UNIVERSIDAD TECNOLÓGICA ISRAEL

ESCUELA DE POSGRADOS "ESPOG"

MAESTRÍA EN
BIG DATA Y CIENCIA DE DATOS

Resolución: RPC-SO-32-No.536-2023.

PROYECTO DE TITULACIÓN EN OPCIÓN AL GRADO DE MAGISTER

Título del proyecto:
Análisis predictivo de ventas utilizando el método CRISP DM para el sector automotriz en el Ecuador
Línea de Investigación:
Ciencias de la ingeniería aplicadas a la producción, sociedad y desarrollo sustentable
Campo amplio de conocimiento:
Tecnologías de la Información y Comunicación TIC
Autor/a:
Diana Lisseth Cueva Villarroel
Tutor/a:
Mg. Mario Pérez (Tutor Técnico) Phd. Renato Toasa (Tutor Metodológico)

Quito – Ecuador

2025

APROBACIÓN DEL TUTOR

Yo, Mario Rubén Pérez Cargua con C.I: 0603251984 en mi calidad de Tutor del proyecto de investigación titulado: **Análisis predictivo de ventas utilizando el método CRISP DM para el sector automotriz en el Ecuador**.

Elaborado por: **Diana Lisseth Cueva Villarroel**, de C.I: **0503612251**, estudiante de la **Maestría en BIG DATA Y CIENCIA DE DATOS**, de la **UNIVERSIDAD TECNOLÓGICA ISRAEL (UISRAEL)**, como parte de los requisitos sustanciales con fines de obtener el Título de Magister, me permito declarar que luego de haber orientado, analizado y revisado el trabajo de titulación, lo apruebo en todas sus partes.

Quito D.M., marzo de 2025

Firmado electrónicamente por:
MARIO RUBEN
PEREZ CARGUA

_____

Firma

APROBACIÓN DEL TUTOR

Yo, Renato Mauricio Toasa Guachi, con C.I: 1804724167 en mi calidad de Tutor del proyecto de investigación titulado: **Análisis predictivo de ventas utilizando el método CRISP DM para el sector automotriz en el Ecuador**.

Elaborado por: **Diana Lisseth Cueva Villarroel**, de C.I: **0503612251**, estudiante de la **Maestría en BIG DATA Y CIENCIA DE DATOS**, de la **UNIVERSIDAD TECNOLÓGICA ISRAEL (UISRAEL)**, como parte de los requisitos sustanciales con fines de obtener el Título de Magister, me permito declarar que luego de haber orientado, analizado y revisado el trabajo de titulación, lo apruebo en todas sus partes.

Quito D.M., marzo de 2025

Firmado electrónicamente por:
RENATO MAURICIO
TOASA GUACHI

_____

Firma

DECLARACIÓN DE AUTORIZACIÓN POR PARTE DEL ESTUDIANTE

Yo, **Diana Lisseth Cueva Villarroel** con C.I: **0503612251,** autor/a del proyecto de titulación denominado: **Análisis predictivo de ventas utilizando el método CRISP DM para el sector automotriz en el Ecuador**. Previo a la obtención del título de Magister en Big Data y Ciencia de Datos.

1. Declaro tener pleno conocimiento de la obligación que tienen las instituciones de educación superior, de conformidad con el Artículo 144 de la Ley Orgánica de Educación Superior, de entregar el respectivo trabajo de titulación para que sea integrado al Sistema Nacional de Información de la Educación Superior del Ecuador para su difusión pública respetando los derechos de autor.
2. Manifiesto mi voluntad de ceder a la Universidad Tecnológica Israel los derechos patrimoniales consagrados en la Ley de Propiedad Intelectual del Ecuador, artículos 4, 5 y 6, en calidad de autor@ del trabajo de titulación, quedando la Universidad facultada para ejercer plenamente los derechos cedidos anteriormente. En concordancia suscribo este documento en el momento que hago entrega del trabajo final en formato impreso y digital como parte del acervo bibliográfico de la Universidad Tecnológica Israel.
3. Autorizo a la SENESCYT a tener una copia del referido trabajo de titulación, con el propósito de generar un repositorio que democratice la información, respetando las políticas de prosperidad intelectual vigentes.

Quito D.M., marzo de 2025

Firmado electrónicamente por:
DIANA LISSETH CUEVA VILLARROEL

_____

Firma

Tabla de contenidos

APROBACIÓN DEL TUTOR..... ii
APROBACIÓN DEL TUTOR.....iii
DECLARACIÓN DE AUTORIZACIÓN POR PARTE DEL ESTUDIANTE iv
INFORMACIÓN GENERAL 1
Contextualización del tema 1
Problema de investigación 2
Objetivo general..... 3
Objetivos específicos..... 3
Vinculación con la sociedad y beneficiarios directos: 3
CAPÍTULO I: DESCRIPCIÓN DEL PROYECTO 5
1.1. Contextualización general del estado del arte 5
1.2. Proceso investigativo metodológico 6
1.2.1 Diseño Metodológico 6
1.2.2. Población y Muestra 7
1.2.3. Instrumento de Recolección de Datos 7
1.3. Análisis de resultados 8
CAPÍTULO II: PROPUESTA 16
2.1. Fundamentos teóricos aplicados 16
2.2. Descripción de la propuesta 18
a. Estructura General 18
b. Explicación del Aporte y Comparación de Metodologías 21
2.3. Validación de la propuesta 24
2.4. Matriz de articulación de la propuesta 27
CONCLUSIONES..... 30
RECOMENDACIONES..... 31
Bibliografía..... 32
ANEXO 1 33
ANEXO 2 37
ANEXO 3 45

Índice de tablas

Tabla 1 Resumen de respuestas de la encuesta realizada. 15
Tabla 2 Comparación de Metodologías de Minería de Datos 21
Tabla 3 Comparación de Herramientas de Procesamiento de Datos.. 22
Tabla 4 Comparación de Herramientas de Visualización. 23
Tabla 5 Comparación de Modelos Predictivos.. 23
Tabla 6 Validación de especialistas. 26
Tabla 7. Matriz de articulación. 27

Índice de figuras

Figura 1 Resultados de la pregunta 1. 8
Figura 2 Resultados de la pregunta 2 9
Figura 3 Resultados de la pregunta 3. 10
Figura 4 Resultados de la pregunta 4 10
Figura 5 Resultados de la pregunta 5 11
Figura 6 Resultados de la pregunta 6. 12
Figura 7 Resultados de la pregunta 7 12
Figura 8 Resultados de la pregunta 8. 13
Figura 9 Resultados de la pregunta 9. 14
Figura 10 Resultados de la pregunta 10. 14
Figura 11 Estructura general de la propuesta. Metodología Crisp DM. 19

INFORMACIÓN GENERAL

Contextualización del tema

En la industria automotriz en Ecuador es crucial planificar y optimizar las ventas para garantizar el crecimiento y la estabilidad de las compañías involucradas en este sector tan competitivo. Hoy en día factores como el comportamiento del consumidor y los avances tecnológicos influyen significativamente en la demanda de vehículos; además de las regulaciones gubernamentales y la disponibilidad de financiamiento. A pesar de esto, siguen existiendo distribuidoras que continúan utilizando métodos convencionales para gestionar sus ventas. Esto limita la capacidad de adaptación al mercado cambiante y afecta su rentabilidad de manera negativa.

En este escenario específico el emplear análisis predictivo fundamentado en datos científicos proporciona a las compañías una ventaja competitiva al posibilitar la identificación de patrones de compra previstos estimando la demandante futura y mejorando la administración del stock. (Jaramillo, 2021). La metodología CRISP-DM es muy útil para este tipo de casos ya que brinda un enfoque con más organización para el análisis de datos abarcando inicialmente con la comprensión del negocio y llegando hasta la evaluación de los modelos predictivos (IBM, 2020).

Aunque el análisis predictivo tiene un gran potencial en la industria automotriz, todavia las empresas se encuentran enfrentando dificultades al recolectar asi como al procesar los datos, lo que les complica la implementación de estrategias basadas en Big Data y Machine Learning (Galán, 2015). Por este motivo, este estudio se enfoca en poder abordar dichos desafíos para asi mejorar la precisión en las predicciones de ventas logrando contribuir en la toma de decisiones (Cali, 2022).

Para la industria automotriz se ha hecho cada vez más relevante el poder emplear modelos predictivos actualmente, en especial al recurrir a enfoques como CRISP-DM, el cual ha sido utilizado en minería de datos, proporcionando una estructura ordenada tanto como para explorar, modelar asi como evaluar datos importantes para la toma de decisiones estratégicas. En este sentido, las empresas y los vendedores del sector automotriz se basan en el análisis predictivo para tener una herramienta que permite mejorar la gestión del stock y elevar la satisfacción del cliente al minimizar los problemas causados por la falta de productos en existencia (Galán, 2015).

En la actualidad, la tecnología, tiene mucho valor, como es el caso de Apache Spark, ya que permite el manejo de cantidades de datos de alto flujo, esto de manera eficiente. Las características de SPARK como lo son la velocidad del procesamiento y su distribución es la parte clave para poder realizar análisis predictivos en el sector automotriz (García, 2022). Con los resultados obtenidos se puede obtener información como son los diferentes hábitos de compra

y también permite analizar la tendencia de la demanda, brindando asi el poder a las empresas de que tengan un mejor manejo en su stock y recursos (García, 2022).

En el tema de la gestión en el campo empresarial, se puede ver que este tipo de modelos de predicción es muy eficiente para la administración de las ventas. En el campo automotriz, se han adoptado ciertos tipos de predicciones de las diferentes demandas, logrando brindar la información necesaria para optimizar el almacenamiento y tener n inventario bien administrado, evitando tener productos de baja rotación o muchos artículos que no se venden, lo cual ayuda a evitar pérdidas económicas. Gracias a los datos organizados asi como las tecnologías para su análisis se pueden tomar decisiones con más precisión y contribuir al desarrollo de las empresas (Robles y Valverde, 2021).

Problema de investigación

En el Ecuador, en el campo automotriz, se tienen muchos desafíos en temas como son la planificación o poder predecir las ventas, considerando los cambios que pueden influir como lo son las preferencias de los clientes o el cambio de mercado. Muchas agencias distribuidoras y fabricantes todavía confían en métodos convencionales para tomar decisiones empresariales, lo que provoca complicaciones en la gestión de inventarios, dificulta la aplicación de estrategias de marketing eficientes y afecta el cálculo de ganancias.

El uso del análisis predictivo se ha vuelto fundamental para mejorar las ventas al poder reconocer esquemas de conductas y proyectar tendencias futuras en función de datos previos (Cali, 2022). A pesar de esto, su aplicación en el ámbito empresarial, aún es reducida debido a la falta que se tiene de un método que sea organizado y que ayude a agilizar la integración de modelos predictivos (Jaramillo, 2021).

Ya que la metodología CRISP-DM tiene un organización que ayuda a poder recopilar información par luego analizarla y asi poder modelar los resultados, dando una precisión de mayor efectividad en las predicciones de las ventas (IBM, 2020). Un punto clave puede ser el lograr reducir cierta incertidumbre con el uso de esta metodología en el campo automotriz, mejorando lo que es la planificación y la gestión de los inventarios (Pinto, 2020).

El desafío mas grande en el presente estudio es lograr la creación asi como la aplicación de un modelo de predicción que utilice la metodología CRISP-DM y asi poder aumentar la planificación del campo comercial y la mejora de la toma de decisiones del sector automotriz, teniendo como objetivo mejorar el enfoque con mejor estructura del análisis de los datos y obtener grandes beneficios para el mercado que es cambiante (Galán, 2015).

Objetivo general

Desarrollar un análisis predictivo de ventas utilizando el método CRISP-DM para el sector automotriz en Ecuador.

Objetivos específicos

- Contextualizar los fundamentos teóricos de la metodología CRISP-DM y de las herramientas de análisis de datos, asi como su aplicabilidad en la predicción de ventas de vehículos en Ecuador.
- Diagnosticar los parámetros necesarios para obtener los datos más importantes sobre las ventas de vehículos en Ecuador.
- Desarrollar el proceso de tratamiento de datos y análisis predictivo de las ventas mediante la metodología CRISP-DM para su visualización en Power BI.
- Validar la presente propuesta mediante criterio de especialistas.

Vinculación con la sociedad y beneficiarios directos:

Este proyecto de titulación planteado tiene como objetivo crear un panel interactivo que presente las estimaciones de ventas de vehículos para el año 2025 y 2026 en Ecuador; destacando las provincias según su actividad comercial y los modelos más solicitados en el mercado automotriz local. Aunque el enfoque principal radica en la presentación visual de información estadística detallada; su repercusión trascendería al beneficiar a la industria automotriz regional junto a la comunidad académica y al sector público en general.

Para su desarrollo se empleará la metodología CRISP-DM que es ampliamente utilizada en el análisis predictivo para garantizar un enfoque estructurado en la recopilación y limpieza de datos (Galán, 2015). Esta herramienta ayudará en la toma de decisiones estratégicas sobre la distribución y comercialización de vehículos al analizar tendencias y patrones de compra (Pinto, 2020).

Impacto en la sociedad y sus contribuciones

Formación: se plantea la realización de distintos talleres asi como de conferencias que tengan como objetivo a las personas emprendedoras asi como los estudiantes, los cuales se pueden beneficiar con educación en el análisis de datos y las diferentes herramientas de big data (Robles & Valverde, 2021). La realización de publicaciones académicas puede ser utilizadas como material técnico para utilizarlo a modo de guías para las futuras investigaciones en el campo de la ciencia de datos en el sector automotriz (Jaramillo, 2021).

Optimizar la industria automotriz: tanto las empresas como los diferentes distribuidores van a tener que mejorar su enfoque en el campo comercial y poco a poco aumentar un mejor manejo de la administración de sus proyecciones de ventas y lograr una optimización en su operatividad.

Planificación gubernamental: esta herramienta va a proporcionar puntos clave para la creación de políticas en el tema de la regulación vehicular (Jaramillo, 2021).

Las compañías y los vendedores podrán ajustar sus enfoques comerciales y perfeccionar la gestión de inventarios según las proyecciones de ventas para mejorar su eficiencia operativa (Jaramillo, 2021).

Apoyo a la planificación gubernamental: Esta herramienta brindará datos fundamentales para que los entes locales elaboren políticas relacionadas a la regulación vehicular y movilidad; alineando estrategias a través de información actualizada (Jaramillo, 2021).

Beneficiarios directos

Industria automotriz - Mejorando la planificación de inventarios y las estrategias comerciales.

Comunidad universitaria dedicada a la capacitación en análisis de datos aplicado a la industria automotriz.

Entidades gubernamentales utilizan información para regular el tráfico de vehículos y tomar decisiones fundamentadas.

CAPÍTULO I: DESCRIPCIÓN DEL PROYECTO

1.1. Contextualización general del estado del arte

En diversos sectores el análisis predictivo de datos ha aumentado su importancia, esto también se puede observar en la industria automotriz. El desarrollo de modelos predictivos los cuales son utilizados para identificar tendencias en las ventas de vehículos en Ecuador ha crecido mediante el aumento de la información digital según Jaramillo en el 2021. Para pronosticar la evolución en el parque automotriz Pontón utilizo técnicas de minería de datos, mediante modelos como STLM y ARIMA, lo cual demostró la eficiencia de un análisis predictivo en la toma de decisiones y estrategias en el sector automotriz (Jaramillo, 2021).

Por otro lado, Cali en el 2022 destaca el uso de redes neuronales convolucionales (CNN) ha permitido mejorar la precisión en la estimación de demanda en la predicción de ventas de productos de primera necesidad. La metodología CRISP-DM fue utilizada para la investigación de Nájera para poder estructurar el proceso de análisis, obteniendo resultados inmejorables al ajustar enormes parámetros en los modelos predictivos. Al realizar estos análisis podemos ver la importancia del machine learning es muy útil para la eficiencia comercial y poder manejar de mejor manera un inventario (Cali, 2022).

Galán en el 2015 estudió como usar la minería de datos para mejorar la planificación estratégica en entornos universitarios, para ello utilizo la metodología CRISP-DM en entornos universitarios, esta metodología gracias a su estructura flexible permite que datos de grandes volúmenes se puedan analizar y explorar de manera. Su investigación propone que este modelo no solo sirve para un ámbito determinado como el entorno de universidades, es muy adaptable diferentes sectores, a partir de datos históricos se puede realizar una amplía extracción de conocimientos (Galán, 2015).

Existen varios estudios que muestran que los modelos predictivos son altamente eficientes y utilizados en varios sectores, en los cuales sea tenido resultados satisfactorios en cuanto a la estimación de tendencias. Sin embargo, la investigación de Pinto en el 2020 realiza una investigación donde demuestra que la selección de algoritmos es decisiva para la eficiencia de los modelos predictivos. Utilizo regresión y clustering en su investigación para optimizar pedidos de repuestos, lo cual demostró que la eficiencia de los modelos se basa en la correcta elección de técnicas demostrando que la correcta elección de técnicas analíticas influye directamente en la eficiencia de los modelos (Pinto, 2020).

Existen desafíos en la interpretación de validación de los resultados a pesar de los avances en la aplicación de modelos predictivos. Robles y Valverde en el 2021 en la empresa GGP Automotriz analizaron la predicción de ventas de accesorios hallaron que el aprendizaje profundo optimizaba la precisión de las predicciones, también concluyeron que es muy

importante validar los modelos con datos recientes para evitar errores, esto se debe a que los modelos se pueden volver menos precisos con el tiempo si no se actualizan los datos (Robles y Valverde, 2021).

Incluso en industrias complejas como lo es la automotriz, se ha demostrado en diferentes investigaciones que existen metodología de análisis predictivo muy eficiente. Jaramillo en el 2021 demostró que los márgenes de error son mínimos al utilizar modelos STLM y ARIMA. De manera similar, Cali en el 2022 corroboró que la combinación de CNN y CRISP-DM mejora la predicción de ventas, esto ha permitido reducir las pérdidas por error en la estimación de la demanda (Jaramillo, 2021), (Cali, 2022).

El análisis de Pinto en el 2020 indica que los algoritmos tradicionales pueden presentar dificultades, para obtener patrones no lineales en las ventas de repuestos, de igual manera Galán en el 2015 destacó que para mejorar la eficiencia en la toma de decisiones es necesario implementar herramientas analíticas más avanzadas (Cali, 2022), (Galán, 2015).

Con las investigaciones realizadas podemos observar que aún nos falta mejorar, lo cual podemos hacerlo con una combinación de aprendizaje automático y métodos estadísticos que son modelos hídricos para mejorar la predicción de ventas. Robles y Valverde en el 2021 sugieren que ante incertidumbres del mercado se podría optimizar el análisis de tendencia usando redes neuronales recurrentes (RNN) (Robles y Valverde, 2021).

Además, la investigación futura debería centrarse en la automatización del proceso de ajuste de modelos para reducir la intervención manual y mejorar la eficiencia en la actualización de predicciones. La metodología CRISP-DM, mencionada en múltiples estudios, sigue siendo una referencia clave, pero se requiere mayor exploración en su aplicación con tecnologías emergentes, como inteligencia artificial y big data (Galán, 2015).

1.2. Proceso investigativo metodológico

Debido a la necesidad de obtener resultados medibles y replicables en un tiempo limitado durante el proceso de investigación se ha decidido adoptar un enfoque cuantitativo. Esta elección se basó en que el análisis de datos resultará más sencillo de llevar a cabo y al ser los resultados cuantificables permitirá una mayor precisión incluso en el ámbito empresarial. Para garantizar la validez y confiabilidad de la información recopilada se emplean técnicas estadísticas para asegurar una recolección precisa de datos (Jaramillo, 2021).

1.2.1 Diseño Metodológico

Ya que el análisis se centra en datos históricos de ventas y la percepción de los expertos del sector automotriz sobre la importancia del análisis predictivo en la toma de decisiones

esto quiere decir que el diseño de la investigación es descriptivo y no es experimental. Por ende, se utilizará la técnica de encuesta estructurada con escala de Likert para realizar la recolección de información primaria, para poder medir la percepción de los encuestados referente a los modelos de análisis predictivos en la optimización de estrategias comerciales (Cali, 2022).

1.2.2. Población y Muestra

Para el estudio en cuestión se tomará en consideración a personas que trabajen en ventas y análisis comercial en agencias de automóviles en Ecuador. Se utilizará un método de muestreo no aleatorio por conveniencia que permitirá seleccionar a los participantes basándose en su accesibilidad y disponibilidad; dando prioridad a aquellos que cuenten experiencia tanto en herramientas analíticas como en la venta de vehículos. El muestreo por conveniencia resulta apropiado para investigaciones exploratorias cuyo objetivo sea obtener información específica de una población sin necesidad de representatividad estadística (Hernández y otros, 2014).

Para la muestra de la encuesta se elegirá un grupo de alrededor de 20 individuos que pertenezcan al grupo definido y que cuenten con la experiencia adecuada para comprender y evaluar los resultados obtenidos en el estudio realizado mediante la encuesta aplicada. Este enfoque asegura que los datos recolectados sean analizados por especialistas en el tema para poder obtener una interpretación más precisa de los descubrimientos.

1.2.3. Instrumento de Recolección de Datos

Esta encuesta fue diseñada para poder evaluar las diferentes preferencias asi como las demandas que existen en el sector automotriz **(ANEXO 1)**. Se ha configurado mediante preguntas de opción múltiple y una escala de Likert que va del 1 al 5; esto posibilita la evaluación del grado de satisfacción y percepción de los clientes acerca de la disponibilidad y calidad de los productos. La utilización de esta metodología facilita la recolecta de datos cuantificables para realizar un análisis más preciso sobre las tendencias del mercado (Robles y Valverde, 2021).

La encuesta se llevó a cabo de manera virtual utilizando una plataforma en línea para asegurar la accesibilidad y obtener una mayor tasa de respuesta. Esto no solo facilitó la recolección de información en tiempo real, sino que también garantizó que las respuestas fueran interpretadas de manera homogénea (Robles y Valverde 2021).

1.3. Análisis de resultados

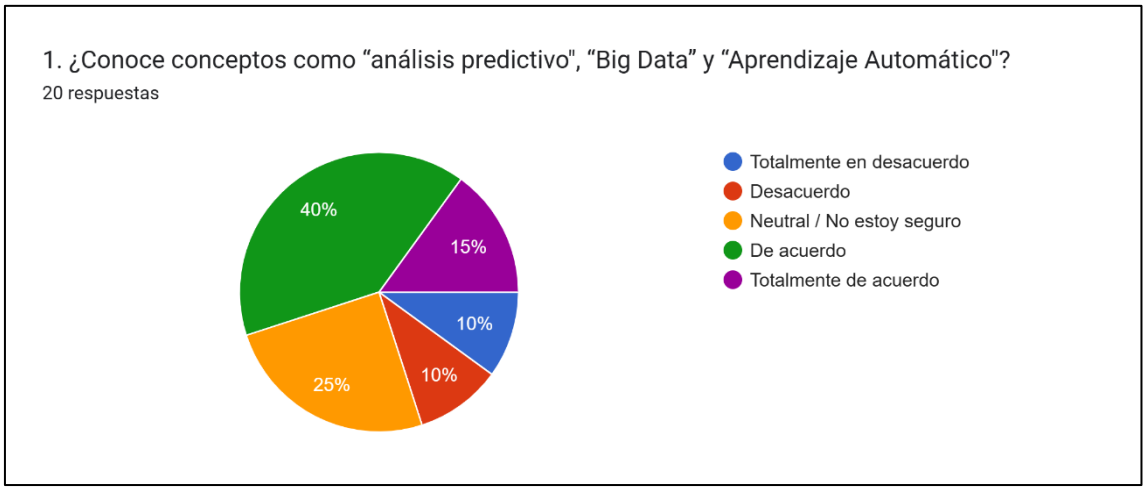
Para el análisis de las encuestas realizas, a través del método de Likert, se establecieron 5 respuestas, las cuales nos van a ayudar a interpretar de manera más efectiva los resultados obtenidos con la muestra considerada. Las respuestas de selección son las siguientes:

- Totalmente en desacuerdo
- Desacuerdo
- Neutral / No estoy seguro
- De acuerdo
- Totalmente de acuerdo

1. ¿Conoce conceptos como "¿Análisis Predictivo", "Big Data" y "Aprendizaje Automático"?

La mayoría de los participantes encuestados muestran cierto nivel de entendimiento sobre estos conceptos: 5 indicaron estar "Totalmente de acuerdo" y 7 dijeron estar simplemente de acuerdo". Sin embargo, un número menor de personas (3 neutrales y 5 en desacuerdo), no tienen total certidumbre o desconocimiento acerca de estos términos como se puede visualizar en la **Figura 1**. A pesar de que existe un nivel aceptable de conocimiento sobre estas tecnologías, aún se observa una falta de comprensión en el sector automotriz. Estos hallazgos indican la importancia de mejorar la formación en Big Data y Machine Learning en particular para su uso en la predicción de ventas y toma de decisiones estratégicas.

Figura 1
Resultados de la pregunta 1.

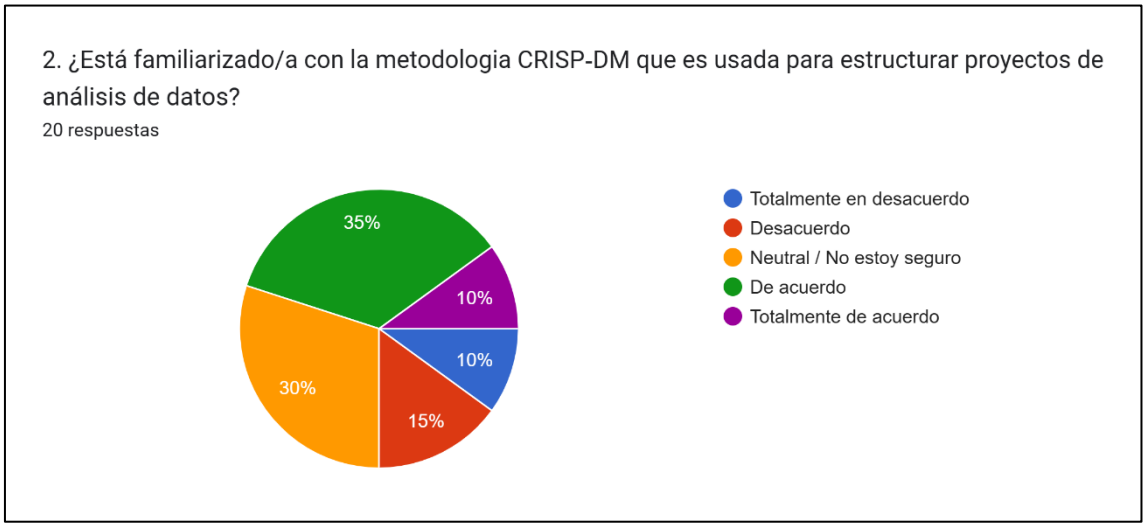


2. ¿Está familiarizado/a con la metodología CRISP-DM?

Solo 4 participantes declararon estar totalmente de acuerdo en comprender esta metodología; por otro lado 6 respondieron que estaban de acuerdo en cierta medida. Un grupo significativo de personas mostró incertidumbre o desconocimiento sobre el tema, 4 personas

señalaron neutralidad y 3 expresaron desacuerdo-. A pesar de ser un marco comúnmente empleado en la minería de datos, CRISP - DM no es muy reconocido por los encuestados. En la **Figura 2** se subraya la importancia de difundir y capacitar más en esta metodología para que los especialistas puedan comprender su relevancia en la planificación de proyectos de análisis de datos y predicción de modelos.

Figura 2
Resultados de la pregunta 2

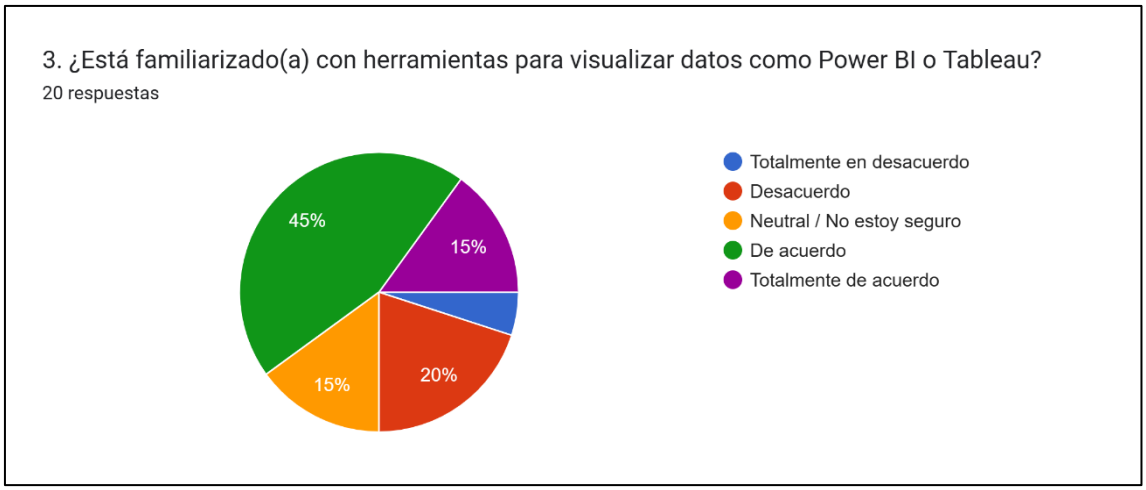


3. ¿Está familiarizado/a con herramientas para visualizar datos como Power BI o Tableau?

Aunque hubo 3 personas que dijeron estar totalmente de acuerdo y 7 que estuvieron de acuerdo, todavía hay un grupo significativo de encuestados que no están seguros o no conocen completamente estas herramientas (con 5 en posición neutral y 5 en desacuerdo). En la **Figura 3** se ve que pesar del amplio uso de Power BI y Tableau en la visualización de datos, su conocimiento y uso dentro del sector automotriz sigue siendo limitado aún. Brindar capacitación a los profesionales para dominar estas herramientas podría mejorar la interpretación de las tendencias de ventas y facilitar la toma de decisiones basadas en datos.

Figura 3

Resultados de la pregunta 3.

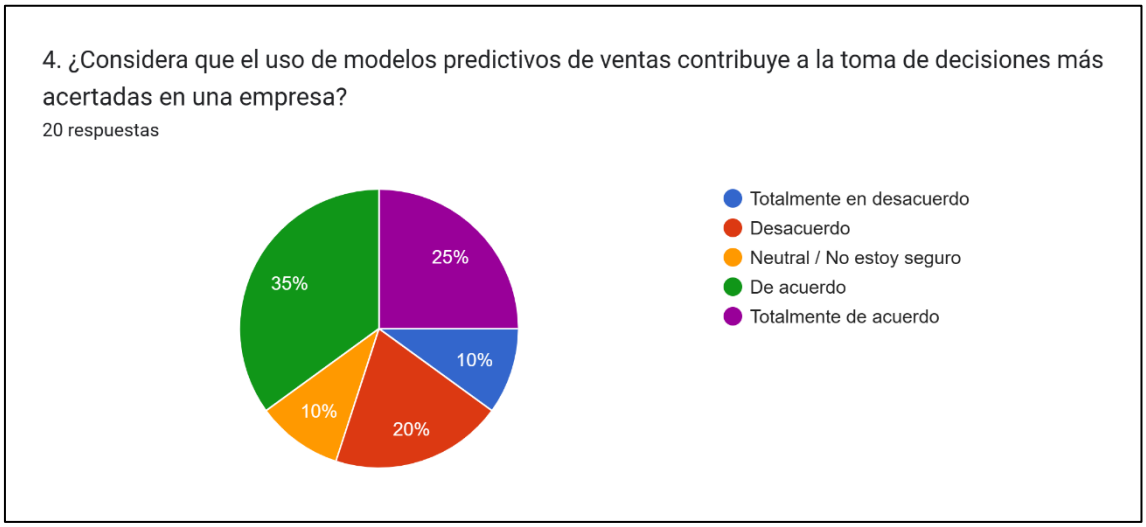


4. ¿Considera que el uso de modelos predictivos de ventas contribuye a la toma de decisiones más acertadas en una empresa?

Seis participantes expresaron su total acuerdo y cinco estuvieron de acuerdo en general en cuanto a la percepción positiva sobre el impacto de los modelos predictivos en la toma de decisiones; solamente tres manifestaron incertidumbre o desacuerdo al respecto. En la **Figura 4** se nota que existe una fuerte convicción de que los modelos predictivos pueden contribuir a mejorar la toma de decisiones en el entorno empresarial; no obstantes para su aplicación efectiva es fundamental que quienes tomen las decisiones comprendan e interpreten adecuadamente los resultados de tales modelos asegurando así que las predicciones sobre ventas sean utilizadas estratégicamente.

Figura 4

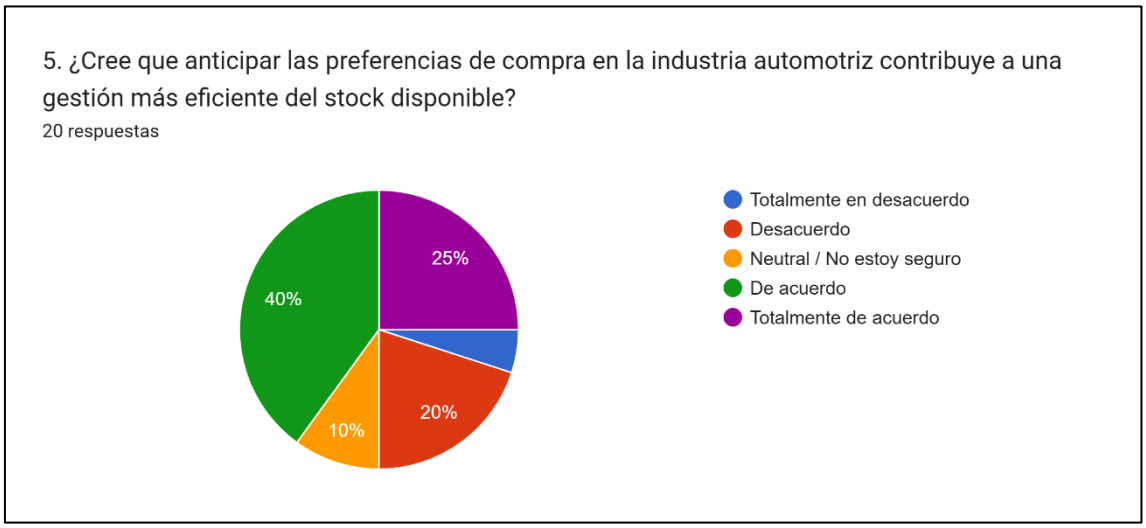
Resultados de la pregunta 4



5. ¿Cree que anticipar las preferencias de compra en la industria automotriz contribuye a una gestión más eficiente del stock disponible?

Hubieron 6 encuestados que estuvieron completamente de acuerdo y 5 que estuvieron de acuerdo en mostrar un consenso sólido sobre la importancia de prever las preferencias de compra. En la **Figura 5** se puede evidenciar que hay un reconocimiento extendido sobre la utilidad de anticipar la demandade los clientes para optimizar el inventario en la industria automotriz. La puesta en marchade modelos predictivos permitiría a las compañías evitar el exceso de producción o la escasez, repercutiendo en una mejora de la rentabilidad y eficiencia operativa.

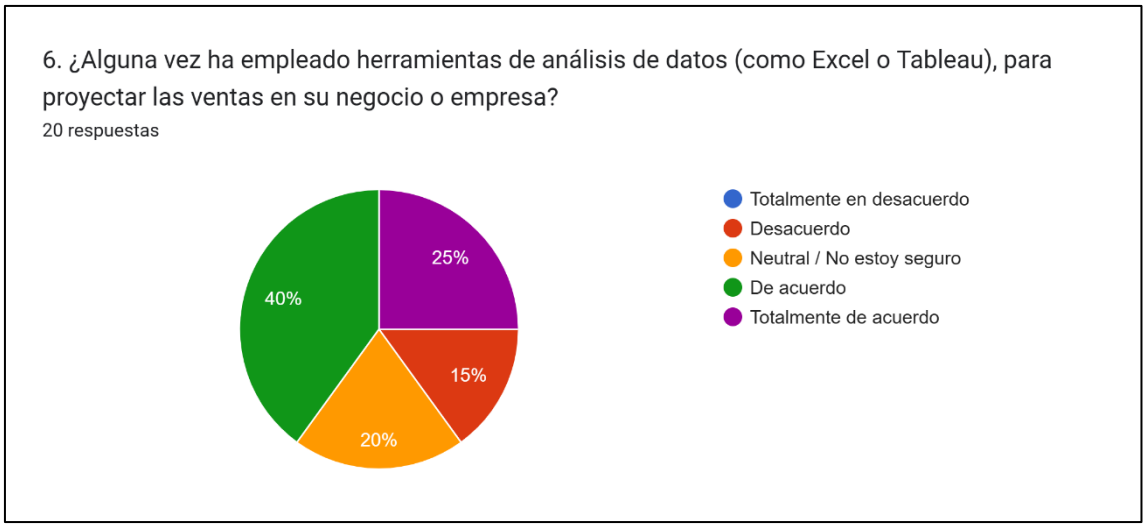
Figura 5
Resultados de la pregunta 5



6. ¿Alguna vez ha empleado herramientas de análisis de datos para proyectar ventas en su negocio o empresa?

Solo 5 personas respondieron “Totalmente de acuerdo", mientras que otras 6 expresaron su conformidad, lo que sugiere que la práctica de emplear herramientas analíticas para prever ventas aún no está ampliamente difundida. En la **Figura 6** se nota que a pesar de que algunas empresas ya han integrado herramientas analíticas en sus operaciones, el uso del análisis predictivo aún no es común en la industria automotriz. Esto abre una oportunidad para promover la adopción de dichas herramientas y mejorar la exactitud de las estrategias comerciales.

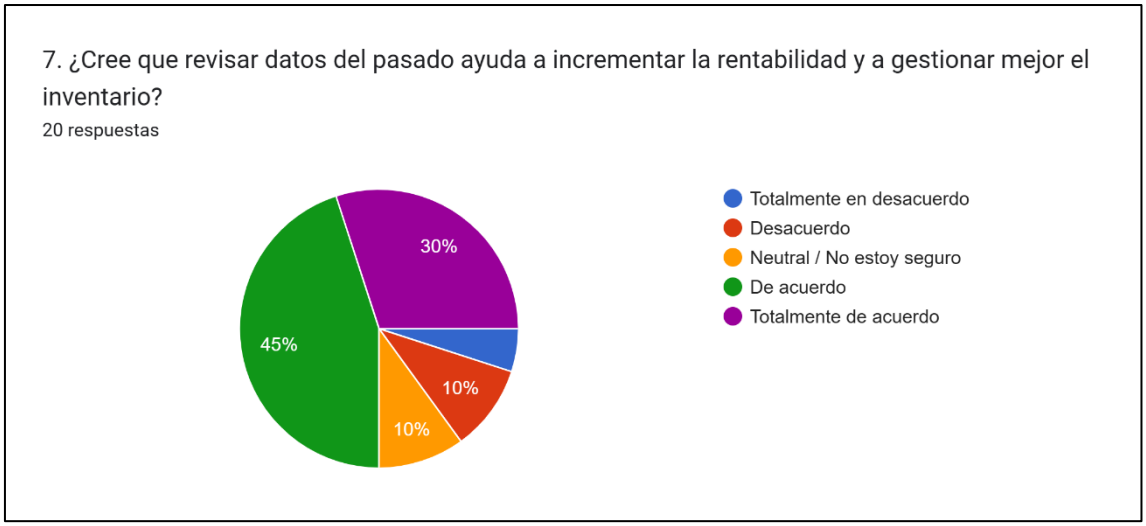
Figura 6
Resultados de la pregunta 6.



7. ¿Cree que revisar datos del pasado ayuda a incrementar la rentabilidad y a gestionar mejor el inventario?

Hubo 7 encuestados que mostraron estar “Totalmente de acuerdo", mientras que 6 indicaron estar “De acuerdo", lo que señala un alto nivel de aceptación hacia esta práctica en particular. La mayoría de los encuestados reconoce la importancia de utilizar datos históricos para aumentar la rentabilidad y gestionar mejor los inventarios en las empresas. En la **Figura 7** se observa que, al emplear modelos basados en información previa, las compañías pueden minimizar pérdidas, optimizar la distribución de vehículos y anticiparse a posibles cambios en la demandante del mercado.

Figura 7
Resultados de la pregunta 7

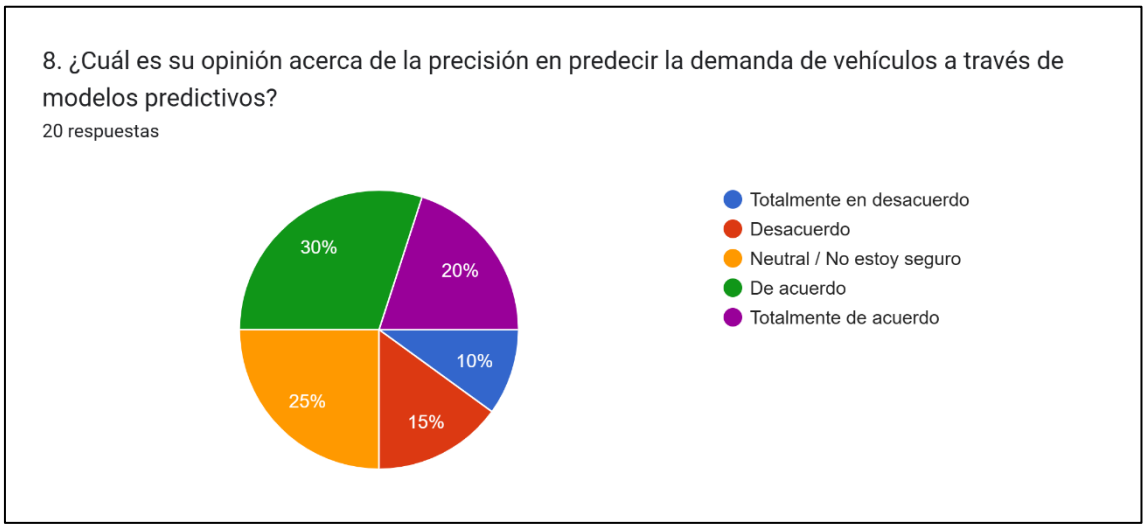


8. ¿Cuál es su opinión acerca de la precisión en predecir la demanda de vehículos a través de modelos predictivos?

Hubo 5 personas que dijeron estar “Totalmente de acuerdo", mientras que 6 manifestaron estar simplemente de acuerdo; solo 3 expresaron neutralidad y pocos demostraron desacuerdo. En la **Figura 8** se percibe cierto nivel de confianza en la capacidad de los modelos predictivos para estimar la demandad de vehículos; sin embargo, persisten dudas acerca de su exactitud lo cual destaca la importancia de evaluaciones periódicas y validaciones utilizando datos reales para fortalecer la confianza en sus resultados.

Figura 8

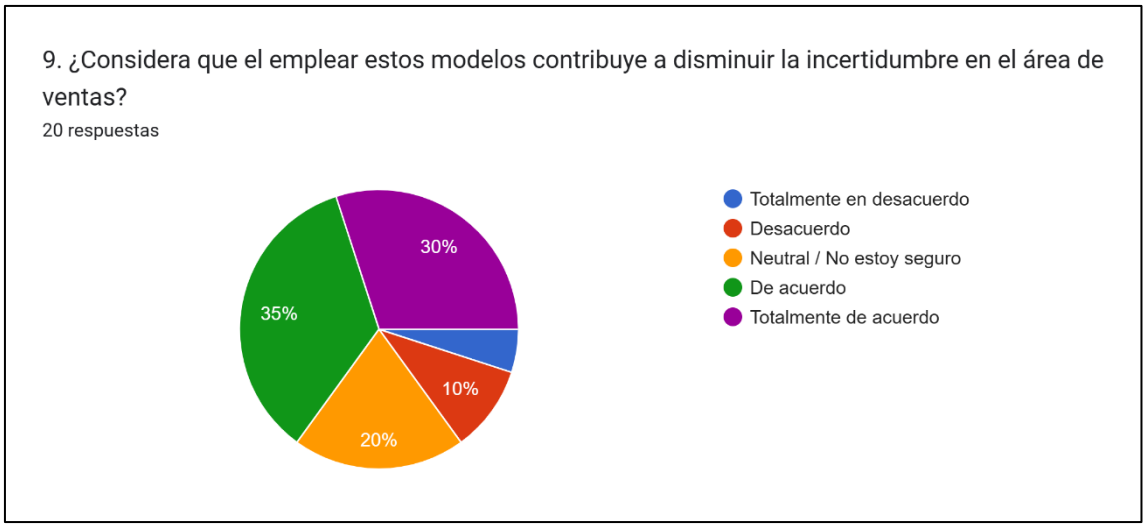
Resultados de la pregunta 8.



9. ¿Considera que el emplear estos modelos contribuye a disminuir la incertidumbre en el área de ventas?

Seis participantes respondieron “Estoy totalmente de acuerdo", mientras que otros seis dijeron “De acuerdo", lo que indica que la mayoría cree en la capacidad de estas herramientas para disminuir el riesgo en las ventas. En la **Figura 9** se percibe un efecto positivo al reducir la incertidumbre mediante el uso de modelos predictivos. Aunque los datos muestran una alta tasa de aceptación, es crucial trabajar en la integración de estas herramientas dentro de las estrategias empresariales para potenciar al máximo su impacto.

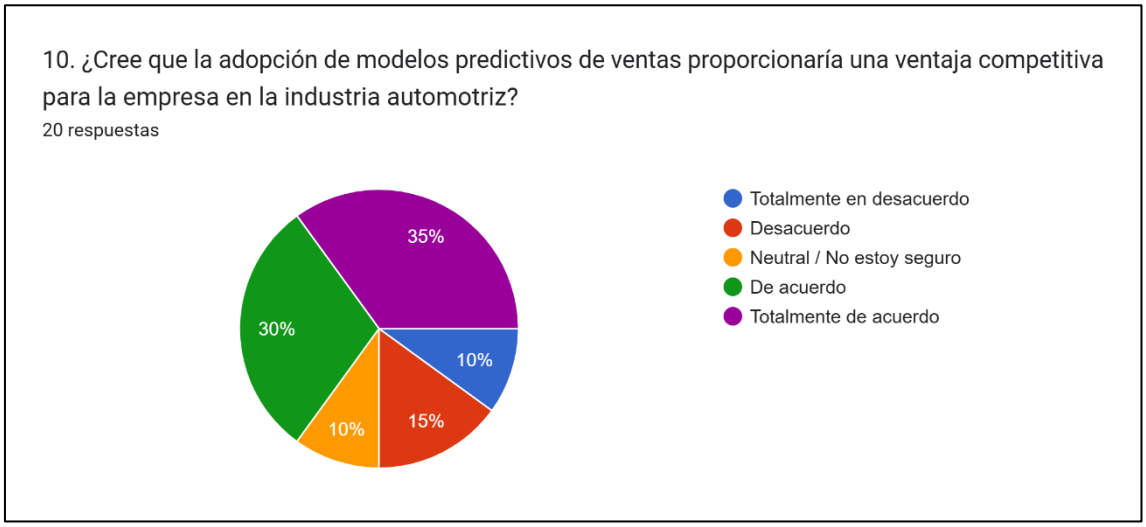
Figura 9
Resultados de la pregunta 9.



10. ¿Cree que la adopción de modelos predictivos proporcionaría una ventaja competitiva en la industria automotriz?

Hubo 7 respuestas que indicaron estar “Totalmente de acuerdo", mientras que 5 expresaron estar “De acuerdo", lo que refleja una opinión positiva sobre la relevancia de estos modelos en el mercado. En la **Figura 10** hay un amplio acuerdo en que los modelos predictivos pueden brindar una ventaja competitiva significativa en la industria automotriz. Las empresas que incorporen análisis de datos en su planificación comercial podrán adaptarse más ágilmente a cambios en la demanda y optimizar sus estrategias de ventas.

Figura 10
Resultados de la pregunta 10.



A continuación, en la **Tabla 1** se puede evidenciar las respuestas recolectadas de la encuesta realizada.

Tabla 1
Resumen de respuestas de la encuesta realizada

#	Pregunta	Totalmente de acuerdo	De acuerdo	Neutral / No estoy seguro	Desacuerdo	Totalmente en desacuerdo
1	¿Conoce conceptos como '¿Análisis Predictivo', 'Big Data' y 'Aprendizaje Automático'?	5	7	3	3	2
2	¿Está familiarizado/a con la metodología CRISP-DM?	4	6	4	3	1
3	¿Está familiarizado/a con herramientas para visualizar datos como Power BI o Tableau?	3	7	5	3	2
4	¿Considera que el uso de modelos predictivos de ventas contribuye a la toma de decisiones más acertadas en una empresa?	6	5	3	3	2
5	¿Cree que anticipar las preferencias de compra en la industria automotriz contribuye a una gestión más eficiente del stock disponible?	6	5	3	3	2
6	¿Alguna vez ha empleado herramientas de análisis de datos para proyectar ventas en su negocio o empresa?	5	6	4	3	2
7	¿Cree que revisar datos del pasado ayuda a incrementar la rentabilidad y a gestionar mejor el inventario?	7	6	3	2	2
8	¿Cuál es su opinión acerca de la precisión en predecir la demanda de vehículos a través de modelos predictivos?	5	6	3	3	2
9	¿Considera que el emplear estos modelos contribuye a disminuir la incertidumbre en el área de ventas?	6	6	3	2	2
10	¿Cree que la adopción de modelos predictivos proporcionaría una ventaja competitiva en la industria automotriz?	7	5	3	2	2

CAPÍTULO II: PROPUESTA

2.1. Fundamentos teóricos aplicados

Análisis Predictivo y Minería de Datos

El uso del análisis predictivo ayuda a prever las tendencias venideras basándose en información previa mediante el empleo de técnicas de minería de datos y aprendizaje automático para la toma de decisiones estratégicas. En la industria automotriz en particular resulta fundamental este enfoque para la planificación de las ventas y la gestión eficiente del inventario (Jaramillo, 2021).

Metodología CRISP-DM

CRISP DM (Proceso Estándar Cruzado para la Minería de Datos en la Industria), es un marco organizado utilizado en proyectos de minería de datos, el cual consta de seis etapas:

Comprensión del negocio, Comprensión de los datos, Preparación de los datos, Modelado, Evaluación y Despliegue. Este método asegura un examen sistemático y repetible en la creación de modelos predictivos (Galán, 2015).

Apache Spark y Big Data

Es una plataforma de distribución de datos que facilitar el manejo efectivo de grandes volúmenes de información. Su capacidad para procesar en memoria lo convierte en mucho más rápido que Hadoop MapReduce (García, 2022).

Apache Spark hace que tareas como estas sean más fáciles:

- Utilización de procesamiento simultáneo para incrementar la rapidez.
- Admite la conexión a diferentes tipos de datos como SQL y NoSQL asi como CSV.
- Puede crecer y ajustarse a grandes cantidades de datos (Benbrahim y otros, 2019).

Facebook Prophet

Esta es una herramienta que realiza el modelado de series temporales que fue creada por la empresa Meta, la cual puede capturar las diferentes tendencias y efectos especiales que se obtienen de datos históricos. Sus características son:

- Automatización al identificar patrones en los datos.
- Gestionar la información que presenta valores inusuales e incidencias externas.
- Modificaciones manuales para aumentar la exactitud de las predicciones.

Prophet se emplee comúnmente para prever tendencias comerciales y la demandada de productos (Jaramillo, 2021).

Power BI

Esta herramienta es de uso empresarial y fue creada por Microsoft la cual facilita el poder de la representación adecuada de datos con el uso de dashboards. Sus características son:

- Conexión a diferentes fuentes de información como Excel, bases de datos en SQL Server y en la nube.
- Funcionalidad avanzada de interacción que facilita el análisis en vivo.
- Habilidad para estructurar datos de manera efectiva y generar informes personalizados que apoyen la toma de decisiones (Jaramillo, 2021).

Metodología SEMMA

SEMMA (Sample, Explore, Modify, Model, Assess) Es un enfoque creado por el Instituto SAS para proyectos de extracción de datos, que constan de cinco etapas:

- Muestreo (Sample): Elección de datos representativos para el análisis.
- Exploración (Explore): Descubrimiento de regularidades y desviaciones.
- Modificación (Modify): Edición y depuración de datos.
- Modelado (Model): Implementación de algoritmos de machine learning.
- Evaluación (Assess): Validación y ajustes de los modelos para incrementar su precisión.

SEMMA se centra en la preparación y modelado de datos sin abordar la comprensión del negocio de la misma manera que lo hace CRISP - DM (Jaramillo, 2021).

Hadoop

Hadoop es una plataforma de software de fuente abierta creada para el manejo y almacenamiento distribuido de cantidades masivas de información, destacando sus elementos fundamentales:

- HDFS (Hadoop Distributed File System): Un sistema para almacenar datos de manera distribuida.
- MapReduce es usada para el procesamiento de forma paralela que utiliza comúnmente los diferentes entornos informáticos.
- YARN (Yet Another Resource Negotiator): Es el encargado de gestionar los diferentes recursos.

El sistema Hadoop facilita la gestión de grandes volúmenes de datos en entornos empresariales y de investigación (Benbrahim y otros, 2019),

MapReduce

Es una herramienta comuna programación distribuida que se creó en la empresa Google, y luego fue implementada por Hadoop y esta herramienta trabaja con 2 diferentes etapas:

- Map: Mapas datos de entrada y los transforman en pares de clave y valor.
- Reduce: Incluye los resultados obtenidos tras el procesamiento del Map y produce su salida definitiva.

Esta herramienta tiene limitaciones a comparación de Apache SPARK, resaltando el tema del procesamiento en memoria (Salas, 2017).

Tableau

Esta herramienta brinda la opción de visualizar los datos asi como la creación de informes que sean interactivos tiene funciones como:

- La capacidad de exploración visual avanzada se logra sin la necesidad de programar previamente.
- Admite la integración de diversas fuentes de información como bases de datos SQL y plataformas en la nube.
- Sencillez en su manejo gracias a una interfaz intuitiva para la creación de paneles de control (Jaramillo, 2021).

ARIMA (Autoregressive Integrated Moving Average)

ARIMA representa un enfoque estadístico que se emplea frecuentemente para prever la evolución de series temporales, y se fundamenta en tres elementos principales:

- AR: Utiliza información previa para hacer predicciones sobre lo que sucederá en el futuro.
- I: Distinto los datos para convertirlos en estacionarios.
- MA: Corrige discrepancias en la predicción basándose en información previa.

ARIMA es una de las técnicas convencionales más empleadas en el análisis de datos; sin embargo, tiene ciertas restricciones al identificar patrones estacionales en contraste a recursos como Facebook Prophet (Jaramillo, 2021).

2.2. Descripción de la propuesta

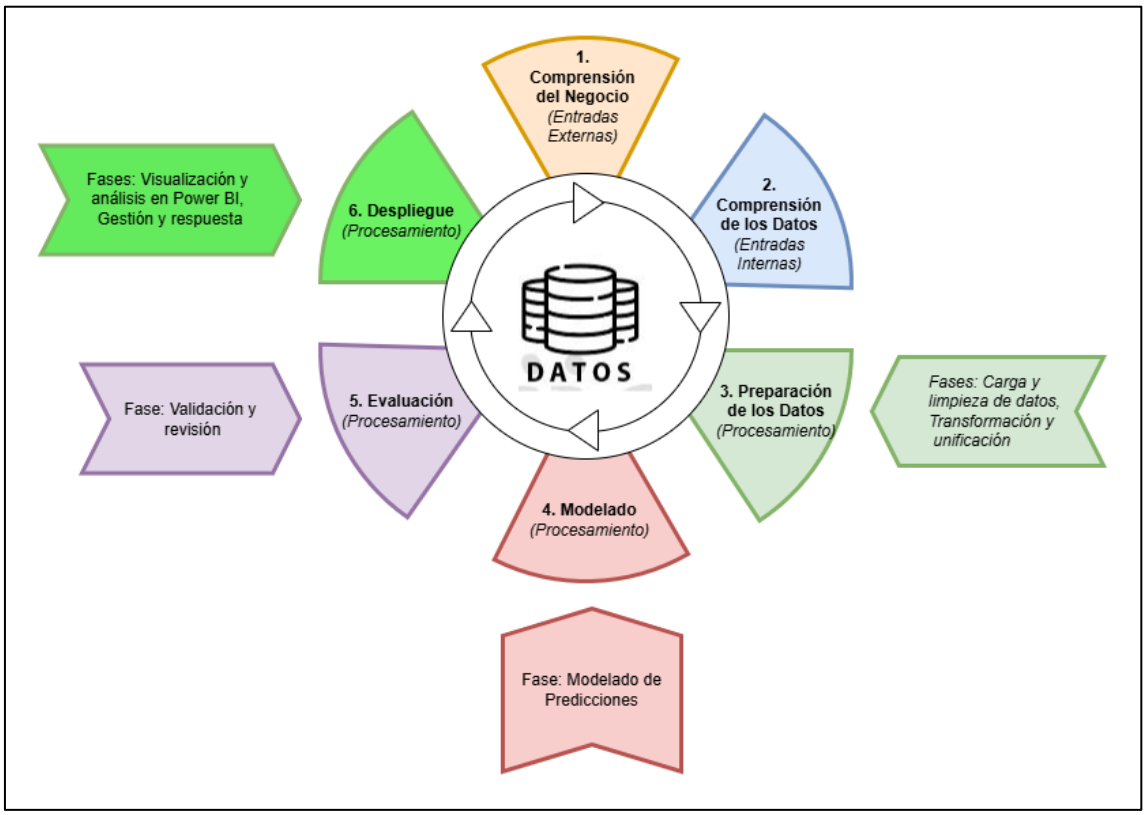
a. Estructura General

El proyecto actual se centra en desarrollar un modelo predictivo para predecir las ventas de automóviles en Ecuador para el año 2025 y 2026. Los métodos utilizados siguen un

enfoque basado en el análisis de datos extensos y la ciencia de datos empleando herramientas avanzadas como Apache Spark, Facebook Prophet y Power BI. El plan del proyecto consta de tres partes fundamentales; entradas, procesamiento(fases) y salidas como se muestra en la **Figura 11**.

Figura 11

Estructura general de la propuesta. Metodología Crisp DM.



Entradas Externas

Información recopilada de instituciones gubernamentales y distribuidores de automóviles experimentados junto a análisis del sector y tendencias actuales en la industria automotriz. Dichos datos albergan información como puede ser precios, clase de vehículos y la demanda a través de la histórica.

Entradas Internas

Datos limpios y organizados en Apache Spark generados a partir de la integración de registros pasados de ventas de automóviles en Ecuador desde 2019 hasta 2024. Durante esta etapa se toman en cuenta elementos clave como el año de venta, la provincia, el modelo, el tipo de vehículo y el avalúo.

Procesamiento (Fases)

Carga y limpieza de datos:

Se combinan y depuran los datos a través de Apache Spark para gestionar de manera efectiva grandes cantidades de información.

Se detectan valores anómalos se eliminan entradas duplicadas se estandarizan variables para mejorar la calidad de los datos.

Transformación y unificación:

Se utilizan métodos de combinación y clasificación de datos a través de Spark SQL para garantizar la coherencia de la información destinada al modelado.

Modelado de Predicciones:

Se emplea Facebook Prophet para realizar proyecciones sobre las ventas en el año 2025 y 2026 teniendo en cuenta las tendencias del mercado y otros elementos como la estacionalidad y factores externos.

Se llevan a cabo diversas repeticiones del modelo para afinar los ajustes de los parámetros y lograr resultados más exactos.

Validación y revisión:

Se contrastan las predicciones con información previa para evaluar la exactitud del modelo.

Se utilizan medidas de error como el RMSE (Error Cuadrático Medio Raíz) y el MAPE (Error Porcentual Absoluto Medio) para evaluar la precisión de las predicciones.

Visualización y análisis en Power BI:

Se ha creado un panel interactivo que facilita la visualización de cómo ha evolucionado el mercado automotriz.

Se puede observar la tendencias de ventas de vehículos desde 2019 hasta 2026.

Provincias donde se registran el top 10 más vendidos y los modelos de vehículos preferidos por los consumidores.

Promedio de valoración de vehículos a lo largo del transcurso del tiempo.

Esto se puede visualizar en el **ANEXO 2.**

Gestión y respuesta:

Se implementan métodos de supervisión para renovar los datos de forma regular y enriquecer el modelo mediante información actualizada.

Salidas Internas

Modelos predictivos basados en datos estructurados y personalizados.

Informe sobre la validación del modelo y los indicadores de precisión.

Panel de control en Power BI con filtros que se ajustan automáticamente y un análisis minucioso.

Salidas Externas

Optimización de la planificación de negocios en agencias y vendedores de automóviles.

Apoyo para tomar decisiones estratégicas fundamentadas en información real.

Utilización del panel de control como recurso de investigación para entidades gubernamentales y académicos que estén interesados en analizar la industria automotriz.

b. Explicación del Aporte y Comparación de Metodologías

Funcionamiento y Empleo de Cada Componente

Este proyecto actual desarrolla un sistema predictivo para pronóstico de ventas de vehículos empleando tecnologías y técnicas avanzadas en Big Data y Ciencia de Datos. La elección de cada elemento se fundamentada en su efectividad, exactitud y capacidad para adaptarse al campo automotriz. Próximamente se expondrá un estudio comparativo sobre las herramientas y metodologías empleadas en el proyecto.

Comparación de Metodologías de Minería de Datos

Si bien hay enfoques similares como SEMMA disponibles para su uso en el análisis de datos y la modelización predictiva de negocios relacionados al sector automotriz se prefiere CRISP DM por su estructura organizada que comienza por comprender las necesidades del negocio y facilitando la posibilidad de iteración entre diferentes etapas del proceso analítico. Este aspecto es vital en contextos donde se debe adaptar los modelos predictivos a las demandas particulares de la industria automotriz. En contraste con SEMMA, CRISP DM incorpora una fase clave enfocada en entender el negocio lo que permite formular hipótesis apropiadas y asegurar que los modelos predictivos sirvan a objetivos comerciales claros como se presenta en la **Tabla 2** (Galán, 2015).

Tabla 2

Comparación de Metodologías de Minería de Datos.

Criterio	CRISP-DM	SEMMA
Inicio del proceso	Comprensión del negocio	Preparación de datos
Estructura	Secuencial con posibilidad de iteración	Lineal
Flexibilidad	Alta	Media
Iteración entre fases	Sí	No
Enfoque	Aplicación en diversos sectores	Enfoque técnico en exploración de datos

Además de esto, el CRISP - DM brinda una estructura flexible y bidireccional, lo que implica que no es obligatorio adherirse fielmente a una secuencia específica de etapas, sino que es posible regresar y perfeccionarse en etapas previas conforme surgen nuevos descubrimientos. Esto favorece la integración de comentarios en cada fase del proyecto, fortaleciendo la solidez del análisis (Galán, 2015).

Comparación de Herramientas de Procesamiento de Datos

En lo que respecta la gestión de datos, puede mencionar diferencias fundamentales entre Apache Spark SQL y Hadoop. Spark SQL se apoya en el procesamiento en memoria lo que permite realizar consultas hasta 11 veces más rápido que Hadoop en ciertos escenarios, mientras que Hadoop se basa en el procesamiento en disco lo cual puede resultar en mayores tiempos de ejecución como se puede visualizar en la **Tabla 3** (Salas, 2017).

Tabla 3

Comparación de Herramientas de Procesamiento de Datos.

Criterio	APACHE SPARK SQL	HADOOP
Velocidad de procesamiento	Alta	Baja
Uso de memoria	Sí	No
Escalabilidad	Alta	Media
Facilidad de uso	Alta	No
Compatibilidad con Big Data	Alta	Alta

Comparación de Herramientas de Visualización

En el campo de las herramientas para analizar y visualizar datos, Power BI y Tableau son dos de las opciones más populares en la actualidad. Según el Cuadrante Mágico de Gartner; Power BI sobresale por su accesibilidad y facilidad de uso ya que ofrece una versión gratuita que permite a los usuarios crear paneles interactivos y realizar análisis de datos sin necesidad de realizar inversiones adicionales. En contraste con esto, Tableau es considerado como una herramienta más sofisticada en términos de personalización y visualización de datos, aunque con un precio más elevado según lo mostrado en la **Tabla 4** (Pinto, 2020).

Tabla 4

Comparación de Herramientas de Visualización.

Criterio	Power BI	Tableau
Facilidad de uso	Alta	Media
Facilidad de uso	Alta	Media
Capacidad de personalización	Media	Alta
Integración con Microsoft	Sí	No
Interactividad	Alta	Alta
Costo	Medio	Alto

Comparación de Modelos Predictivos

En lo que respecta a los modelos predictivos en el análisis de series temporales se emplean dos metodologías destacadas: Facebook Prophet y ARIMA. Prophet se caracteriza por su facilidad de implementación y robustez ante datos irregulares; por lo tanto, es ideal para predecir ventas en sectores como el automotriz que experimentan variaciones estacionales significativas. Por otro lado, ARIMA brinda un mayor control sobre los parámetros estadísticos, pero requiere una intervención manual más extensa para ajustar los modelos y evaluar su rendimiento según se muestra en la **Tabla 5** (Cali, 2022).

Tabla 5

Comparación de Modelos Predictivos.

Criterio	Facebook Prophet	ARIMA
Facilidad de implementación	Alta	Media
Precisión en datos con tendencias	Alta	Alta
Manejo de datos faltantes	Sí	No
Automatización de procesos	Alta	Baja
Adaptabilidad	Alta	Media

En resumen, la unión de CRISP-DM, Power BI, Apache Spark SQL y Facebook Prophet ofrece una metodología completa para realizar análisis predictivos de ventas en la industria automotriz, optimizando tanto la eficiencia en el manejo de datos como la interpretación de los resultados.

Contribución integral del producto:

Fiabilidad de los Datos: Apache Spark SQL garantiza que la información procesada sea confiable.

Exactitud en las Predicciones - Facebook Prophet emplea esta información depurada para crear modelos predictivos altamente precisos.

Entendimiento y toma de decisiones simplificados por Power BI al mostrar los resultados de manera visualmente atractiva para ayudarte a identificar rápidamente las provincias y marcas más solicitadas.

c. Estrategias y/o Técnicas Empleadas

Para desarrollar el producto, se pusieron en práctica las siguientes tácticas:

- Para llevar a cabo la estrategia de Integración de Datos se recurrió a Apache Spark SQL para manejar y depurar cantidades significativas de datos asegurando la calidad de la información que nutre el modelo predictivo.
- En el análisis de series temporales se empleó Facebook Prophet como técnica de modelado predictivo; esta herramienta permite desglosar los datos en tendencias y estacionalidades junto a efectos especiales.
- Enfoque de Representación Visual: Se crearon paneles interactivos en Power BI que posibilitan la visualización de los resultados de manera clara y fácil de entender para facilitar la interpretación y la toma de decisiones.
- Enfoque de Trabajo: La estrategia se fundamenta en el ciclo de vida de CRISP-DM para asegurar una organización coherente y flexible a las exigencias del negocio; facilitando la iteración y la modificación del modelo a medida que se recibe retroalimentación.

2.3. Validación de la propuesta

La Presente validación de la propuesta se lo realizó a través del método de criterios de especialistas. Esto con el fin de tener criterios actuales y con experiencia en el campo del big data. Esto se puede visualizar en el **ANEXO 3**, donde tenemos los documentos de los especialistas con sus criterios.

A continuación, se muestra los perfiles del personal especialista que realizó la validación.

Especialista 1

Validado por: Sebastian Pozo Eugenio

Título obtenido: INGENIERO EN SISTEMAS

C.I.: 1804365276

E-mail: sebastianpe93@hotmail.com

Institución de Trabajo: Importadora Alvarado

Cargo: Jefe de Programación

Años de experiencia en el área: 9

Especialista 2

Validado por: PAUL DANIEL CHERRES CARGUACUNDO

Título obtenido: INGENIERIA EN ELECTRONICA TELECOMUNICACIONES Y REDES

C.I.: 0202134367

E-mail: pcherres@rectima.com.ec

Institución de Trabajo: RECTIMA INDUSTRY CIA. LTDA.

Cargo: JEFE DEL DEPARTAMENTO DE SISTEMAS

Años de experiencia en el área: 5 AÑOS

Especialista 3

Validado por: Jorge Enrique Carrasco U.

Título obtenido: Tecnólogo en sistemas

C.I.: 1803861440

E-mail: jorge86_car@hotmail.com

Institución de Trabajo: Importadora Alvarado Vásconez Ltda.

Cargo: Coordinador de Tecnologías

Años de experiencia en el área: 9 años

En la **Tabla 6** se muestra los resultados de los criterios de los especialistas, asi como sus observaciones y comentarios sobre la propuesta.

Tabla 6
Validación de especialistas

	CALIFICACION	OBSERVACIONES	RECOMENDACIONES
ESPECIALISTA 1	30/35	Resulta de mucha utilidad predecir las ventas de vehículos para todos los giros de negocio dentro de la industria automotriz; especialmente en la venta de repuestos, nos permite tener un panorama claro de las autopartes a importar.	Predecir las ventas dentro de los próximos 6 meses, es de utilidad sobre todo para planificar la importación de repuestos a mediano plazo.
ESPECIALISTA 2	32/35	La fundamentación en datos históricos y proyecciones de ventas es un buen punto de partida para el análisis de los datos, también se pudiera añadir factores externos para completas el estudio y asi tener un resultado mejor en el análisis predictivo.	Se puede evidenciar con los datos proporcionados que el mercado automotriz ecuatoriano se ha mantenido al alza a partir del 2020 y con ciertas preferencias por algunos segmentos, lo cual puede ser aprovechado por las concesionarias para enfocar esfuerzos en estos segmentos y obtener mayores ganancias.
ESPECIALISTA 3	32/35	El análisis se encuentra correctamente segmentado con los clasificadores adecuados para un análisis, el diseño del dashboard me parece que es el apropiado y usa colores que se permite distinguir los elementos usados, los datos presentados son claros y coherentes tomando en cuenta que se analizan también los años de pandemia.	Paralelamente a esta metodología sería interesante usar otra alternativa de predicción con el mismo origen de datos, para compararlos y verificar si el resultado tuviese alguna diferencia.

2.4. Matriz de articulación de la propuesta

En la **tabla 7** se puede visualizar la articulación del producto realizado teniendo como sustentos teóricos, la metodología y técnicas.

Tabla 7

Matriz de articulación.

Fases de CRISP-DM (EJES O PARTES PRINCIPALES)	SUSTENTO TEÓRICO	SUSTENTO METODOLÓGICO	ESTRATEGIAS / TÉCNICAS	DESCRIPCIÓN DE RESULTADOS	INSTRUMENTOS APLICADOS
1. Comprensión del Negocio	La extracción de datos ayuda a encontrar tendencias escondidas y optimizar la toma de decisiones en áreas como la industria automotriz (Jaramillo, 2021).	Revisión de documentos sobre la industria automotriz y las tendencias de venta de vehículos en Ecuador, así como la identificación de los principales factores que influyen en la decisión de compra (Galán, 2015).	Análisis del mercado y estudio de las tendencias en ventas de automóviles, así como la identificación de variables importantes para el sector (año, provincia, modelo, clase de vehículo, avalúo).	Los objetivos del análisis predictivo se definen al determinar la información pertinente para el modelo.	Fuentes externas como informes sobre ventas y estudios del mercado automotriz.
2. Comprensión de los Datos	La utilización de grandes conjuntos de datos y el procesamiento distribuido para manejar grandes cantidades de información en	Utilización de Apache Spark para la exploración de datos, detección de valores anómalos e inspección de la calidad	Extracción y examinación de registros de ventas pasadas (2019 - 2024). Revisión de datos	Se logra una comprensión precisa de la organización y la excelencia de los datos para poder	Apache Spark SQL se utiliza para explorar datos y Pandas para realizar un análisis inicial de los mismos.

	tiempo real (Benbrahim y otros, 2019)	de la información (García, 2022)	ausentes e identificación de discrepancias.	detectar posibles modificaciones requeridas.	
3. Preparación de los Datos	Datos estructurados y no estructurados se procesan en la minería de datos, comparando el uso de Hadoop y Spark para optimizar la información recolectada (Salas, 2017).	Procesamiento de información utilizando Apache Spark SQL: estandarización de variables y depuración de datos innecesarios para organizar la información de manera efectiva en un formato apropiado (García, 2022)	Filtrado y organización de información en Apache Spark para combinar diversas fuentes de datos en un solo conjunto de datos para su análisis.	Se adquiere un conjunto de datos ordenado y estructurado, preparado para utilizar en la creación de un modelo predictivo.	Apache Spark SQL se utiliza para el procesamiento distribuido y Pandas para la manipulación de datos en Jupyter Notebook.
4. Modelado	Facebook Prophet es un modelo de vanguardia utilizado para predecir series temporales que incluyen tanto tendencias como estacionalidades, y se compara su automatización y precisión respecto a ARIMA (Cali, 2022)	Aplicación del modelo predictivo en Facebook Prophet y ajustes de los parámetros para aumentar la exactitud de las predicciones (Jaramillo, 2021)	Entrenamiento del modelo utilizando información recopila desde el año 2019 hasta el 2024, empleando técnicas de validación cruzada y evaluación de errores, así como la comparación de resultados obtenidos	En 2025 y 2026 se prevé una proyección precisa de las ventas de automóviles, optimizando la planificación de inventario y estrategias comerciales.	Facebook Prophet se utiliza para prever tendencias futuras mediante el uso de RMSE y MAPE como medidas de rendimiento.

			contra datos históricos para confirmar la fiabilidad de las predicciones realizadas.		
5. Evaluación del Modelo	Evaluar modelos predictivos a través de métricas estadísticas y analizar la precisión en las predicciones de series temporales (Cali, 2022).	Validación del modelo al contrastar las predicciones realizadas frente a datos previos, y evaluación del error utilizando métricas como RMSE y MAPE (Jaramillo, 2021).	La revisión continua del modelo implica mejorar las predicciones mediante la modificación de los parámetros y la corrección de sesgos en los datos.	Se ha creado un modelo afinado que ofrece predicciones y está listo para su implementación en el análisis de mercado.	Facebook Prophet se utiliza para hacer predicciones mientras que Pandas se emplea para evaluarlas con métricas como el RMSE y el MAPE.
6. Despliegue y Visualización	Power BI es una solución de análisis de datos empresariales que facilitan visualizar la información clave y la toma de decisiones estratégicas, mientras Tableau se destaca por su accesibilidad y variedad de funciones para el análisis de datos (Pinto, 2020).	Creación de paneles interactivos en Power BI para mostrar las predicciones y evolución de las ventas (Jaramillo, 2021).	Creación de gráficos interactivos y filtros a medida para clasificar por región geográfica, modelo y categoría de vehículo. Generación de informes para respaldar la toma de decisiones estratégicas.	Se logra la creación de un panel interactivo en Power BI que simplifica el análisis visual y ayuda a mejorar las estrategias comerciales.	Power BI se utiliza para visualizar datos a través de archivos CSV que contienen predicciones y segmentaciones interactivas.

CONCLUSIONES

El estudio teórico permitió identificar que CRISP-DM es un método organizado y adaptable para la extracción de datos que se utiliza extensamente en proyectos de pronóstico. Se comprobó que recursos como Apache Spark, Facebook Prophet y Power BI son fundamentales para manejar grandes cantidades de información, crear modelos predictivos y representar visualmente los datos. Sin embargo, se ha observado que en la industria automotriz de Ecuador todavía hay una falta de comprensión y aplicación de estas herramientas tecnológicas; esto supone un desafío para introducir modelos predictivos de manera efectiva en el sector.

Parámetros clave identificados para predecir las ventas incluyen el año de referencia de los vehículos en cada provincia junto al modelo específico del automóvil evaluado en el mercado local de automóviles nuevos o usados. Para garantizar la precisión del modelo predictivo se recopilaron datos históricos desde 2019 hasta 2024 para crear una base sólida de información fiable. Además de esto se reconoció la importancia fundamental de normalizar los datos recolectados a fin de eliminar cualquier inconsistencia o valor atípico que pueda afectar negativamente la exactitud de las predicciones futuras.

El procesamiento de datos se llevó a cabo utilizando Apache Spark para mejorar la limpieza y transformación de grandes cantidades de información almacenada eficientemente. Luego se desarrollaron modelos predictivos basados en series temporales utilizando Facebook Prophet y se ajustaron los parámetros mediante métricas como RMSE y MAPE. Finalmente se presentaron los resultados en Power BI a través de paneles interactivos que muestran tendencias de ventas y segmentación por regiones para facilitar la toma de decisiones estratégicas.

Los hallazgos de la encuesta realizada a expertos en la industria automotriz confirman que los modelos predictivos son muy apreciados para la planificación de ventas y la gestión de inventarios. Sin embargo, el estudio reveló una falta de comprensión sobre el enfoque CRISP-DM y herramientas como Power BI y Tableau lo cual indica que su adopción aún es limitada. Además, los especialistas estuvieron de acuerdo en que la implementación de modelos predictivos podría disminuir la incertidumbre en las ventas y otorgaría una ventaja competitiva significativa en el sector automotriz de Ecuador.

RECOMENDACIONES

Se sugiere la realización de cursos de formación y programas de capacitación en CRISP-DM, Apache Spark, Facebook Prophet y Power BI dirigidos a profesionales del ámbito automovilístico. Esto ayudará a reducir la brecha de conocimiento actual y facilitará la adopción de modelos predictivos en concesionarios y empresas del sector. Asimismo, se propone llevar a cabo talleres prácticos en los que los participantes puedan emplear estas herramientas en situaciones cotidianas, mejorando así su capacidad para tomar decisiones estratégicas fundamentadas en datos.

Se sugiere seguir adelante con la automatización del proceso de recopilación y depuración de datos para garantizar que la información utilizada en los modelos sea siempre actualizada y libre de incongruencias importantes expandir las variables analizadas al incorporar datos sobre tendencias económicas comportamiento del consumidor y políticas gubernamentales además se propone la integración de fuentes de datos en tiempo real lo que permitiría que los modelos predictivos se ajustaran dinámicamente los cambios en el mercado automotriz.

Se sugiere considerar la combinación de enfoques tradicionales como ARIMA junto a modelos avanzados de machine learning para mejorar la precisión en las predicciones futuras. Es esencial llevar a cabo pruebas regulares para validar los resultados del modelo y ajustar sus parámetros según datos de la vida real. Además de esto se recomienda explorar técnicas más avanzadas como las redes neuronales recurrentes RNNs para capturar patrones complejos en la demanda de vehículos.

Se sugiere promover la adopción de modelos predictivos a través de programas piloto y casos de éxito documentados que muestren cómo su uso puede mejorar la gestión del inventario y aumentar la rentabilidad en las empresas del sector. Es fundamental que las decisiones empresariales se basen en datos sólidos en lugar de depender únicamente de la intuición. Además, se plantea la idea de establecer alianzas estratégicas entre los distribuidores automovilísticos y las instituciones académicas y de investigación para fomentar el avance en el campo del análisis predictivo.

Bibliografía

Benbrahim, H., Hachimi, H., & Amine, A. (3| de 2019). *Comparison between Hadoop and Spark.* https://www.researchgate.net/publication/333647852_Comparison_between_Hadoop_and_Spark

Cali, C. (2022). *MODELO PREDICTIVO DE LAS VENTAS DE PRODUCTOS DE PRIMERA NECESIDAD EN EL SECTOR COMERCIAL BASADO EN TÉCNICAS DE MACHINE LEARNING*. https://repositorio.utmachala.edu.ec/bitstream/48000/19707/1/TMFIC-2022-SOFT-DE00003.pdf

Galán, V. (9 de 2015). *APLICACIÓN DE LA METODOLOGÍA CRISP-DM A UN PROYECTO DE MINERÍA DE DATOS EN EL ENTORNO UNIVERSITARIO.* https://e-archivo.uc3m.es/rest/api/core/bitstreams/714c5452-962e-44cf-993f-ebb3088d4aa5/content

García, A. (2022). *APACHE SPARK - Big Data.* https://dit.gonzalonazareno.org/gestiona/proyectos/2016-17/Apache_spark-alejandro_palomino.pdf

Hernández, R., Fernández, C., & Baptista, M. (2014). *METODOLOGIA DE LA INVESTIGACION* (6 ed.). INTERAMERICANA EDITORES, S.A. DE C.V. https://doi.org/: 978-1-4562-2396-0

IBM. (2020). *Guía de CRISP-DM de IBM SPSS Modeler*. https://www.ibm.com/docs/es/SS3RA7_18.4.0/pdf/ModelerCRISPDM.pdf

Jaramillo, J. (2021). *ANÁLISIS PREDICTIVO DE DATOS APLICADO AL SECTOR AUTOMOTRIZ.* UNIVERSIDAD TÉCNICA DE MACHALA: https://repositorio.utmachala.edu.ec/bitstream/48000/17228/1/Juan%20Jaramillo.pdf

Pinto, A. (12 de 8 de 2020). *Modelo de analisis predictivo para el mejoramiento de pedidos de repuestos en una importadora automotriz en el Ecuador*. https://repositoriobe.espe.edu.ec/server/api/core/bitstreams/15f62ff4-b37d-4307-be3e-3c7a614cc5ba/content

Robles, M., & Valverde, M. (2021). *SISTEMA DE PREDICCIÓN PARA INCREMENTAR LAS VENTAS DE ACCESORIOS Y REPUESTOS AUTOMOTRICES EN LA EMPRESA GGP AUTOMOTRIZ.* https://repositorio.usmp.edu.pe/handle/20.500.12727/7424

Salas, R. (2017). *Análisis comparativo de herramientas Open Source para Big Data Apache Hadoop y Apache Spark.* https://www.dspace.uce.edu.ec/server/api/core/bitstreams/284b49c3-686c-437a-abde-070ba0431efe/content

ANEXO 1

Análisis de datos

* Indica que la pregunta es obligatoria

1. 1. ¿Conoce conceptos como "análisis predictivo", "Big Data" y "Aprendizaje Automático"? *

Marca solo un óvalo.

- ◯ Totalmente en desacuerdo
- ◯ Desacuerdo
- ◯ Neutral / No estoy seguro
- ◯ De acuerdo
- ◯ Totalmente de acuerdo

2. 2. ¿Está familiarizado/a con la metodologia CRISP-DM que es usada para estructurar proyectos de análisis de datos? *

Marca solo un óvalo.

- ◯ Totalmente en desacuerdo
- ◯ Desacuerdo
- ◯ Neutral / No estoy seguro
- ◯ De acuerdo
- ◯ Totalmente de acuerdo

3. 3. ¿Está familiarizado(a) con herramientas para visualizar datos como Power BI o Tableau? *

Marca solo un óvalo.

- Totalmente en desacuerdo
- Desacuerdo
- Neutral / No estoy seguro
- De acuerdo
- Totalmente de acuerdo

4. 4. ¿Considera que el uso de modelos predictivos de ventas contribuye a la toma de decisiones más acertadas en una empresa? *

Marca solo un óvalo.

- Totalmente en desacuerdo
- Desacuerdo
- Neutral / No estoy seguro
- De acuerdo
- Totalmente de acuerdo

5. 5. ¿Cree que anticipar las preferencias de compra en la industria automotriz contribuye a una gestión más eficiente del stock disponible? *

Marca solo un óvalo.

- Totalmente en desacuerdo
- Desacuerdo
- Neutral / No estoy seguro
- De acuerdo
- Totalmente de acuerdo

6. 6. ¿Alguna vez ha empleado herramientas de análisis de datos (como Excel o Tableau), para proyectar las ventas en su negocio o empresa? *

Marca solo un óvalo.

- Totalmente en desacuerdo
- Desacuerdo
- Neutral / No estoy seguro
- De acuerdo
- Totalmente de acuerdo

7. 7. ¿Cree que revisar datos del pasado ayuda a incrementar la rentabilidad y a gestionar mejor el inventario? *

Marca solo un óvalo.

- Totalmente en desacuerdo
- Desacuerdo
- Neutral / No estoy seguro
- De acuerdo
- Totalmente de acuerdo

8. 8. ¿Cuál es su opinión acerca de la precisión en predecir la demanda de vehículos a través de modelos predictivos? *

Marca solo un óvalo.

- Totalmente en desacuerdo
- Desacuerdo
- Neutral / No estoy seguro
- De acuerdo
- Totalmente de acuerdo

9. 9. ¿Considera que el emplear estos modelos contribuye a disminuir la incertidumbre en el área de ventas? *

Marca solo un óvalo.

- Totalmente en desacuerdo
- Desacuerdo
- Neutral / No estoy seguro
- De acuerdo
- Totalmente de acuerdo

10. 10. ¿Cree que la adopción de modelos predictivos de ventas proporcionaría una ventaja competitiva para la empresa en la industria automotriz? *

Marca solo un óvalo.

- Totalmente en desacuerdo
- Desacuerdo
- Neutral / No estoy seguro
- De acuerdo
- Totalmente de acuerdo

Google no creó ni aprobó este contenido.

Google Formularios

ANEXO 2

DASHBOARD EN POWER BI ACERCA DE EL ANÁLISIS PREDICTIVO DE VENTAS UTILIZANDO EL MÉTODO CRISP DM PARA EL SECTOR AUTOMOTRIZ EN EL ECUADOR

Ing. Diana Cueva

UNIVERSIDAD TECNOLÓGICA ISRAEL

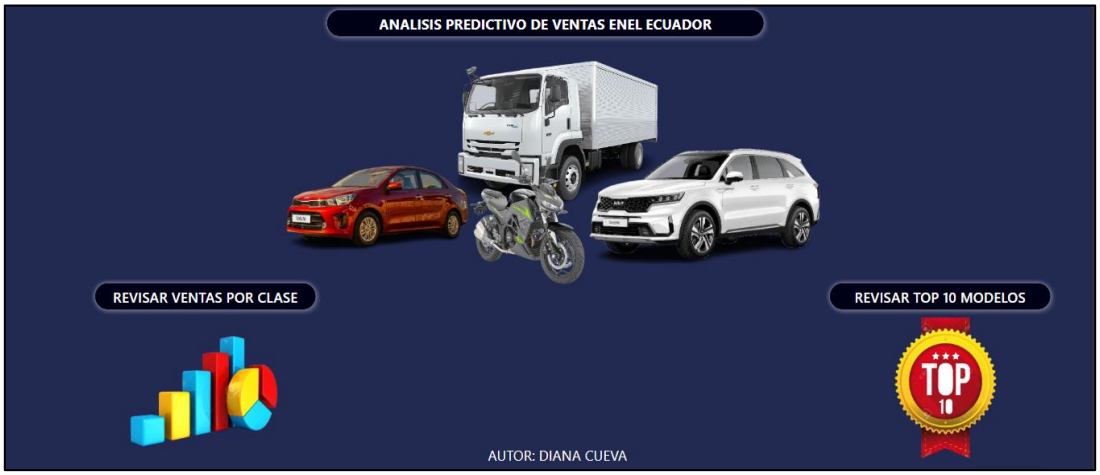
2025

PAGINA PRINCIPAL

Aquí se muestra la portada principal, **Figura 1**, con los dos reportes que se presentan en el proyecto, el primero es acerca de las predicciones de las ventas por clase de vehículo y el segundo es el Top 10 de los modelos de vehículos más vendidos, todo esto para los años desde el 2019 hasta el 2024, asi como las predicciones del 2025 y 2026.

Figura 1

Portada principal.



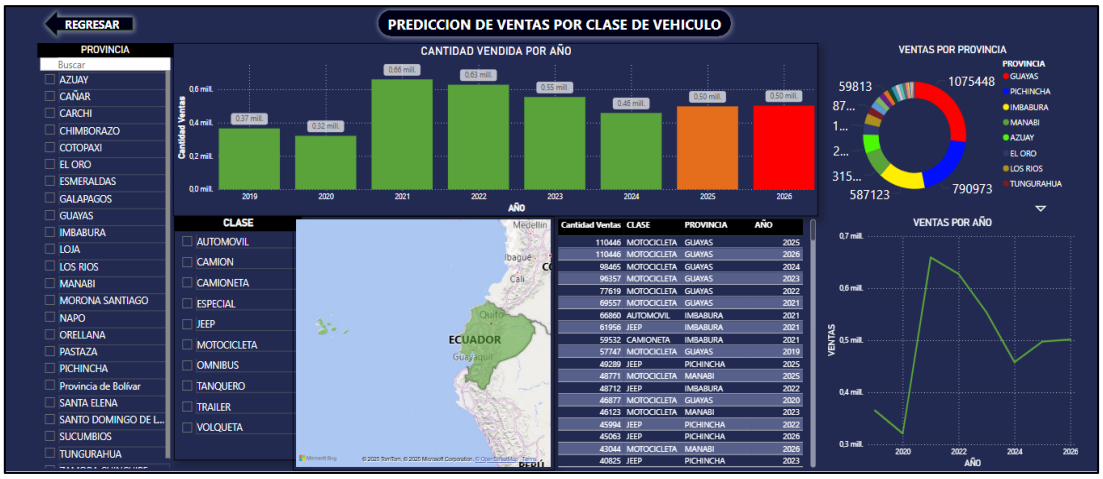
PREDICCION DE CANTIDAD DE VENTAS POR CLASE DE VEHICULO

En la **Figura 2** se muestra un análisis de la cantidad de vehículos vendidos en Ecuador a lo largo de los años, junto con su distribución geográfica.

Este análisis es clave para identificar patrones de compra a lo largo del tiempo y permite a los actores del sector automotriz tomar decisiones estratégicas basadas en datos históricos y proyecciones futuras.

Figura 2

Vehículos vendidos en el Ecuador, 2019 – 2026(predicción).

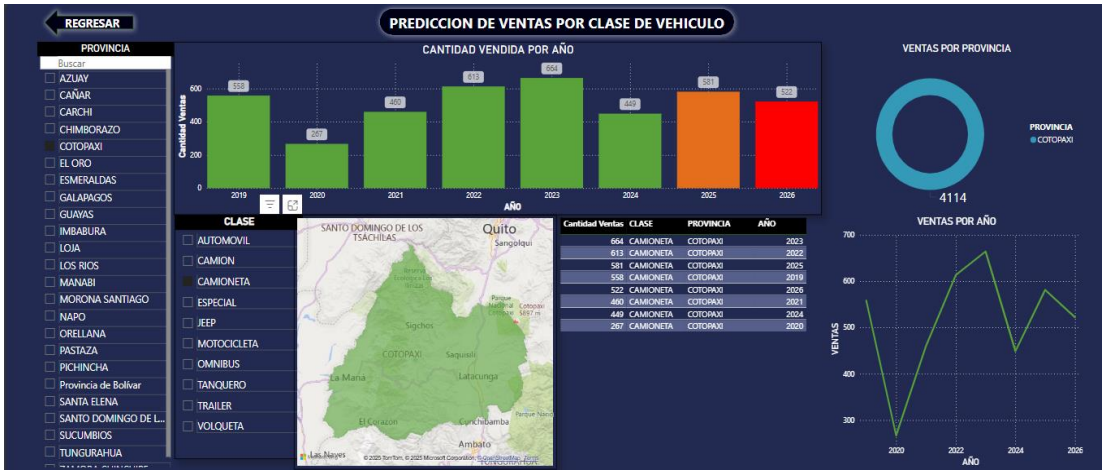


En la **Figura 3** se muestra análisis del gráfico con los filtros aplicados a la provincia de **Cotopaxi** y los tipos de vehículos **Camioneta**

- La provincia de Cotopaxi ha experimentado una recuperación del mercado automotriz después de la caída de 2020, alcanzando su mayor nivel de ventas en 2023.
- La proyección para 2026 muestra una tendencia estable, lo que indica que el mercado puede haber alcanzado un punto de equilibrio.
- Es importante analizar qué factores influyeron en la baja de ventas en 2024 y si existen oportunidades para impulsar el crecimiento en los próximos años.

Figura 3

Ventas de camionetas en Cotopaxi a lo largo de los años.



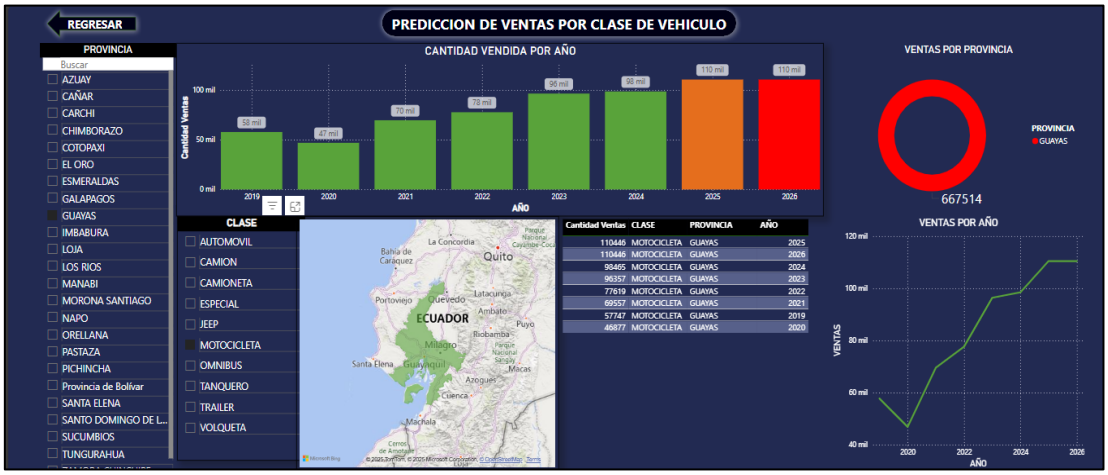
En la **Figura 4** muestra un análisis basado en los filtros aplicados al **año (2022 y 2026), la provincia de Guayas y la clase de vehículo (motocicleta)**

- La provincia de Guayas muestra una clara tendencia de crecimiento en la venta de motocicletas.

Este análisis es clave para concesionarias y distribuidores, ya que indica un mercado en expansión, lo que puede generar oportunidades para estrategias de ventas, financiamiento y servicios postventa.

Figura 4

Comparación entre el 2022 y la predicción 2026, Guayas, Motocicletas.



PREDICCION DEL TOP 10 DE LOS VEHICULOS MAS VENDIDOS

En la **Figura 5** muestra un análisis del top 10 de vehículos según el modelo, basado en los datos filtrados para la **provincia de Cañar en el año 2023**, considerando únicamente la clase de vehículo **automóvil.**

- El mercado en Cañar en 2023 muestra una clara preferencia por modelos compactos y eficientes (JOY BLACK, SOLUTO, ONIX, YARIS, PICANTO).
- Las concesionarias pueden centrar sus estrategias de venta en estos modelos, ofreciendo promociones, financiamiento o incentivos dirigidos a estos segmentos.

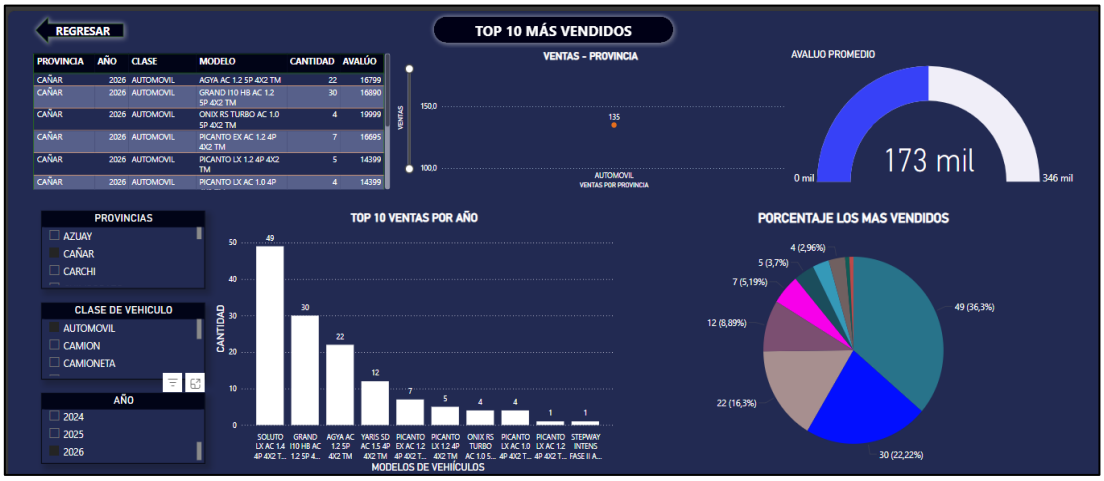
Figura 5

Top 10 automóviles en Cañar en el 2023

Mientas que para el año 2026, **Figura 6**, se prevé que el modelo del Kia Soluto va a liderar las ventas en dicha provincia, esto beneficiaria a las importadoras tanto de vehículos, repuestos y accesorios.

Figura 6

Predicción Top 10 automóviles en Cañar en el 2026



En la **Figura 7** y **Figura 8** se muestra un análisis basado en los datos filtrados para la **provincia de Chimborazo en el año 2022 y 2026 respectivamente, haciendo énfasis en los automóviles**:

Figura 7

Top 10 automóviles en Chimborazo en el 2022

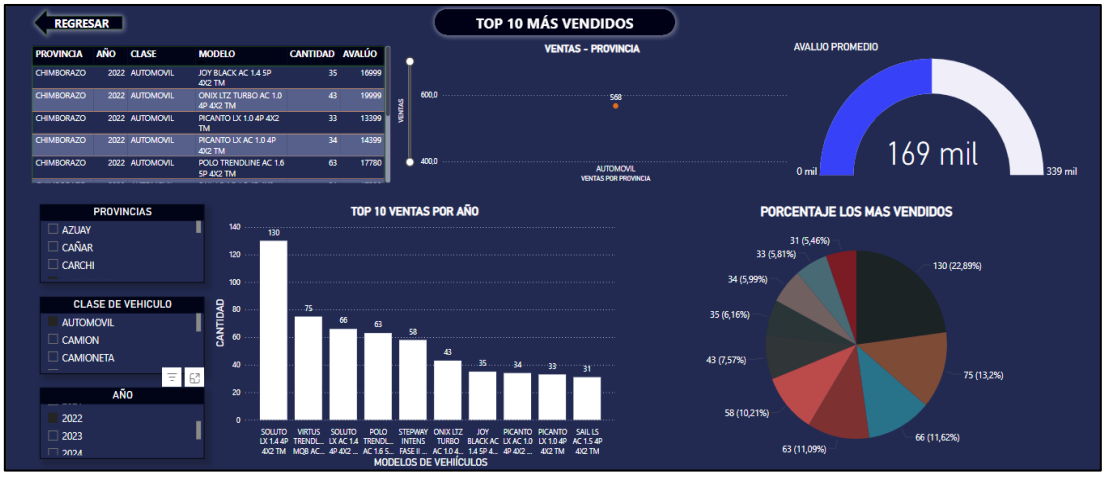
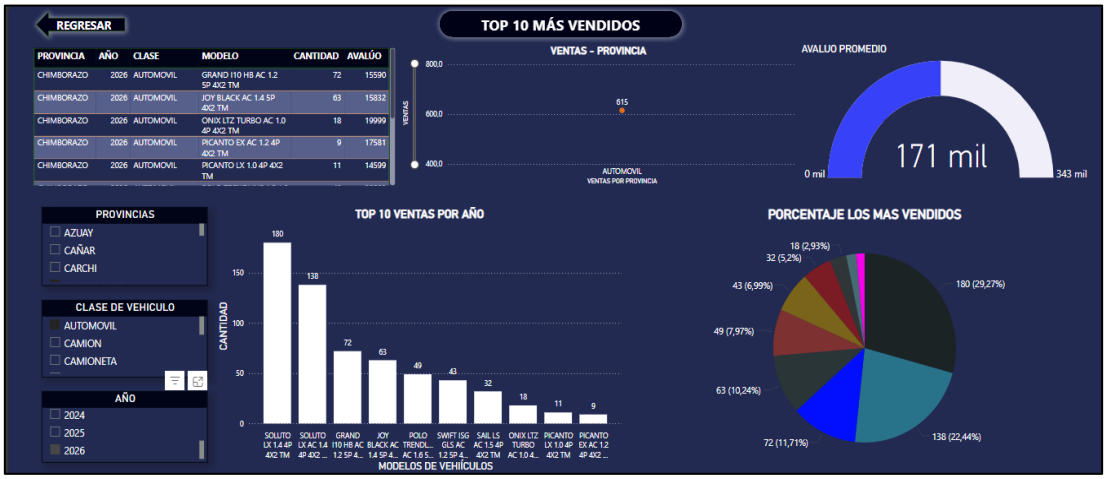


Figura 8

Predicción del Top 10 automóviles en Chimborazo en el 2026



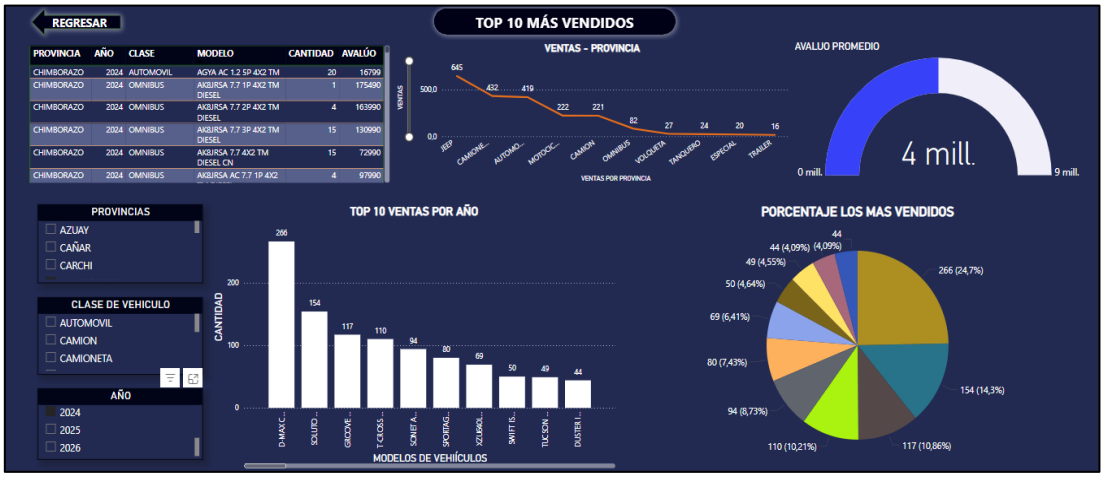
- En 2022, se vendieron 568 unidades, mientras que, en 2026, la cifra aumentó a 615 unidades.
- Esto representa un incremento del 8.3%, indicando un crecimiento en la demanda de automóviles en la provincia.
- En ambos años, el Soluto LX 1.4 4P 4X2 TM fue el modelo más vendido.
- Sin embargo, en 2026, su popularidad creció considerablemente, pasando de 66 unidades en 2022 a 180 unidades en 2026, consolidándose como el modelo más preferido por los compradores.

En la **Figura 9** muestra un análisis basado en los datos filtrados para la **provincia de Chimborazo en el año 2024**, considerando todas las clases de vehículos:

- La D-MAX domina el mercado en Chimborazo, confirmando la preferencia por camionetas de trabajo.
- Los SUVs y automóviles compactos también tienen un espacio relevante en la demanda, especialmente modelos eficientes y accesibles.

Figura 9

Top 10 todos los vehículos de Chimborazo 2024



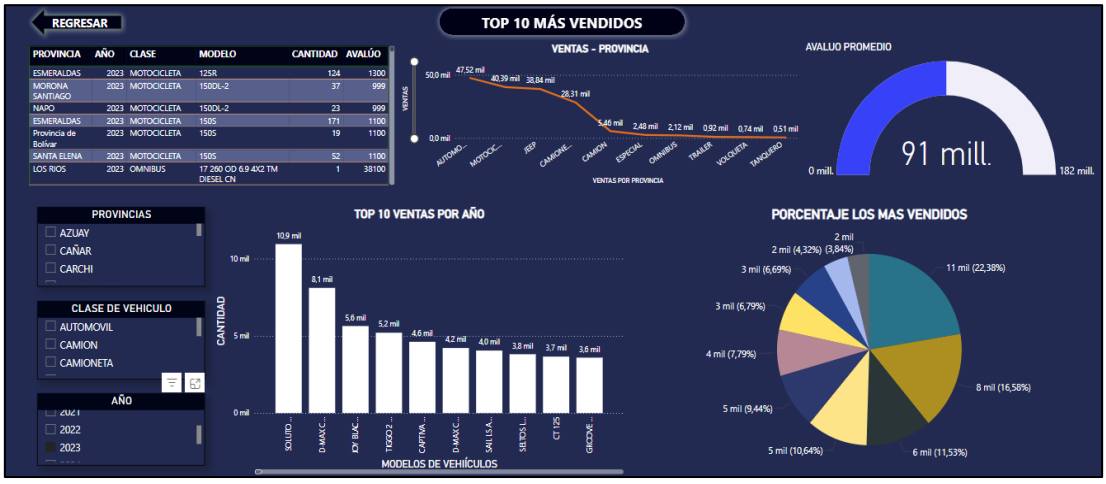
En la **Figura 10** muestra un análisis basado en la información de **los vehículos más vendidos a nivel nacional en el año 2023:**

- Los sedanes económicos dominan el mercado, lo que sugiere que las marcas deben seguir ofreciendo opciones accesibles y eficientes.
- Las camionetas continúan siendo clave en el sector comercial, por lo que las concesionarias pueden dirigir estrategias de financiamiento y promoción a empresas y negocios.

Los SUVs están en aumento, por lo que fabricantes y distribuidores pueden enfocarse en características como seguridad, espacio y tecnología para atraer a más compradores.

Figura 10

Revisión General Top 10 a nivel Nacional en el 2023, seleccionando el primero en la lista para el avalúo.

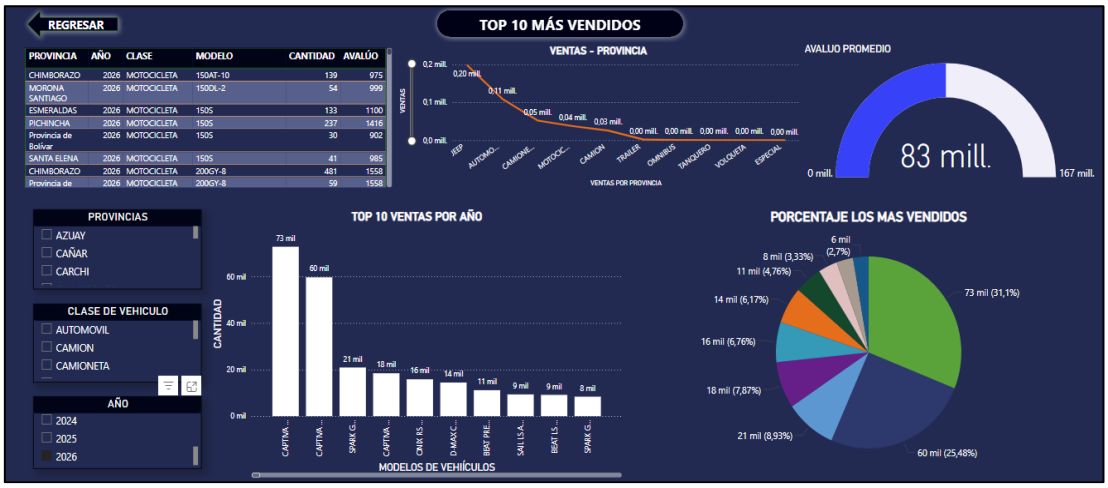


En la **Figura 11** se muestra un análisis basado en los datos filtrados para los **vehículos más vendidos a nivel nacional en el año 2026:**

- El **mercado ecuatoriano en 2026 estará dominado por SUVs y camionetas**, lo que sugiere que las concesionarias deben enfocarse en estos segmentos.
- Los autos compactos siguen presentes, pero con una participación menor, por lo que las marcas deberan optimizar su oferta con modelos más eficientes y tecnológicos.
- Las concesionarias pueden potenciar su estrategia de ventas ofreciendo financiamiento atractivo para SUVs y camionetas, así como incentivos para vehículos eléctricos o híbridos.

Figura 11

Predicción de la Revisión General Top 10 a nivel Nacional en el 2026, seleccionando el primero en la lista para el avalúo.



ANEXO 3

VALIDACION DE LOS ESPECIALISTAS

Universidad Israel

ESPOG | Escuela de Posgrados

UNIVERSIDAD TECNOLÓGICA ISRAEL

ESCUELA DE POSGRADOS "ESPOG"

MAESTRÍA EN BIG DATA Y CIENCIA DE DATOS

INSTRUMENTO PARA VALIDACIÓN DE LA PROPUESTA

Estimado colega:

Se solicita su valiosa cooperación para evaluar la calidad del siguiente contenido digital **"Análisis predictivo de ventas utilizando el método CRISP DM para el sector automotriz en el Ecuador"**. Sus criterios son de suma importancia para la realización de este trabajo, por lo que se le pide que brinde su cooperación contestando las preguntas que se realizan a continuación.

Datos informativos

Validado por: Sebastian Pozo Eugenio
Titulo obtenido: INGENIERO EN SISTEMAS
C.I.: 1804365276
E-mail: sebastianpe93@hotmail.com
Institución de Trabajo: Importadora Alvarado
Cargo: Jefe de Programación
Años de experiencia en el área: 9

Universidad Israel

ESPOG | Escuela de Posgrados

Instructivo:

- Responda cada criterio con la máxima sinceridad del caso.
- Revisar, observar y analizar la propuesta de la plataforma virtual, blog o sitio web.
- Coloque una X en cada indicador, tomando en cuenta que Muy adecuado equivale a 5, Bastante Adecuado equivale a 4, Adecuado equivale a 3, Poco Adecuado equivale a 2 e Inadecuado equivale a 1.

Tema: "Análisis predictivo de ventas utilizando el método CRISP DM para el sector automotriz en el Ecuador"

Indicadores	Muy adecuado	Bastante Adecuado	Adecuado	Poco adecuado	Inadecuado
Pertinencia	X				
Aplicabilidad	X				
Factibilidad		X			
Novedad			X		
Fundamentación pedagógica		X			
Fundamentación tecnológica		X			
Indicaciones para su uso	X				
TOTAL	30				

Observaciones: Resulta de mucha utilidad predecir las ventas de vehículos para todos los giros de negocio dentro de la industria automotriz; especialmente en la venta de repuestos, nos permite tener un panorama claro de las autopartes a importar.

Recomendaciones: Predecir las ventas dentro de los próximos 6 meses, es de utilidad sobre todo para planificar la importación de repuestos a mediano plazo.

Lugar, fecha de validación: Ambato, 05 marzo de 2025

Firma del especialista
Nombre: SEBASTIAN POZO EUGENIO

Página **2** de **2**

UNIVERSIDAD TECNOLÓGICA ISRAEL

ESCUELA DE POSGRADOS "ESPOG"

MAESTRÍA EN BIG DATA Y CIENCIA DE DATOS

INSTRUMENTO PARA VALIDACIÓN DE LA PROPUESTA

Estimado colega:

Se solicita su valiosa cooperación para evaluar la calidad del siguiente contenido digital **"Análisis predictivo de ventas utilizando el método CRISP DM para el sector automotriz en el Ecuador"**. Sus criterios son de suma importancia para la realización de este trabajo, por lo que se le pide que brinde su cooperación contestando las preguntas que se realizan a continuación.

Datos informativos

Validado por: PAUL DANIEL CHERRES CARGUACUNDO
Título obtenido: INGENIERIA EN ELECTRONICA TELECOMUNICACIONES Y REDES
C.I.: 0202134367
E-mail: pcherres@rectima.com.ec
Institución de Trabajo: RECTIMA INDUSTRY CIA. LTDA.
Cargo: JEFE DEL DEPARTAMENTO DE SISTEMAS
Años de experiencia en el área: 5 AÑOS

Universidad Israel

ESPOG | Escuela de Posgrados

Instructivo:

- Responda cada criterio con la máxima sinceridad del caso.
- Revisar, observar y analizar la propuesta de la plataforma virtual, blog o sitio web.
- Coloque una X en cada indicador, tomando en cuenta que Muy adecuado equivale a 5, Bastante Adecuado equivale a 4, Adecuado equivale a 3, Poco Adecuado equivale a 2 e Inadecuado equivale a 1.

Tema: "Análisis predictivo de ventas utilizando el método CRISP DM para el sector automotriz en el Ecuador"

Indicadores	Muy adecuado	Bastante Adecuado	Adecuado	Poco adecuado	Inadecuado
Pertinencia	X				
Aplicabilidad		X			
Factibilidad		X			
Novedad	X				
Fundamentación pedagógica		X			
Fundamentación tecnológica	X				
Indicaciones para su uso	X				
TOTAL	20	12			

Observaciones: La fundamentación en datos históricos y proyecciones de ventas es un buen punto de partida para el análisis de los datos, también se pudiera añadir factores externos para completas el estudio y así tener un resultado mejor en el análisis predictivo.

Recomendaciones: Se puede evidenciar con los datos proporcionados que el mercado automotriz ecuatoriano se ha mantenido al alza a partir del 2020 y con ciertas preferencias por algunos segmentos, lo cual puede ser aprovechado por las concesionarias para enfocar esfuerzos en estos segmentos y obtener mayores ganancias.

Lugar, fecha de validación: Ambato 5 de marzo 2025

Firma del especialista

Nombre: Paúl Cherres

UNIVERSIDAD TECNOLÓGICA ISRAEL

ESCUELA DE POSGRADOS "ESPOG"

MAESTRÍA EN BIG DATA Y CIENCIA DE DATOS

INSTRUMENTO PARA VALIDACIÓN DE LA PROPUESTA

Estimado colega:

Se solicita su valiosa cooperación para evaluar la calidad del siguiente contenido digital **"Análisis predictivo de ventas utilizando el método CRISP DM para el sector automotriz en el Ecuador"**. Sus criterios son de suma importancia para la realización de este trabajo, por lo que se le pide que brinde su cooperación contestando las preguntas que se realizan a continuación.

Datos informativos

Validado por: Jorge Enrique Carrasco U.
Título obtenido: Tecnólogo en sistemas
C.I.: 1803861440
E-mail: jorge86_car@hotmail.com
Institución de Trabajo: Importadora Alvarado Vásconez Ltda.
Cargo: Coordinador de Tecnologías
Años de experiencia en el área: 9 años

Universidad Israel

ESPOG | Escuela de Posgrados

Instructivo:

- Responda cada criterio con la máxima sinceridad del caso.
- Revisar, observar y analizar la propuesta de la plataforma virtual, blog o sitio web.
- Coloque una X en cada indicador, tomando en cuenta que Muy adecuado equivale a 5, Bastante Adecuado equivale a 4, Adecuado equivale a 3, Poco Adecuado equivale a 2 e Inadecuado equivale a 1.

Tema: "Análisis predictivo de ventas utilizando el método CRISP DM para el sector automotriz en el Ecuador"

Indicadores	Muy adecuado	Bastante Adecuado	Adecuado	Poco adecuado	Inadecuado
Pertinencia	X				
Aplicabilidad	X				
Factibilidad	X				
Novedad	X				
Fundamentación pedagógica		X			
Fundamentación tecnológica		X			
Indicaciones para su uso		X			
TOTAL	20	12			

Observaciones: El análisis se encuentra correctamente segmentado con los clasificadores adecuados para un análisis, el diseño del dashboard me parece que es el apropiado y usa colores que se permite distinguir los elementos usados, los datos presentados son claros y coherentes tomando en cuenta que se analizan también los años de pandemia.

Recomendaciones: Paralelamente a esta metodología sería interesante usar otra alternativa de predicción con el mismo origen de datos, para compararlos y verificar si el resultado tuviese alguna diferencia.

Lugar, fecha de validación: Ambato, 06 de marzo del 2025

JORGE ENRIQUE CARRASCO URQUIZO

Firma del especialista
Nombre: Jorge Enrique Carrasco Urquizo