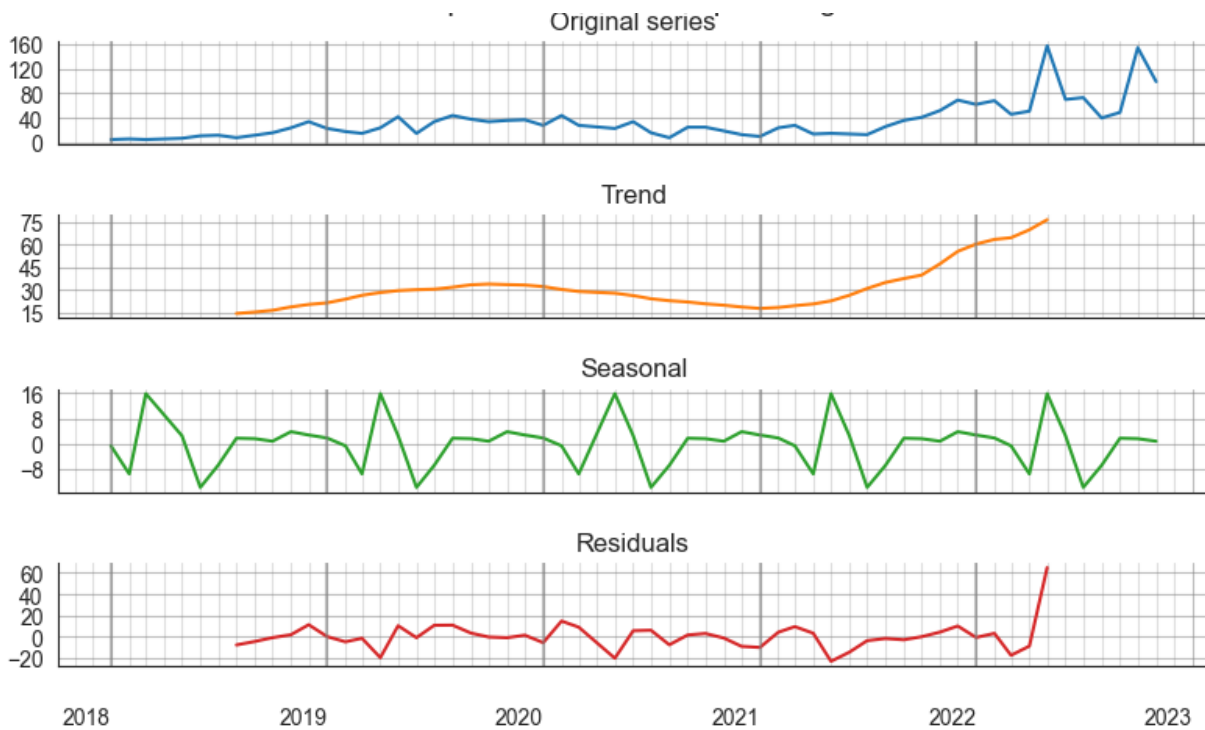


### **Bogotá**

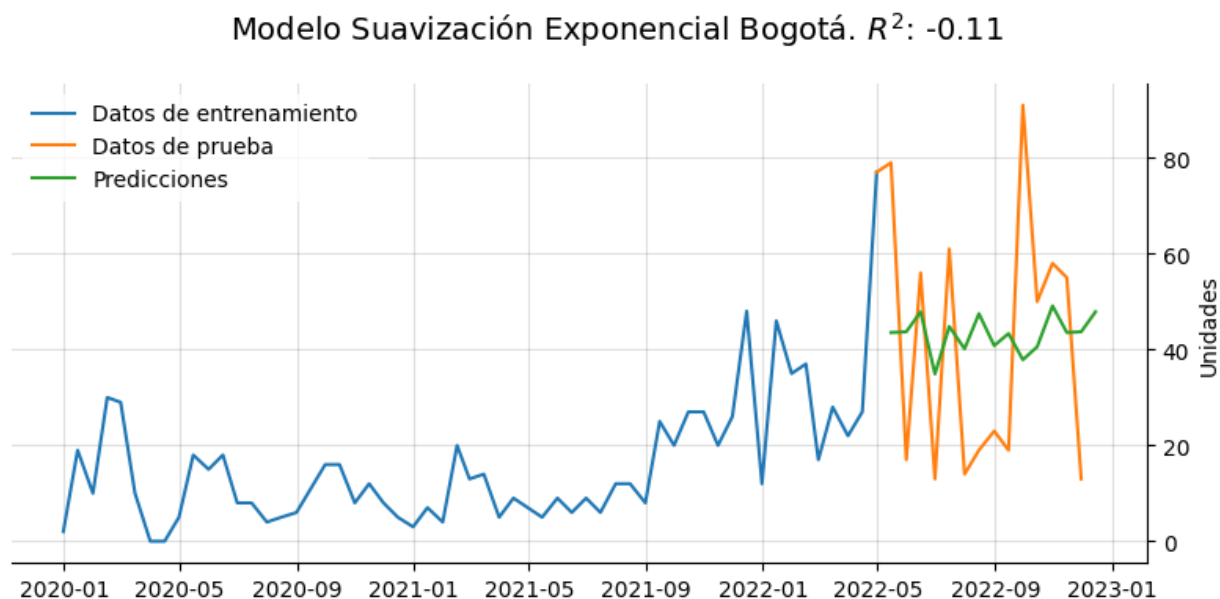
La serie con los datos del distrito capital exhibe un primer incremento en la tendencia en el año 2019, experimentando una contracción posterior a 2020 para posteriormente reactivarse en 2021, fenómeno que coincide con la recesión y reactivación de la economía nacional durante la declaración de pandemia de 2020 (**Figura 30**).

**Figura 30**

*Descomposición de la Serie de Tiempo para Bogotá D. C.*



El modelo aplicado obtuvo un  $R^2$  de -0.11, estimando volúmenes de ventas cercanos a las 40 unidades mensuales.

**Figura 31***Datos para Bogotá D. C.*

A continuación, se optimiza los parámetros para la aplicación de los modelos AR, MA, ARMA y ARIMA, se obtuvo los siguientes resultados:

**Tabla 6***Resultados de Optimización Modelos de Regresión Bogotá*

p	d	q	AIC	$R^2$
5	1	5	434.810642	50.75
5	1	4	434.808763	49.01
4	1	4	433.479859	47.36

Con base en el valor de  $R^2$ , observamos que el modelo ARIMA (5,1,5) presenta el mejor rendimiento en los datos de entrenamiento, alcanzando un  $R^2$  de 50.75 %.

**Tabla 7**

*Resultados de Optimización de Modelos de Regresión para Bogotá*

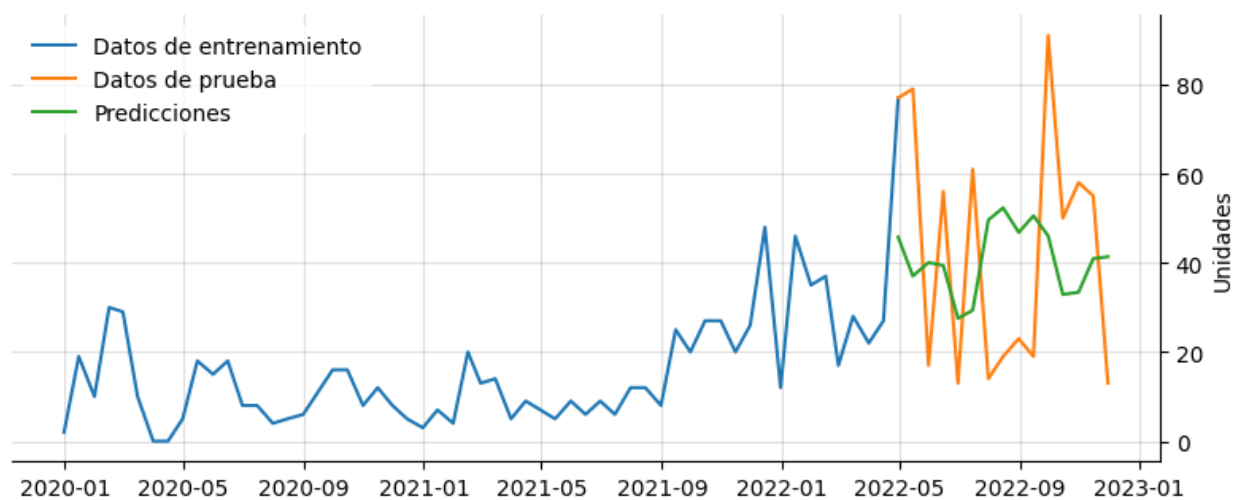
p	d	q	AIC	R2
0	2	4	428.684240	30.32
0	2	5	429.077812	32.64
0	1	5	429.820519	44.83

La tabla presentada muestra los resultados de la optimización organizados de menor a mayor valor de AIC. Observamos que el modelo de orden (0,2,4) presenta un AIC de 428.6. Dado que el valor de p es igual a cero, esto indica que el mejor modelo según el criterio AIC es un modelo MA con diferenciación de segundo orden. Este modelo obtuvo un coeficiente de determinación ( $R^2$ ) del 30.32 %, el cual es inferior al obtenido por el modelo ARIMA (5, 1, 5).

Asimismo, el segundo mejor modelo es un ARIMA (0, 2, 5), equivalente a un modelo MA con diferenciación de segundo orden, con un  $R^2$  de 32.6 %

**Figura 32**

*Datos para Bogotá D. C. (0, 2, 5)*



El modelo ARIMA muestra un rendimiento superior al de los modelos AR, MA y ARMA. Sin embargo, las predicciones aún están significativamente alejadas de los valores de prueba. El  $R^2$  obtenido fue de -0.2, lo que refleja un rendimiento muy bajo y evidencia que el modelo no es adecuado para la predicción de venta de EV en Bogotá D.C.

Para la implementación del modelo Prophet, se hace necesario configurar parámetros adicionales relacionados al efecto de los puntos de cambio, estacionalidad y eventos puntuales que pudieron afectar las ventas de estos vehículos.

En este sentido, se definió un periodo de COVID en el cual se configura como un regresor adicional del modelo de predicción las fechas establecidas fueron desde el primero de abril del 2020 hasta el 31 de diciembre del 2021. Los regresores permiten complementar el modelo subyacente del algoritmo, sumándose o multiplicándose a los otros componentes como la tendencia o estacionalidad (Facebook, 2024).

Al igual que en los modelos de regresión, para la implementación de los modelos de Prophet se realizó la optimización de los parámetros de prioridad de los puntos de cambio, prioridad de la escala de estacionalidad y modo de estacionalidad.

Al realizar el remuestreo de los datos a quincenales, se obtuvieron los siguientes valores óptimos.

**Tabla 8**

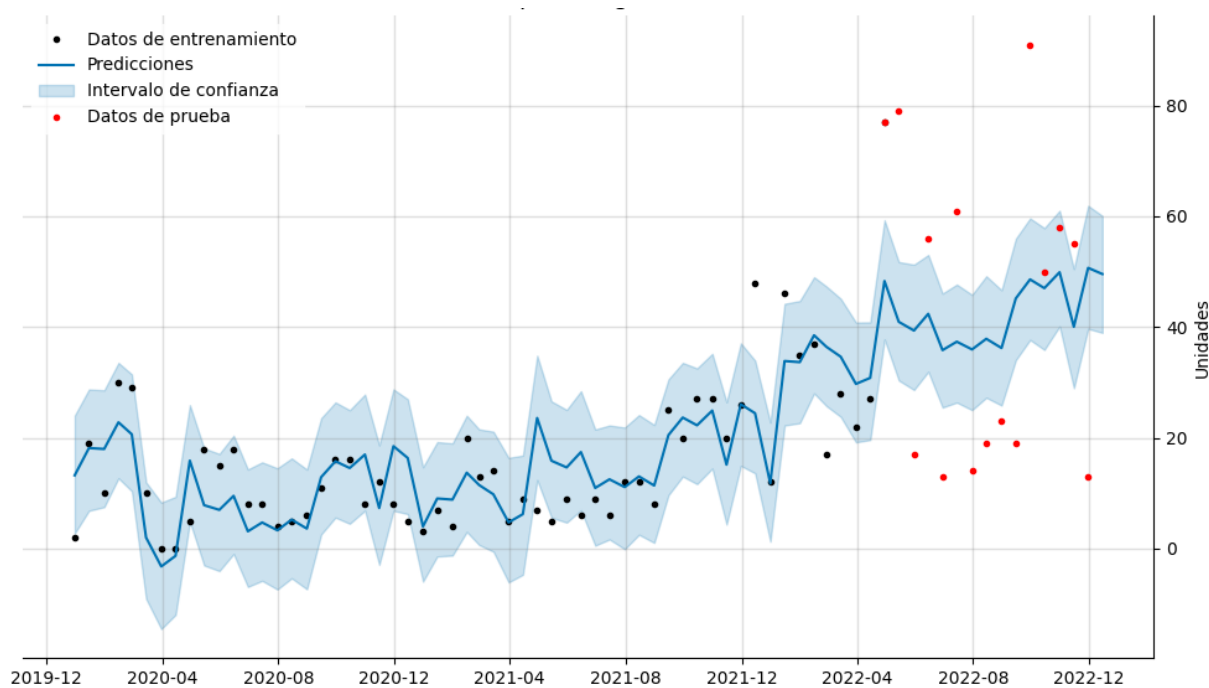
*Valores Óptimos de Prophet con Datos Quincenales*

changepoint_prior_scale	seasonality_prior_scale	seasonality_mode	RMSE
0.01	0.1	additive	13.719057
0.10	0.1	additive	13.859508
0.50	0.1	additive	14.179906

La tabla 7 muestra los resultados de la optimización de hiperparámetros del modelo Prophet, ordenados por RMSE (error cuadrático medio) de menor a mayor. El mejor desempeño (RMSE más bajo: 13.71) se logra con  $\text{change\_point\_prior\_scale} = 0.01$  y  $\text{seasonality\_prior\_scale} = 0.1$ , lo que indica que esta combinación permite un equilibrio adecuado entre la flexibilidad para capturar cambios en la tendencia y la regularización de la estacionalidad. Los valores más altos de  $\text{seasonality\_prior\_scale}$  (0.5 y 1.0) resultan en un rendimiento ligeramente inferior.

### Figura 33

*Resultados del Modelo Prophet para Bogotá ( $R^2: -0.04$ )*



El coeficiente de determinación  $R^2$  obtenido es negativo (-0.04). Esto indica que el modelo no está capturando correctamente la variabilidad de los datos, y sus predicciones son menos precisas que usar la media de los datos como referencia. Aunque el modelo está optimizado en términos de sus hiperparámetros, los resultados sugieren que Prophet, en este caso, tiene limitaciones para capturar la dinámica de las ventas de EV en Bogotá. Es necesario analizar los datos y considerar enfoques adicionales o complementarios para mejorar las predicciones.

**Tabla 9**

*Métricas de los Modelos*

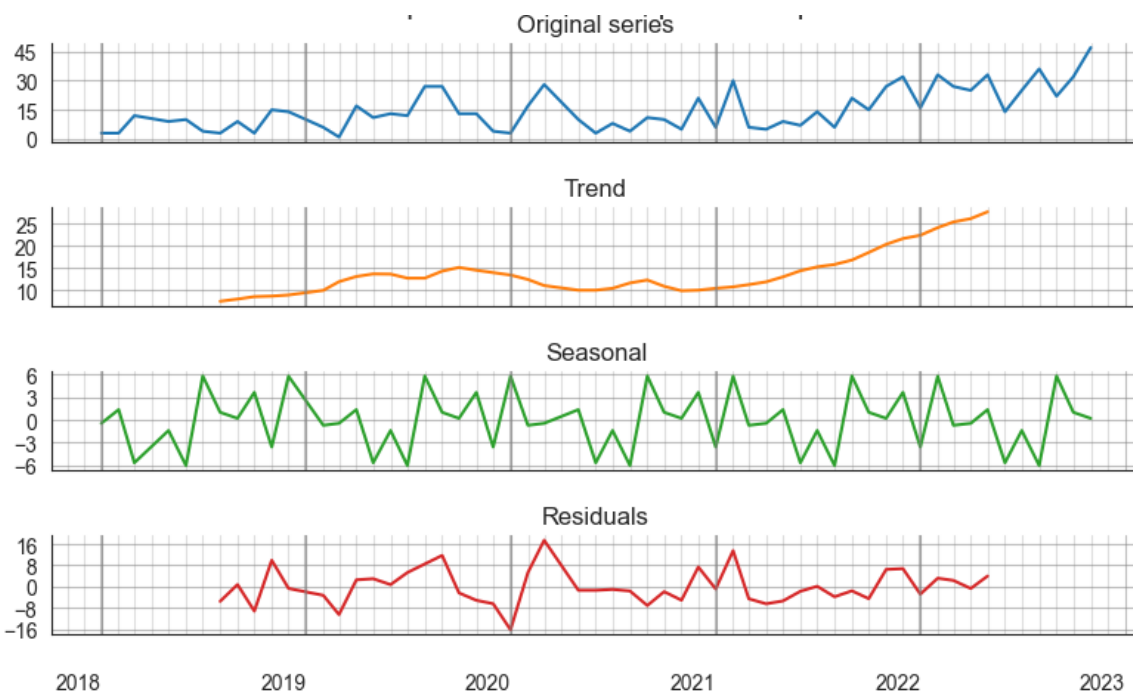
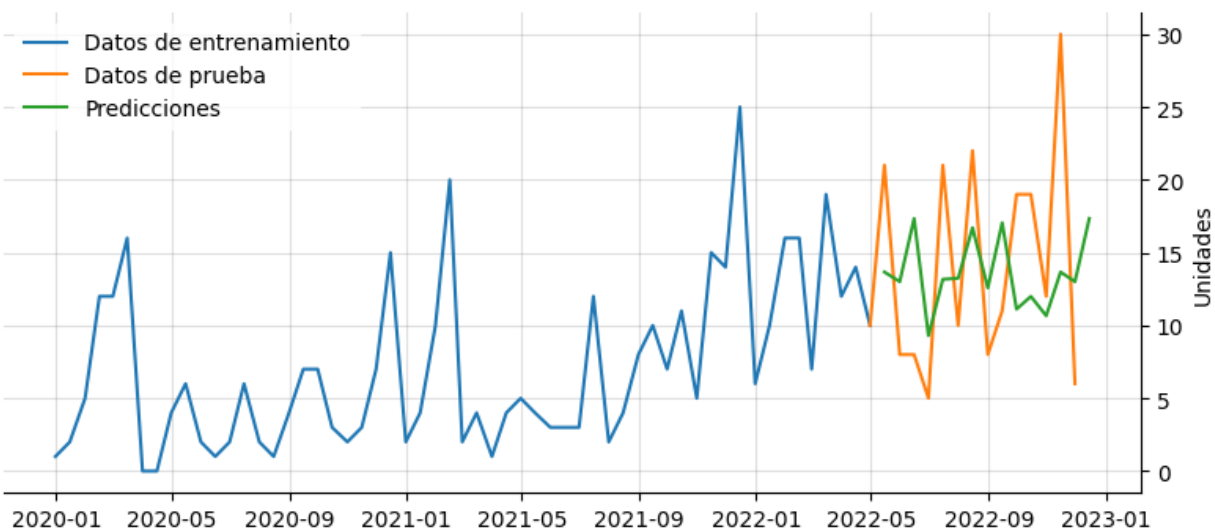
Métrica	Suavización exponencial	ARIMA	Prophet
$R^2$	-0.11	-0.2	-0.04
MSE	778.13	841.1	732.65
RMSE	27.89	29.0	27.07
MAE	25.32	27.49	24.43

La demanda de Bogotá presenta retos como la alta variabilidad de las ventas, además, factores externos como el gobierno o la entrada de competidores fuertes como BYD, generan fuertes cambios en la demanda. Los modelos anteriormente expuestos presentan en general un bajo rendimiento, al comparar los resultados del  $R^2$  notamos que en ningún caso este es mayor a cero, esto sugiere que las predicciones están lejos de los valores de prueba e incluso el promedio de los datos puede estar más cerca a la realidad.

En el caso de Bogotá, notamos que el modelo que mejor se comporta es Prophet, este modelo presenta métricas de error ligeramente inferiores a los resultados de la suavización exponencial y el modelo ARIMA.

## **Antioquia**

La serie de datos de Antioquia se transforma a estacionaria obteniendo un estadístico ADF de -1.9761, sobre la que a continuación se aplica SE, obteniendo una predicción de aproximadamente 45 EV mensuales en circulación adicionales, con un valor p de solo -1.9761.

**Figura 34***Descomposición de la Serie de Tiempo de Antioquia***Figura 35***Predicción del Modelo de SE para Antioquia ( $R^2: -0.37$ )*

Tras analizar los resultados obtenidos con los modelos de suavización exponencial, los cuales ofrecen una aproximación basada en patrones históricos de los datos, procedemos a explorar los modelos regresivos.

**Tabla 10**

*Resultados de Optimización de Modelos de Regresión para Antioquia ( $R^2$ )*

p	d	q	AIC	$R^2$
5	0	4	212.521622	40.85
4	0	4	211.038773	39.95
5	0	5	217.309728	37.46

Priorizando el valor de  $R^2$  calculado sobre los datos de entrenamiento, se determina que el modelo con mayor rendimiento es el ARIMA (5, 0, 4), equivalente a un modelo ARMA (5, 4). Este modelo obtuvo un valor de AIC de 212.52, que es 8.91 puntos superior al valor óptimo si priorizamos exclusivamente el AIC.

**Tabla 11**

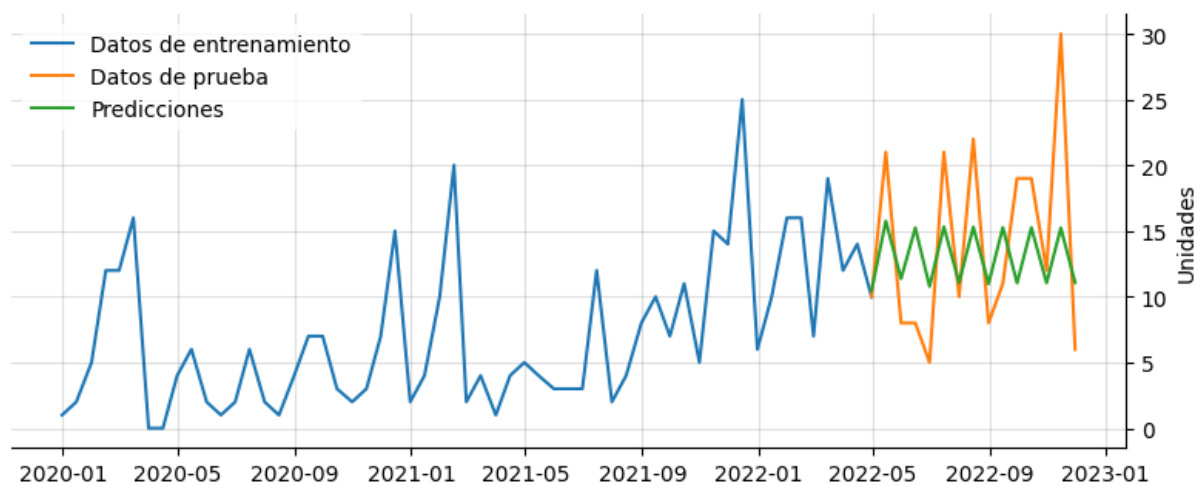
*Resultados de Optimización de Modelos de Regresión para Antioquia (AIC)*

p	d	q	AIC	$R^2$
3	1	2	345.059569	31.94
1	1	3	346.070595	28.53
5	1	3	346.203489	36.42

La tabla 10 muestra los tres modelos con menor AIC. Se observa que el mejor modelo según este criterio es el ARIMA (3, 1, 2), con un  $R^2$  de 31.94 %. Asimismo, el segundo mejor modelo es el ARIMA (0, 1, 1), aunque este modelo presenta un buen valor de AIC, su  $R^2$  es muy bajo, lo que indica un rendimiento deficiente sobre los datos de entrenamiento.

**Figura 36**

*Modelo ARIMA para Antioquia (3, 1, 2)*



El modelo ARIMA aplicado a los datos de venta de vehículos eléctricos en Antioquia muestra un rendimiento bajo. Se evaluaron dos configuraciones:

- Orden (3, 1, 2): Prioriza la métrica AIC y alcanza un  $R^2$  final de 28 %.
- Orden (5, 0, 4): Prioriza el  $R^2$  en los datos de entrenamiento, pero obtiene un  $R^2$  de  $-0.5$  en los datos de prueba, siendo inferior al modelo con el orden (3, 1, 2).

Para la implementación del modelo Prophet en el departamento de Antioquia se validó la posibilidad de cambiar el remuestreo de mensual a quincenal, a pesar de que este departamento es el segundo con mayor cantidad de EV registrados, se presentan muchos periodos en donde las ventas fueron cero, esto limita el remuestreo de los datos debido a la adición de ruido. Por lo tanto, se mantuvo el remuestreo mensual, manteniendo el periodo de COVID.

**Tabla 12**

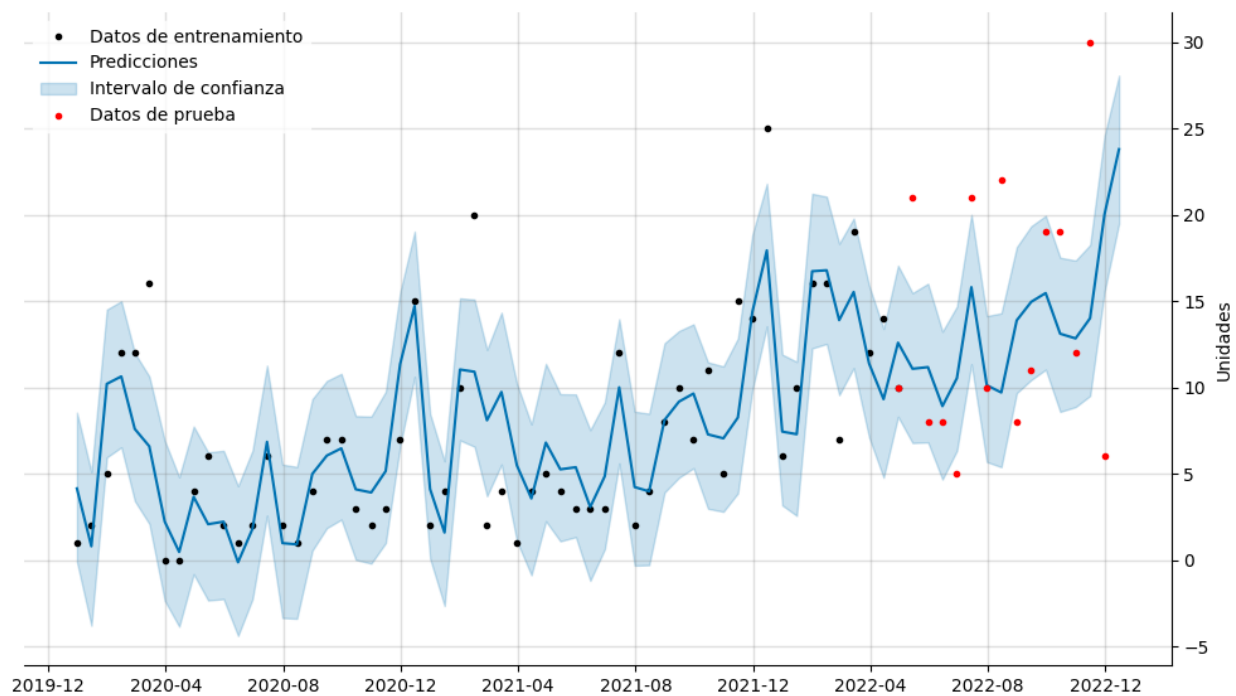
*Valores Óptimos para Prophet con Datos Mensuales de Antioquia*

Changepoint_prior_scale	seasonality_prior_scale	seasonality_mode	error
0.01	0.1	additive	6.105105
0.10	0.1	additive	6.106906
0.50	0.1	additive	6.353097

En este caso, el mejor desempeño (RMSE más bajo: 6.1) se obtiene con  $\text{changepoint\_prior\_scale} = 0.01$ ,  $\text{seasonality\_prior\_scale} = 0.1$  y  $\text{seasonality\_mode} = \text{additive}$ . Esto sugiere que el modelo funciona mejor con una mayor regularización en los cambios de tendencia ( $\text{changepoint\_prior\_scale}$  bajo) y con una representación aditiva de la estacionalidad.

**Figura 37**

*Resultados del Modelo Prophet para Antioquia ( $R^2: -0.26$ )*



**Tabla 13***Métricas de los Modelos para Antioquia*

Métrica	Suavización exponencial	ARIMA	Prophet
R <sup>2</sup>	-0.37	0.28	-0.26
MSE	69.86	36.83	64.46
RMSE	8.36	6.07	8.03
MAE	7.26	5.02	6.51

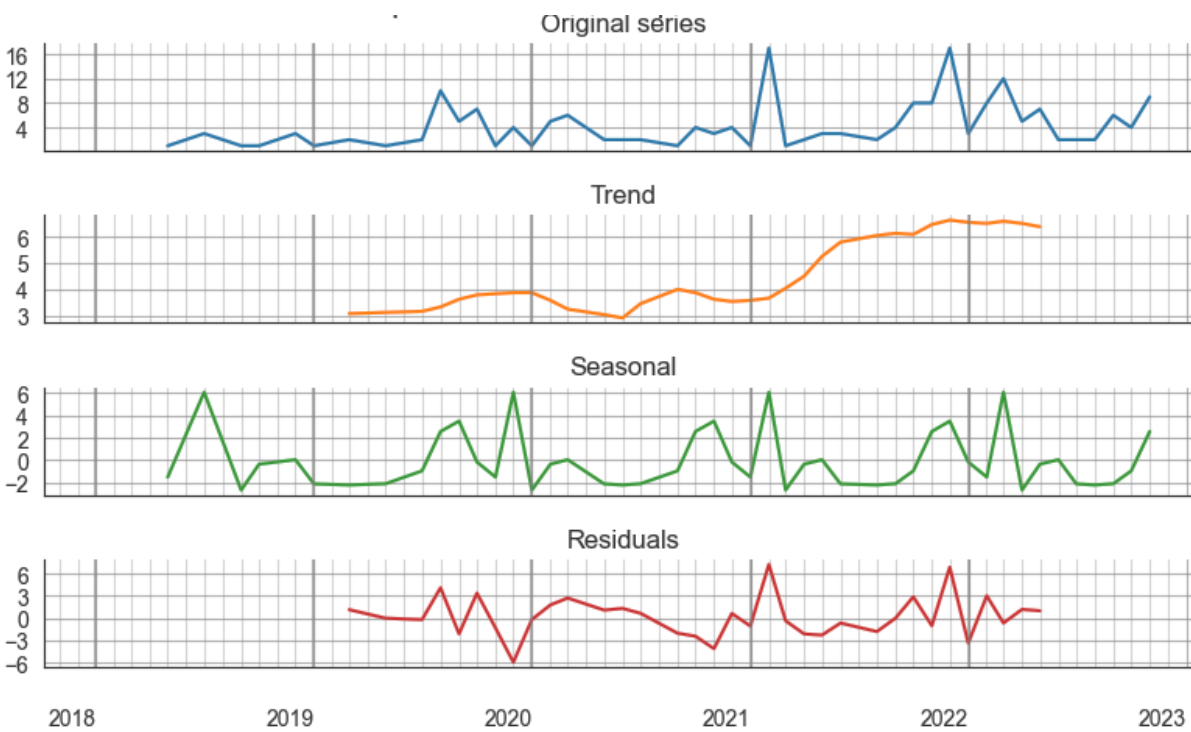
El modelo ARIMA resulta tener el mejor rendimiento para las predicciones de las ventas de vehículos eléctricos en el departamento antioqueño. ARIMA presenta valores ligeramente inferiores en términos del error de raíz cuadrada media y error absoluto medio, sin embargo, el error cuadrático medio presenta un valor significativamente inferior a lo observado en los modelos Prophet y la suavización exponencial.

### **Cundinamarca**

En la serie se observan picos estacionales de ventas durante los últimos cuatro años (**Figura 38**) lo cual supone continuidad en la voluntad de adopción del EV por parte de los consumidores, así como consistencia en la tendencia positiva de las ventas a partir de 2018, por lo tanto, los modelos de predicción se realizaron a partir del 2018 y con muestreo mensual debido a la adición de ruido en muestreos inferiores. La serie es estacionaria, por lo que se pasa directamente a la aplicación del modelo de SE para generar las predicciones.

**Figura 38**

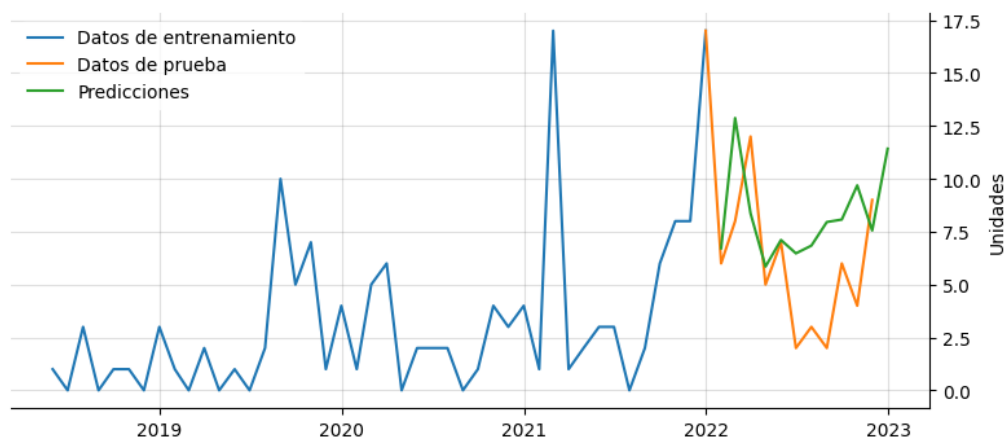
*Descomposición de la Serie de Tiempo de EV en Cundinamarca*



La predicción indica que en el departamento de Cundinamarca serán adquiridos entre 3 y 17 vehículos mensuales durante los próximos 12 meses, con un  $R^2$  de -0.49.

**Figura 39**

*Predicción del Modelo SE para la venta de EV en Cundinamarca ( $R^2$ : -0.49)*



A continuación, se presentan los resultados de la implementación de los modelos regresivos al conjunto de datos de venta de vehículos eléctricos en Cundinamarca.

**Tabla 14**

*Resultados de Optimización de Modelos de Regresión para Cundinamarca ( $R^2$ )*

p	d	q	AIC	$R^2$
5	1	5	252.234528	23.85
4	0	5	255.855953	23.42
4	0	4	253.908046	23.37

El modelo ARIMA óptimo para los datos de cantidad de vehículos eléctricos en Cundinamarca tiene un orden de (5, 1, 5), ya que este conjunto de parámetros presenta el  $R^2$  más alto entre los evaluados, se observa que el segundo y el tercer modelo son ARMA, donde el parámetro de orden de diferenciación es igual a cero.

**Tabla 15**

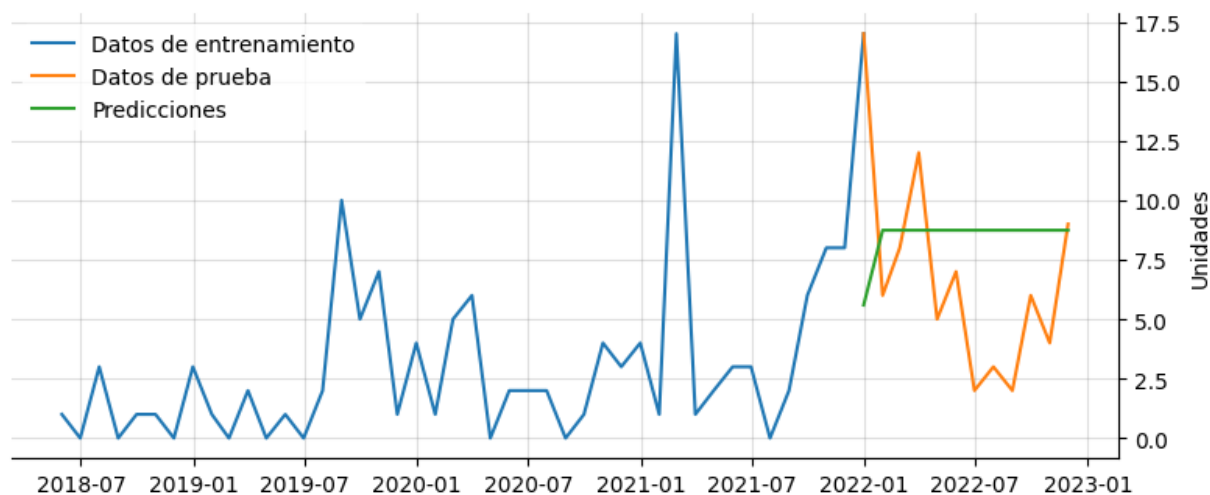
*Resultados de Optimización de Modelos de Regresión para Cundinamarca (AIC)*

p	d	q	AIC	R <sup>2</sup>
0	1	1	241.956656	5.08
1	1	1	243.860693	4.88
0	1	2	243.952704	5.01

En cuanto al AIC, el menor valor encontrado fue 241.95, asociado al orden (0, 1, 1). Sin embargo, este último orden genera un R<sup>2</sup> cercano al 5.08 %, lo que indica que no es capaz de capturar adecuadamente la variación en los datos de entrenamiento.

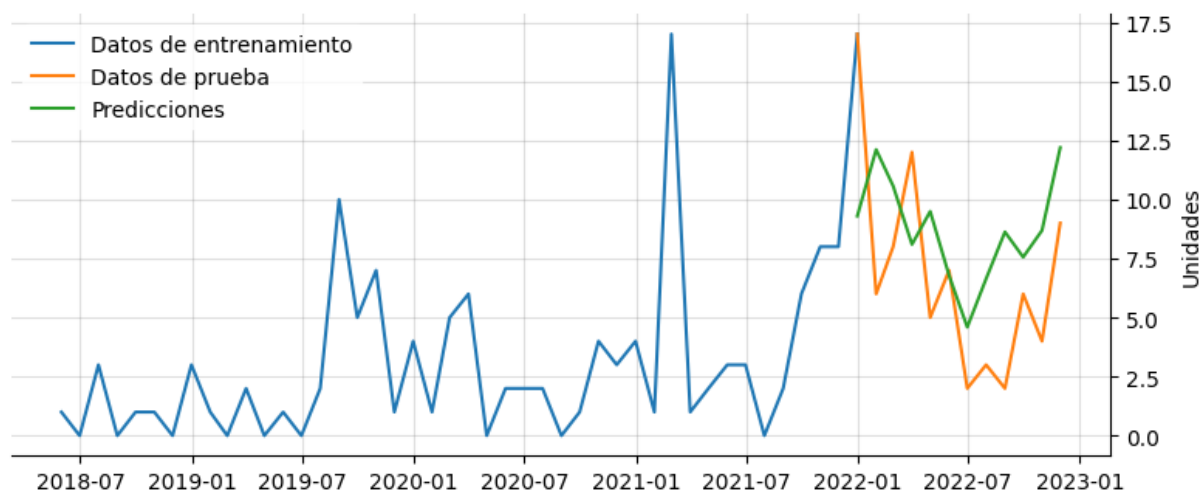
**Figura 40**

*Modelo ARIMA (0, 1, 1) para la Serie de Tiempo de Cundinamarca*



**Figura 41**

*Modelo ARIMA (5, 1, 5) para la Serie de Tiempo de Cundinamarca*



Gráficamente, el modelo ARIMA (5, 1, 5) parece capturar la tendencia general de los datos en el departamento de Cundinamarca. Sin embargo, no logra representar correctamente los puntos altos y bajos de las ventas reales. Se observa que, aunque las predicciones se acercan a los datos de prueba en algunos casos, el modelo subestima los valores en períodos donde las ventas reales son altas, prediciendo valores significativamente más bajos. Esto se refleja en un  $R^2$  de -0.13, indicando un rendimiento deficiente.

Respecto al modelo ARIMA (0, 1, 1), se obtiene un  $R^2$  menor, gráficamente las predicciones generan una línea lo cual no describe correctamente los datos de prueba.

Finalmente, se implementa el modelo Prophet a nivel departamental.

**Tabla 16**

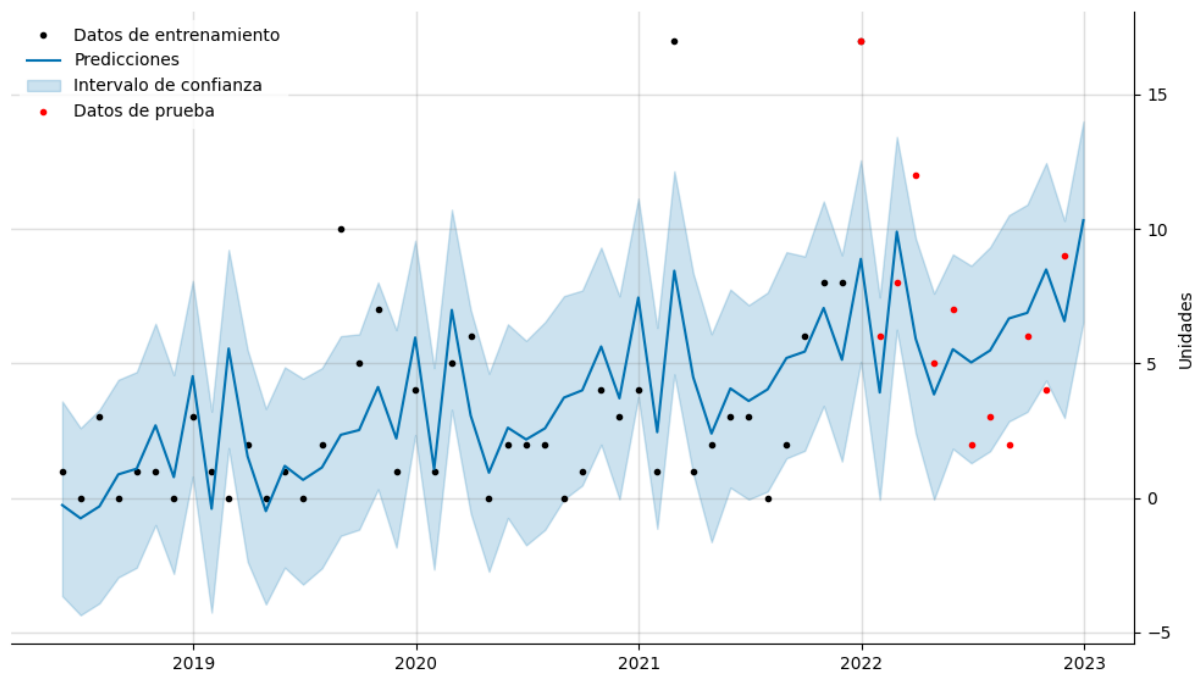
*Valores Óptimos Prophet para Datos Mensuales de Cundinamarca*

changeoint_prior_scale	seasonality_prior_scale	seasonality_mode	RMSE
0.10	0.1	additive	4.584037
0.01	0.1	additive	4.598603
0.50	0.1	additive	5.539640

El modelo Prophet para las ventas de EV's en Cundinamarca alcanza su mejor desempeño con un enfoque aditivo de estacionalidad y una fuerte regularización de los cambios de tendencia (changeoint\_prior\_scale bajo).

**Figura 42**

*Resultados del Modelo Prophet para Cundinamarca ( $R^2: -0.55$ )*



El modelo obtiene un  $R^2$  de -0.55, lo que indica un bajo rendimiento. A diferencia de la implementación del modelo Prophet en otros departamentos, en el caso de Cundinamarca, los datos no parecen estar significativamente afectados por la pandemia de COVID-19. De hecho, al retirar el regresor relacionado con el COVID, el rendimiento del modelo mejora, alcanzando su mejor valor de  $R^2 = -0.55$ .

### Figura 43

*Métricas de los Modelos Aplicados a la Serie de Cundinamarca*

Métrica	Suavización exponencial	ARIMA	Prophet
$R^2$	-0.49	-0.13	-0.55
MSE	26.06	19.77	27.09
RMSE	5.1	4.45	5.2
MAE	4.32	3.94	4.01

Se observa que las ventas de EV's en el departamento de Cundinamarca son descritas mejor por el modelo ARIMA, a pesar de tener un  $R^2$  negativo, este es superior al obtenido de los modelos de suavización exponencial y Prophet. El modelo ARIMA presentó un rendimiento ligeramente superior resultando en menores valores de error.

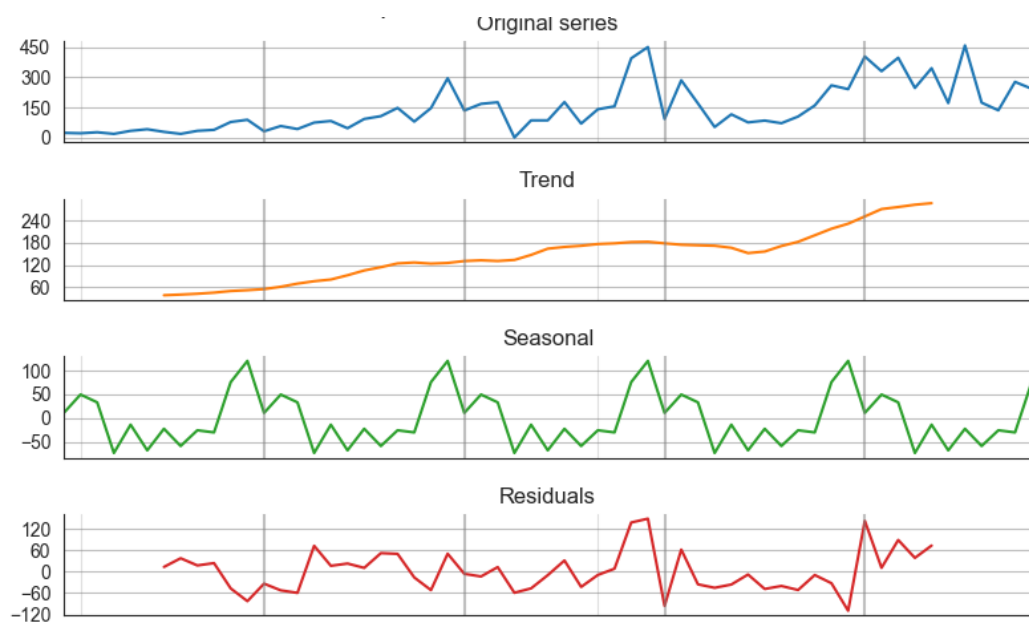
### Predicción de Ventas a 2030

Para la predicción de ventas al año 2030 se toman los registros correspondientes a la totalidad de los vehículos híbridos y eléctricos registrados en el *dataset* a partir de 2018, para todas las marcas y departamentos del país. La serie exhibe un descenso sobre el año 2020 y un repunte hacia 2022, con una tendencia positiva hacia este año y picos estacionales dos veces por año. Los residuos de la serie se comportan aleatoriamente, sin apreciarse una tendencia en los mismos, como se aprecia en la **Figura 44**.

Se aplica el modelo de suavización exponencial a la serie de tiempo original, que registra un valor de -2.3865 en la prueba Dickey Fuller aumentada (ADF) y un valor p de 0.1455, confirmando la carencia de estacionalidad. La serie se diferencia una vez, adquiriendo estacionariedad que se confirma aplicando la prueba ADF nuevamente obteniendo un valor de -5.0743 y un valor p de 1.5788.

#### Figura 44

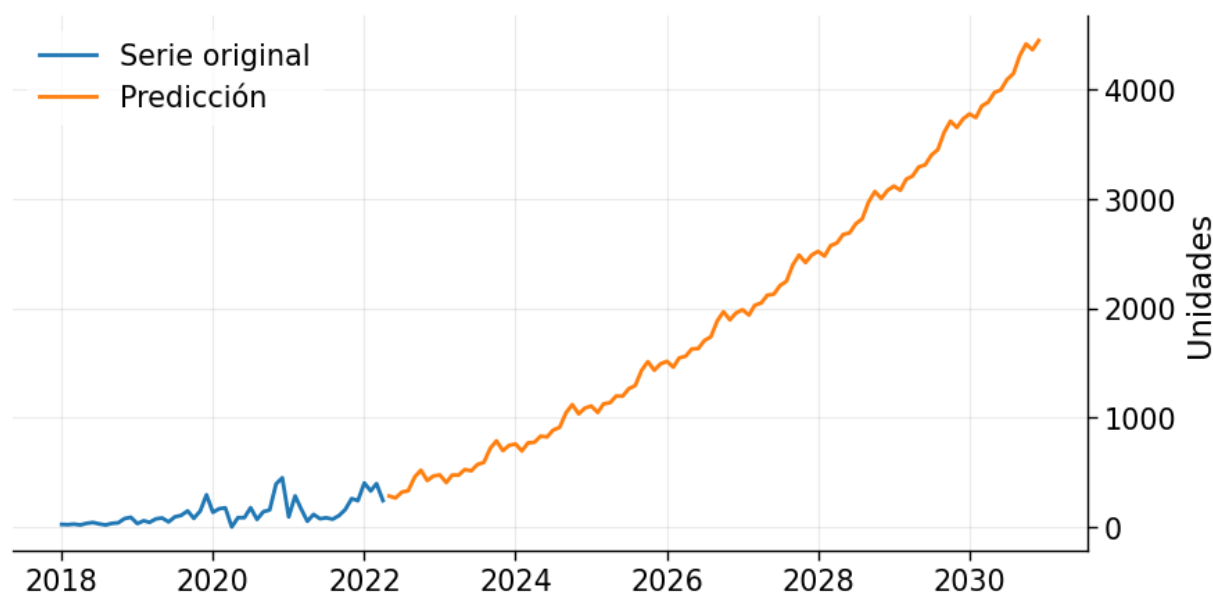
##### *Descomposición Estacional de la Serie Total EV e híbridos*



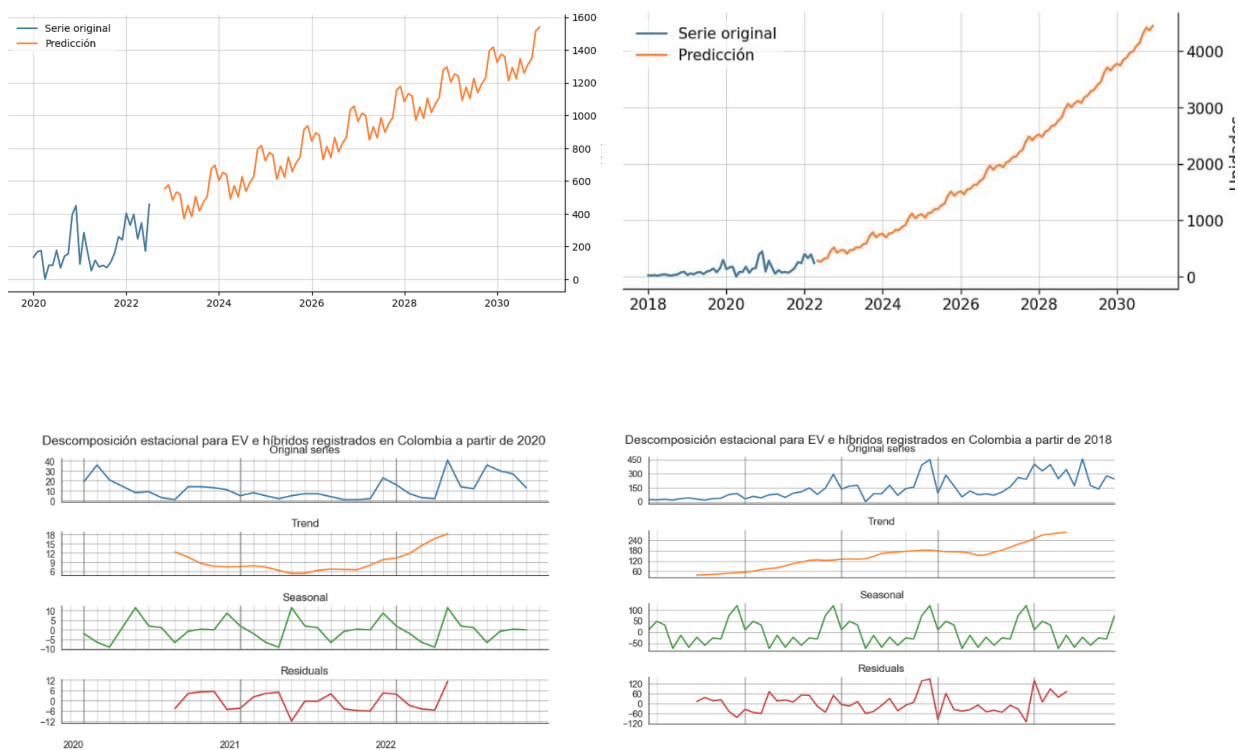
Luego de restaurar los valores de la serie, se grafican las predicciones, estimando que hacia el año 2030 se registrarán cerca de 4.000 vehículos entre híbridos y EV. Desde luego, esta es una cifra condicionada a varias variables como las arancelarias, impositivas, la masificación de la red eléctrica de carga, inclusive las proyecciones de diversificación eléctrica que soporten el impacto sobre la red de suministro, asociado al incremento de movilidad alternativa. Estas proyecciones exceden el alcance de este trabajo y bien merecen la pena ser analizadas con detenimiento.

### Figura 45

*Proyección de Ventas de EV e Híbridos hasta 2030 Basada en SE*



Durante la etapa de modelado de la serie se observó el cambio en la forma de la predicción como resultado de la extensión del número de registros (**Figura 46**). En particular, se aprecia una mayor captura de la estacionalidad en el modelo con datos desde 2020, en tanto que el modelo con datos iniciados desde 2018 capta mejor la tendencia, además se percibe una mejora en el nivel, propio de una sustancial reducción de ruido.

**Figura 46***Efecto del Número de Registros Sobre la Predicción*

*Nota.* Serie con datos a partir de 2020 (Izquierda), Serie con datos a partir de 2018 (Derecha).

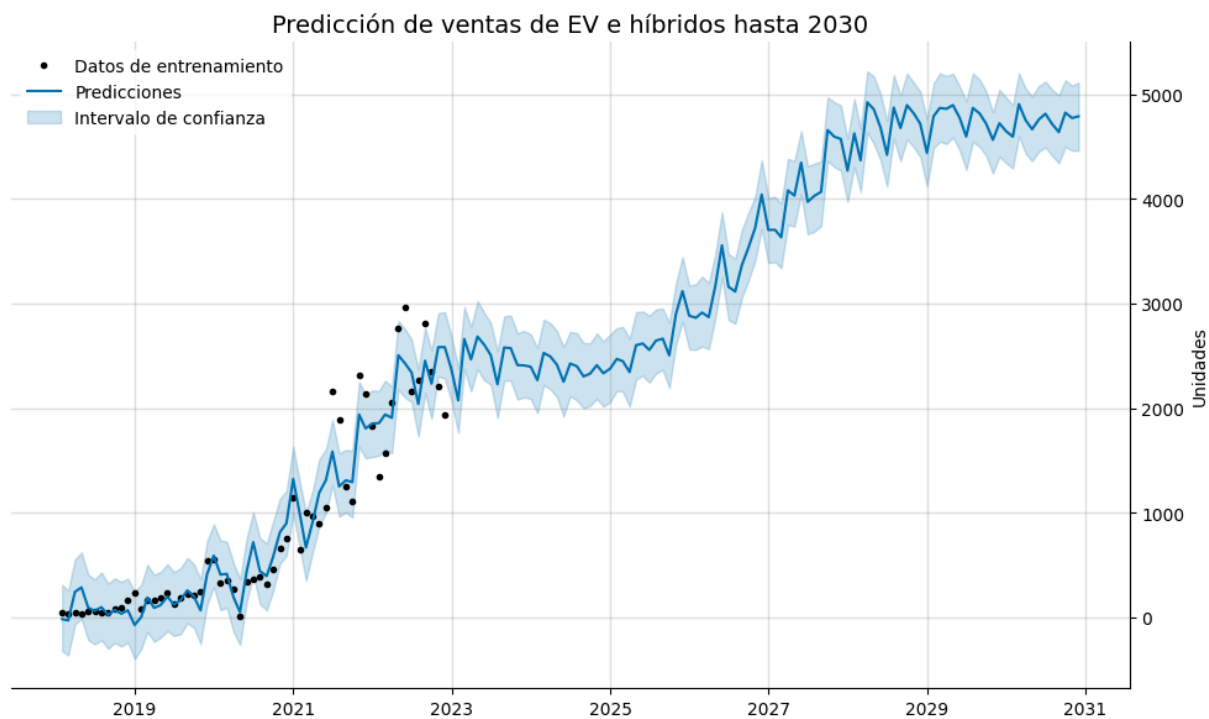
**Proyección con Prophet**

Para el desarrollo del modelo con Prophet se optimizaron los parámetros con base en el 90 % de los datos iniciales (desde 2018 hasta el 2022), adicional, se incluyó un factor de estacionalidad mensual, semestral y anual con diferentes pesos para incrementar la precisión del modelo. Los parámetros utilizados fueron los siguientes:

- yearly\_seasonality=False
- changepoint\_prior\_scale=0.1
- seasonality\_prior\_scale=1
- seasonality\_mode=additive

**Figura 47**

*Proyección de Ventas Basada en el Método Prophet*



Se utilizó el modelo Prophet para proyectar las ventas de vehículos eléctricos hasta el 2030. El modelo indica un crecimiento a lo largo de los años pasando de 3.500 vehículos vendidos en promedio por mes en el año 2023 hasta las 4.800 unidades vendidas en promedio por mes para el año 2030. Debemos destacar que esta proyección se realizó solamente sobre los vehículos particulares, por lo cual, aquellas iniciativas del gobierno no son incluidas, esto, como se evidenció con la introducción de los buses eléctricos en Bogotá, puede afectar enormemente el desempeño del modelo. La gran variación de los datos a finales del 2022 dificulta la precisión del modelo, generando errores altos y cuestionando los valores predichos en horizontes muy lejanos.

De acuerdo con la proyección usando el modelo Prophet, se estima que para el 2030 haya un total de 227.254 vehículos eléctricos e híbridos, cifra inferior a lo proyectado por el gobierno colombiano en donde se establece una meta de 600.000 vehículos.

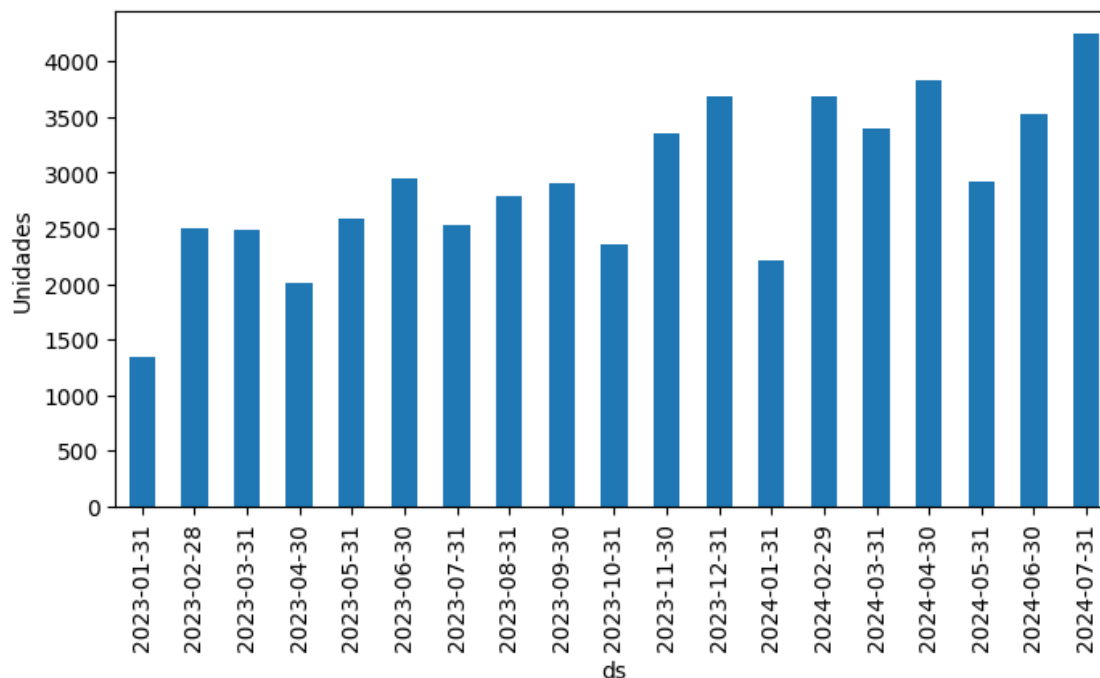
### Proyección Ventas 2023

Según (Fenalco y Andi, 2024) en 2023, a nivel nacional se vendieron en total 31.490 vehículos eléctricos, y hasta julio del 2024 las ventas alcanzaban las 23.818 unidades. En la **Figura 46** observamos que las ventas de EV continúan creciendo después del 2022 y no se presentan cambios significativos como los observados en años anteriores.

Adicional, observamos que las ventas son bajas al inicio de cada año y para julio del 2024 las ventas superan las 4000 unidades, cifra que el modelo alcanza a finales del 2026.

### Figura 48

*Ventas de Vehículos Eléctricos e Híbridos en Colombia*

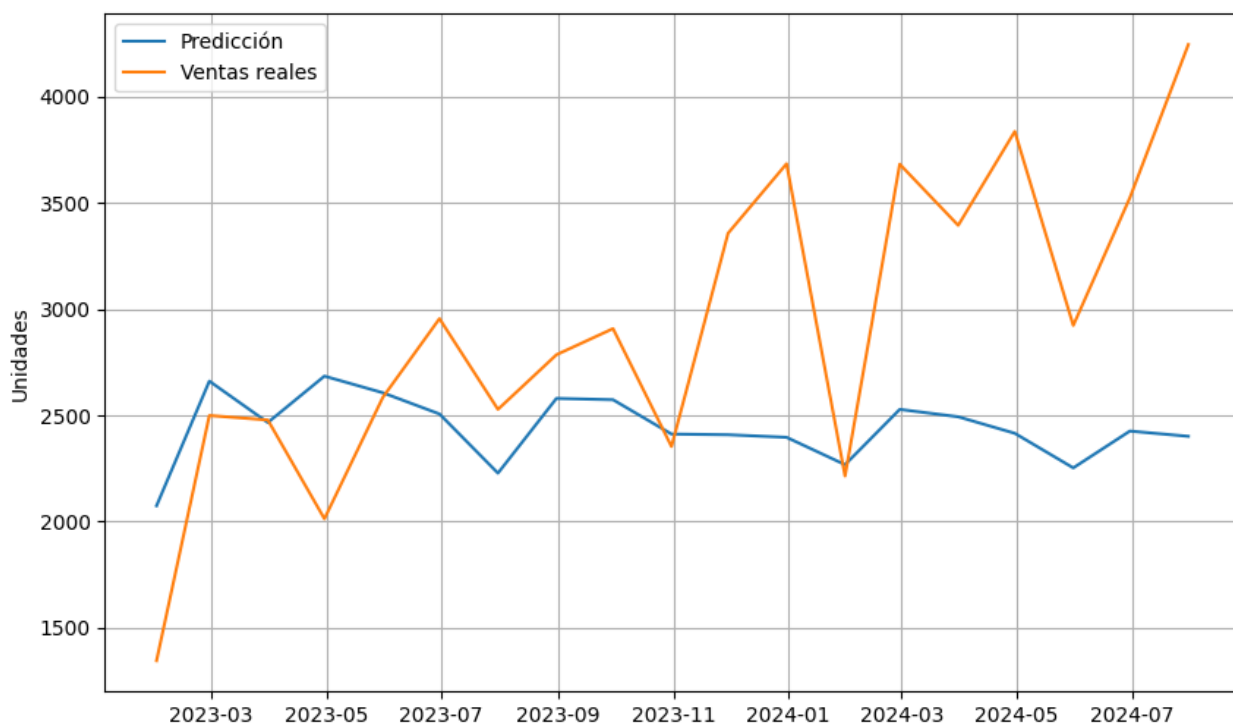


*Nota.* Datos de Fenalco (2023).

Con base en el modelo desarrollado con Prophet, las ventas de EV's en el país para el 2023 y el primer semestre del 2024 las ventas presentan un comportamiento estable, mientras que las ventas reales exhiben picos y caídas abruptas, esto puede deberse a factores externos que no fueron considerados en el modelo como nuevos incentivos o la introducción de una nueva marca en el país, las métricas de este modelo muestran un bajo rendimiento del modelo lo que soporta lo observado en la **Figura 47**,  $R^2$ : -0.42, MSE: 696619, RMSE: 834 y MAE 648.

### Figura 49

*Proyección de Ventas 2023 con Prophet*



Otros hiperparámetros que generan un alto rendimiento del modelo sobre los datos de entrenamiento incluyen los siguientes valores:

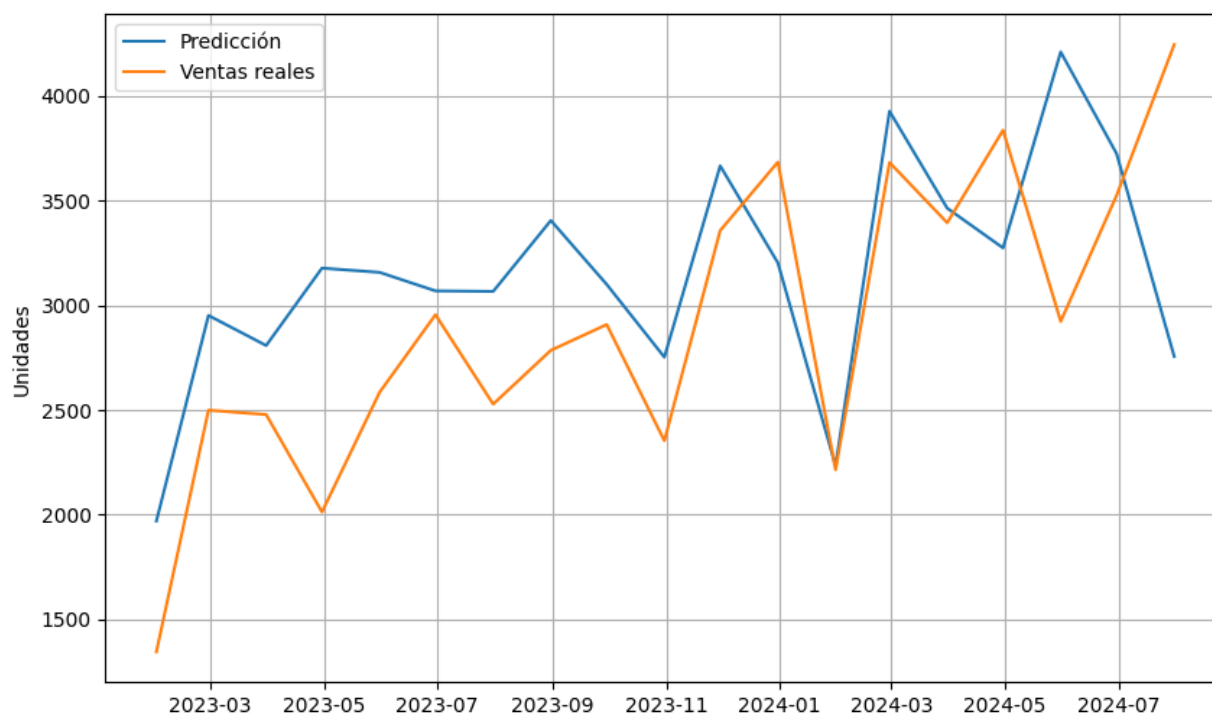
- yearly\_seasonality=False
- changepoint\_prior\_scale=1

- seasonality\_prior\_scale=5
- seasonality\_mode=multiplicative

Aunque este modelo produce predicciones más cercanas a las ventas reales en el corto plazo, a largo plazo presenta valores erráticos, con predicciones que oscilan entre 100 unidades y 13,000 unidades por mes, cifras muy alejadas de la realidad. Las métricas calculadas para este modelo, aplicadas a las ventas del 2024, son:  $R^2$  de 0.15, MSE de 415,388, RMSE de 644 y MAE de 508, reflejando una mejora respecto al modelo previamente analizado.

### Figura 50

*Proyección de Ventas 2023 con Variación Hiperparámetros en Prophet*



## Conclusiones

Para el pronóstico se aplicaron modelos de suavización exponencial, modelos regresivos y el modelo Prophet, siendo este último, el modelo más avanzado por el nivel de detalle y parámetros que se pueden configurar en el desarrollo de los modelos. El rendimiento de cada modelo está sujeto a las características de los datos tales como el nivel de ruido, cantidad de datos, entre aquellos factores inherentes a las series de tiempo.

Los vehículos eléctricos en Colombia se clasifican en vehículos eléctricos e híbridos, donde la principal diferencia está en el uso de un combustible alterno en conjunto con la energía eléctrica. Para el desarrollo de los modelos y los análisis por marcas y departamentos se hizo especial énfasis en los vehículos eléctricos siendo estos aquellos que tienen menor penetración en el mercado colombiano.

La baja penetración de vehículos eléctricos en Colombia se debe principalmente al bajo desarrollo de la infraestructura de las zonas de carga. Esto ha generado que la cantidad de vehículos dependa del departamento, el 78 % de los vehículos eléctricos se encuentran en los tres departamentos con mayor desarrollo del país, siendo Bogotá D.C. el departamento con alrededor del 45 % de los vehículos eléctricos e híbridos.

Las ventas de vehículos eléctricos en Colombia pueden ser descrita de dos formas, por marcas de vehículos y por zona geográfica con mayor cantidad de unidades registradas. Frente a las marcas, se evidenció que BYD, Renault y BMW dominan el 80 % del mercado colombiano.

La cantidad de ventas de EV en la capital del país y en Antioquia presenta un incremento significativo entre 2021 y 2022, esto es consecuencia de dos factores, el primero fue la compra de buses eléctricos para el servicio público y la introducción posterior de la marca BYD en el mercado, ofreciendo SUV de muy buena relación calidad-precio.

La proyección de las ventas de EV por marca y por departamentos resulta ser desafiante debido a las características de los datos y el comportamiento de la demanda. Los modelos de suavización exponencial, ARIMA y Prophet presentan rendimientos bajos y se hace necesario un estudio más exhaustivo enfocado en un solo departamento/marca para explorar otros modelos que permitan capturar mejor las características de los datos.

Los modelos generados para la predicción de ventas por departamento fueron adecuados con el fin de reducir el ruido de los datos que generaba desempeños muy bajos frente al grupo de datos de prueba, los cambios realizados contemplan la reducción del horizonte de datos pasando a usar los datos únicamente del 2018 al 2022, también se consideró la reducción de los datos de prueba para permitir que los modelos capturaran los aquellos en la demanda ocurridos después del 2021.

El modelo Prophet resulta ser el modelo más complejo, donde se permite configurar parámetros como la tendencia, temporalidad, días festivos, entre otros. En los modelos planteados, se incluyó un periodo denominado covid\_periodo el cual permitió considerar los cambios en las ventas de EV debido al COVID-19, esto fue fundamental en los departamentos de Bogotá y Antioquia para lograr mejores resultados.

La cantidad de datos disponibles fue determinante, influyendo significativamente en la sensibilidad de los modelos, lo cual se aprecia en la serie de predicción de ventas a 2030.

Los datos sobre los que se basó este estudio abarcan registros desde el 29 de enero de 2010 hasta el 30 de noviembre de 2022. Pese a que se solicitó una versión actualizada del *dataset* al Ministerio de Transporte, no se obtuvo respuesta.

Como se vio en la proyección a 2030 mediante el modelo de suavización exponencial, la extensión de la serie de tiempo es crucial para determinar predicciones confiables, de tal manera que el algoritmo logre captar la variabilidad en los datos.

Pese a las proyecciones conservadoras obtenidas con los modelos de suavización exponencial y autorregresivos, se prevé un incremento en el tiempo, de las marcas que lideran el mercado. Tanto BMW, líder del segmento de lujo, como BYD, que lidera los modelos SUV, y Renault, que se sostiene como una opción multinivel, exhiben crecimientos en el mediano plazo.

Las proyecciones para el 2023 muestran que, aunque el modelo logra capturar ciertas tendencias generales del comportamiento de las ventas de vehículos eléctricos en Colombia, existen limitaciones en la precisión debido a la incapacidad de reflejar las fluctuaciones abruptas observadas en los datos reales. Esto indica que el modelo es adecuado para proporcionar una visión general, pero carece de la sensibilidad necesaria para predecir variaciones significativas impulsadas por factores externos.

## Recomendaciones

Ampliar la cantidad de datos y la granularidad de los datos de registros de vehículos eléctricos e híbridos, así como asegurar la actualización frecuente, lo cual permitirá capturar tendencias reales y reducir errores en los modelos predictivos, favoreciendo la profundización de estudios ulteriores basados en la adopción de medios de transporte alternativos.

Explorar con mayor detalle cómo cambios regulatorios, impuestos y precios de la energía eléctrica impactan en la voluntad de compra de los usuarios. Estos hallazgos permitirían realizar análisis que reflejen estimaciones más precisas.

Debido a la variedad de factores que pueden incidir en los volúmenes de venta de EV, se sugiere la exploración de modelos de *machine learning* de desempeño superior como LSTM (*long short-term memory*), buscando la mejora en la captación de la variabilidad en los datos, así como la adaptación a la estructura de la serie en predicciones de largo plazo.

## Referencias

- Afandizadeh, S., Sharifi, D., Kalantari, N., & Mirzahosseini, H. (2023). Using machine learning methods to predict electric vehicles penetration in the automotive market. *Scientific Reports*, 13(1), 8345. <https://doi.org/10.1038/s41598-023-35366-3>
- Bamos, Z. N., Laitos, V. M., Afentoulis, K. D., Vagropoulos, S. I., & Biskas, P. N. (2024). Electric vehicles load forecasting for day-ahead market participation using machine and deep learning methods. *Applied Energy*, 360, 122801. <https://doi.org/10.1016/j.apenergy.2024.122801>
- Bonaccorso, G. (2017). *Machine learning algorithms: A reference guide to popular algorithms for data science and machine learning*. Packt Publishing.
- Buzna, L., Falco, P. D., Khormali, S., Proto, D., & Straka, M. (2019). *Electric vehicle load forecasting: A comparison between time series and machine learning approaches*.
- Castro Mediavilla, J. J., Orbea Hinojosa, L. X., Toapaxi Csanoba, J. A., & Guano Calvache, C. A. (n.d.). *Análisis de incentivos y proyecciones del vehículo 100% eléctrico en el Ecuador*. 2(4), 112–124. <https://doi.org/10.33890/innova.v2.n4.2017.243>
- Dabbaghjamesh, M., Moeini, A., & Kavousi-Fard, A. (2021). Reinforcement Learning-Based Load Forecasting of Electric Vehicle Charging Station Using Q-Learning Technique. *IEEE Transactions on Industrial Informatics*, 17(6), 4229–4237. <https://doi.org/10.1109/TII.2020.2990397>
- Datos Abiertos Colombia. (2022). *Numero de Vehículos Eléctricos - Híbridos*. [https://www.datos.gov.co/Transporte/Numero-de-Veh-culos-El-ctricos-Hibridos/7qfh-tkr3/about\\_data](https://www.datos.gov.co/Transporte/Numero-de-Veh-culos-El-ctricos-Hibridos/7qfh-tkr3/about_data)

Datos Abiertos Colombia. (2024, June 18). *Crecimiento del parque automotor Runt2.0*.

<https://www.datos.gov.co/Transporte/CRECIMIENTO-DEL-PARQUE-AUTOMOTOR-RUNT2-0/u3vn-bdcy/data>

Facebook. (2024). *Seasonality, Holiday Effects, And Regressors*.

[http://facebook.github.io/prophet/docs/seasonality,\\_holiday\\_effects,\\_and\\_regressors.html](http://facebook.github.io/prophet/docs/seasonality,_holiday_effects,_and_regressors.html)

Fenalco y Andi. (2024). *Informe Vehículos Eléctricos e Híbridos Julio 2024*.

<https://www.fenalco.com.co/blog/gremial-4/informe-vehiculos-electricos-e-hibridos-julio-2024-7393>

Giansoldati, M., Rotaris, L., Scorrano, M., & Danielis, R. (2020). Does electric car knowledge influence car choice? Evidence from a hybrid choice model. *Research in Transportation Economics*, 80, 100826. <https://doi.org/10.1016/j.retrec.2020.100826>

Golsefidi, A. H., Hüttel, F. B., Peled, I., Samaranayake, S., & Pereira, F. C. (2023). A joint machine learning and optimization approach for incremental expansion of electric vehicle charging infrastructure. *Transportation Research Part A: Policy and Practice*, 178, 103863. <https://doi.org/10.1016/j.tra.2023.103863>

González Mares, M. (2019). Hernández-Sampieri, R. & Mendoza, C (2018). Metodología de la investigación. Las rutas cuantitativa, cualitativa y mixta. *Revista Universitaria Digital de Ciencias Sociales (RUDICS)*, 10(18), 92–95. <https://doi.org/10.22201/fesc.20072236e.2019.10.18.6>

González, R. (2022). ¿Por qué BYD se consolida en Colombia como el líder en vehículos eléctricos? In *V12 Magazine*. <https://v12magazine.com/tecnologia/por-que-byd-se-consolida-en-colombia-como-el-lider-en-vehiculos-electricos/>

- Hernández-Ambato, J., Fernández, R., Mora, A., & Alvarado, J. (2022a). Evaluación de la huella de carbono de vehículos con motor eléctrico y de combustión interna según la matriz energética de Ecuador: Caso de estudio KIA Soul vs KIA Soul EV. *NOVASINERGIA REVISTA DIGITAL DE CIENCIA, INGENIERÍA Y TECNOLOGÍA*, 5(2), 58–75.  
<https://doi.org/10.37135/ns.01.10.04>
- Hernández-Ambato, J., Fernández, R., Mora, A., & Alvarado, J. (2022b). Evaluación de la huella de carbono de vehículos con motor eléctrico y de combustión interna según la matriz energética de Ecuador: Caso de estudio KIA Soul vs KIA Soul EV. *NOVASINERGIA REVISTA DIGITAL DE CIENCIA, INGENIERÍA Y TECNOLOGÍA*, 5(2), 58–75.  
<https://doi.org/10.37135/ns.01.10.04>
- IDEAM, Natura, F., PNUD, MADS, DNP, & Cancillería. (2022). *Informe del inventario nacional de gases efecto invernadero 1990-2018 y carbono negro 2010-2018 de Colombia* (Issue No. 3).  
<https://unfccc.int/sites/default/files/resource/Annex%20BUR3%20COLOMBIA.pdf>
- Kamis, A., & Susan Abraham, P. (2024). Predictive models of electric vehicle adoption in the United States: Charging ahead with renewable energy. *Transportation Research Interdisciplinary Perspectives*, 24, 101041. <https://doi.org/10.1016/j.trip.2024.101041>
- Krivevski, B. (2022). BMW Group doubles global sales of fully-electric vehicles in first half-year. In *Electric Cars Report*. <https://electriccarsreport.com/2022/07/bmw-group-doubles-global-sales-of-fully-electric-vehicles-in-first-half-year/>
- Kyriakides, G., & Margaritis, K. G. (2019). *Hands-on ensemble learning with Python: Build highly optimized ensemble machine learning models using scikit-learn and Keras*. Packt Publishing.

- Lo Franco, F., Ricco, M., Cirimele, V., Apicella, V., Carambia, B., & Grandi, G. (2023a). Electric Vehicle Charging Hub Power Forecasting: A Statistical and Machine Learning Based Approach. *Energies*, *16*(4), 2076. <https://doi.org/10.3390/en16042076>
- Lo Franco, F., Ricco, M., Cirimele, V., Apicella, V., Carambia, B., & Grandi, G. (2023b). Electric Vehicle Charging Hub Power Forecasting: A Statistical and Machine Learning Based Approach. *Energies*, *16*(4), 2076. <https://doi.org/10.3390/en16042076>
- Luis, F., José, S., Sacristán, R., Javier, A., Orozco, W., Rosa, C., & Suarez, M. (2012). “Informe de visita fiscal.” *Contraloría de Bogota D.C.* [www.contraloriabogota.gov.co](http://www.contraloriabogota.gov.co)
- Madhavan, S. (2015). *Mastering Python for data science: Explore the world of data science through Python and learn how to make sense of data*. Packt Publishing.
- Marino, C. A., & Marufuzzaman, M. (2020). Unsupervised learning for deploying smart charging public infrastructure for electric vehicles in sprawling cities. *Journal of Cleaner Production*, *266*, 121926. <https://doi.org/10.1016/j.jclepro.2020.121926>
- Mauricio, J. (2007). *Introducción al Análisis de Series Temporales*. Universidad Complutense de Madrid.
- Ministerio de Ambiente y Desarrollo Sostenible. (2018). *Con medidas estratégicas a corto y mediano plazo Medellín mejora su calidad de aire | Ministerio de Ambiente y Desarrollo Sostenible*. <https://archivo.minambiente.gov.co/index.php/noticias-minambiente/3653-con-medidas-estrategicas-a-corto-y-mediano-plazo-medellin-mejora-su-calidad-de-aire>
- Ministerio de transporte. (2021, October 19). *En 2022 Colombia tendrá 1.589 buses eléctricos operando en sus sistemas masivos, la mayor flota eléctrica de Latinoamérica*. <https://mintransporte.gov.co/publicaciones/10346/en-2022-colombia-tendra-1589-buses-electricos-operando-en-sus-sistemas-masivos-la-mayor-flota-electrica-de-latinoamerica/>

- Mouhy-Ud-Din, M. A., Azeem, F., Memon, Z., & Arshad, J. (2024). Comparative techno-economic assessment of electric vehicle charging preferences: A quantitative study for sustainable EV policy in Pakistan. *Energy for Sustainable Development*, 78, 101355. <https://doi.org/10.1016/j.esd.2023.101355>
- Núñez Hernandez, F., & Arcos-Vargas, Á. (2018). *ANÁLISIS COMPARATIVO A NIVEL INTERNACIONAL DE LA EXPANSIÓN DEL VEHÍCULO ELÉCTRICO*.
- Pal, A., & Prakash, P. (2017). *Practical Time Series Analysis* (1st ed.). Packt Publishing.
- Pamidimukkala, A., Kermanshachi, S., Rosenberger, J. M., & Hladik, G. (2023). Evaluation of barriers to electric vehicle adoption: A study of technological, environmental, financial, and infrastructure factors. *Transportation Research Interdisciplinary Perspectives*, 22, 100962. <https://doi.org/10.1016/j.trip.2023.100962>
- Quintero Abril, L. F. (2020). *Evaluación del impacto de los vehículos con tecnología híbrida en el sector asegurador para la empresa CESVI Colombia S. A.* [Fundación Universidad de América]. <https://repository.uamerica.edu.co/bitstream/20.500.11839/7893/1/4081053-2020-1-IM.pdf>
- Ramos Robayo, J. G., & Herrera Romero, J. A. (2024a). *Análisis de viabilidad de infraestructura de recarga de vehículos eléctricos en Colombia* [Universidad Distrital Francisco José De Caldas]. <http://hdl.handle.net/11349/40144>
- Ramos Robayo, J. G., & Herrera Romero, J. A. (2024b). *Análisis de viabilidad de infraestructura de recarga de vehículos eléctricos en Colombia* [Universidad Distrital Francisco José De Caldas]. <http://hdl.handle.net/11349/40144>

- Ramos-Pérez, F., Vivas Venegas, C., & Rubio, F. R. (2023). Predicción de demanda y generación renovable con Deep Learning: Aplicación a la optimización de estaciones de carga de vehículos eléctricos. *XLIV Jornadas de Automática: Libro de Actas*, 346–351.  
<https://doi.org/10.17979/spudc.9788497498609.346>
- Rojas Quiroga, K. E., Nieto Vargas, J. F., & Marulanda, G. A. (2019). Impactos técnicos y económicos para comercializadores de electricidad debido a la implementación de vehículos eléctricos en Colombia. *Revista Ingenierías Universidad de Medellín*, 18(34), 219–236.  
<https://doi.org/10.22395/rium.v18n34a13>
- Selvaraj, V., & Vairavasundaram, I. (2023). A comprehensive review of state of charge estimation in lithium-ion batteries used in electric vehicles. *Journal of Energy Storage*, 72, 108777. <https://www.sciencedirect.com/science/article/pii/S2352152X23021746>
- Steer. (2020). *Realizar un estudio que permita identificar las clases de vehículos y modalidades de transporte susceptibles de realizar el ascenso tecnológico hacia tecnologías de cero y bajas emisiones a nivel nacional.*
- Vargas, L. (2023). ¿En qué anda la movilidad eléctrica en Colombia? In *Forbes Colombia*.  
<https://forbes.co/2023/10/23/actualidad/en-que-anda-la-movilidad-electrica-en-colombia>
- Xing, Q., Chen, Z., Zhang, Z., Wang, R., & Zhang, T. (2021). Modelling driving and charging behaviours of electric vehicles using a data-driven approach combined with behavioural economics theory. *Journal of Cleaner Production*, 324, 129243.  
<https://doi.org/10.1016/j.jclepro.2021.129243>
- Zarazua De Rubens, G. (2019). Who will buy electric vehicles after early adopters? Using machine learning to identify the electric vehicle mainstream market. *Energy*, 172, 243–254.  
<https://doi.org/10.1016/j.energy.2019.01.114>

Zhao, M., Fang, Y., & Dai, D. (2023). Forecast of the Evolution Trend of Total Vehicle Sales and Power Structure of China under Different Scenarios. *Sustainability*, *15*(5), 3985.

<https://doi.org/10.3390/su15053985>