

Abstract geometric lines in black on a white background, forming various overlapping polygons and triangles.

PROYECTO FINAL

# ANÁLISIS DE ROBOS DE AUTOS EN ARGENTINA (2022-2023)

Entrenamiento y optimización de modelos de ML

Alumna: Gimena Sozzi

Curso: Data Science | Comisión: 49140

Fecha de entrega: 10 Junio de 2024

# ÍNDICE

3| Introducción

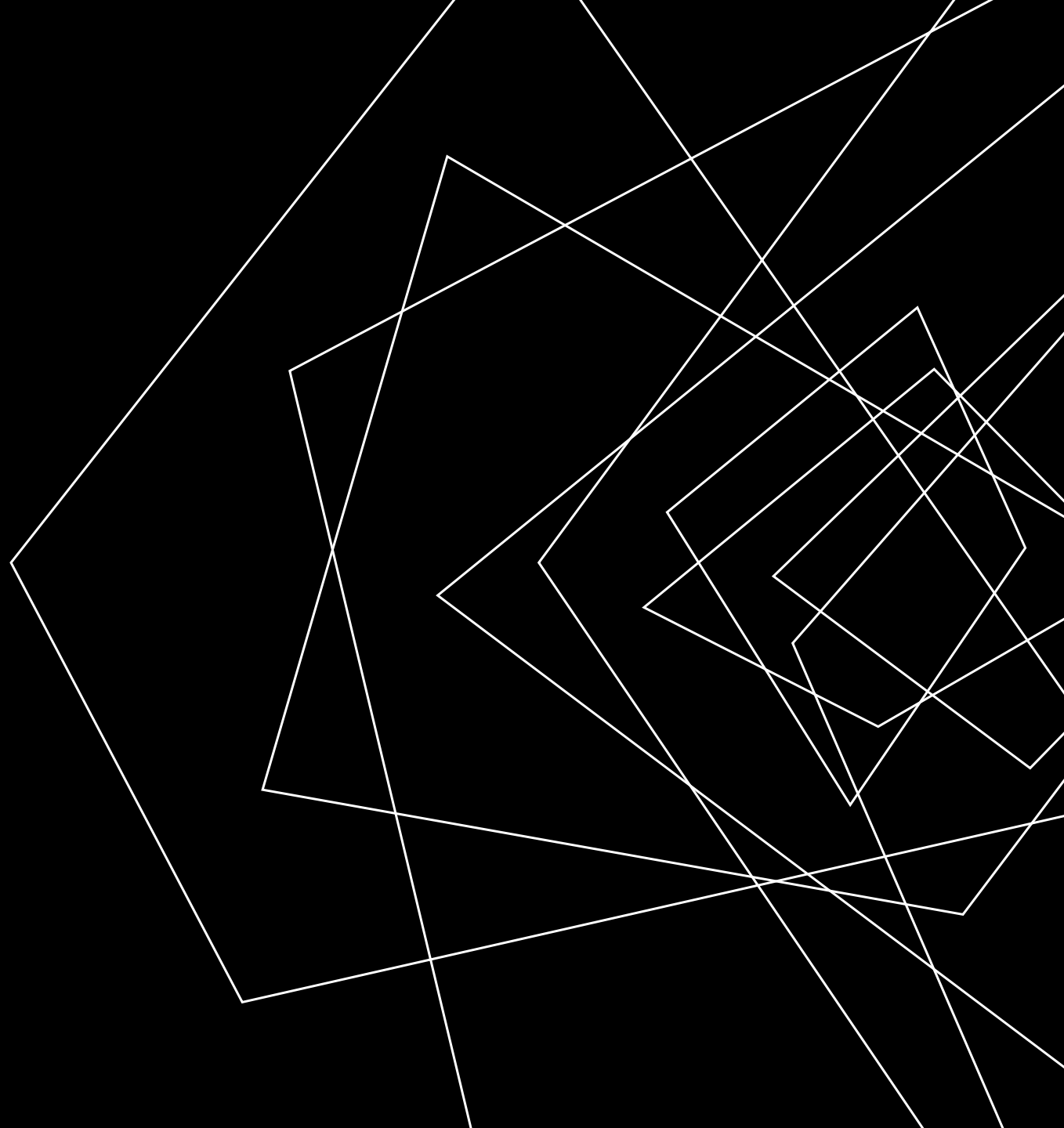
6| Exploracion de Datos

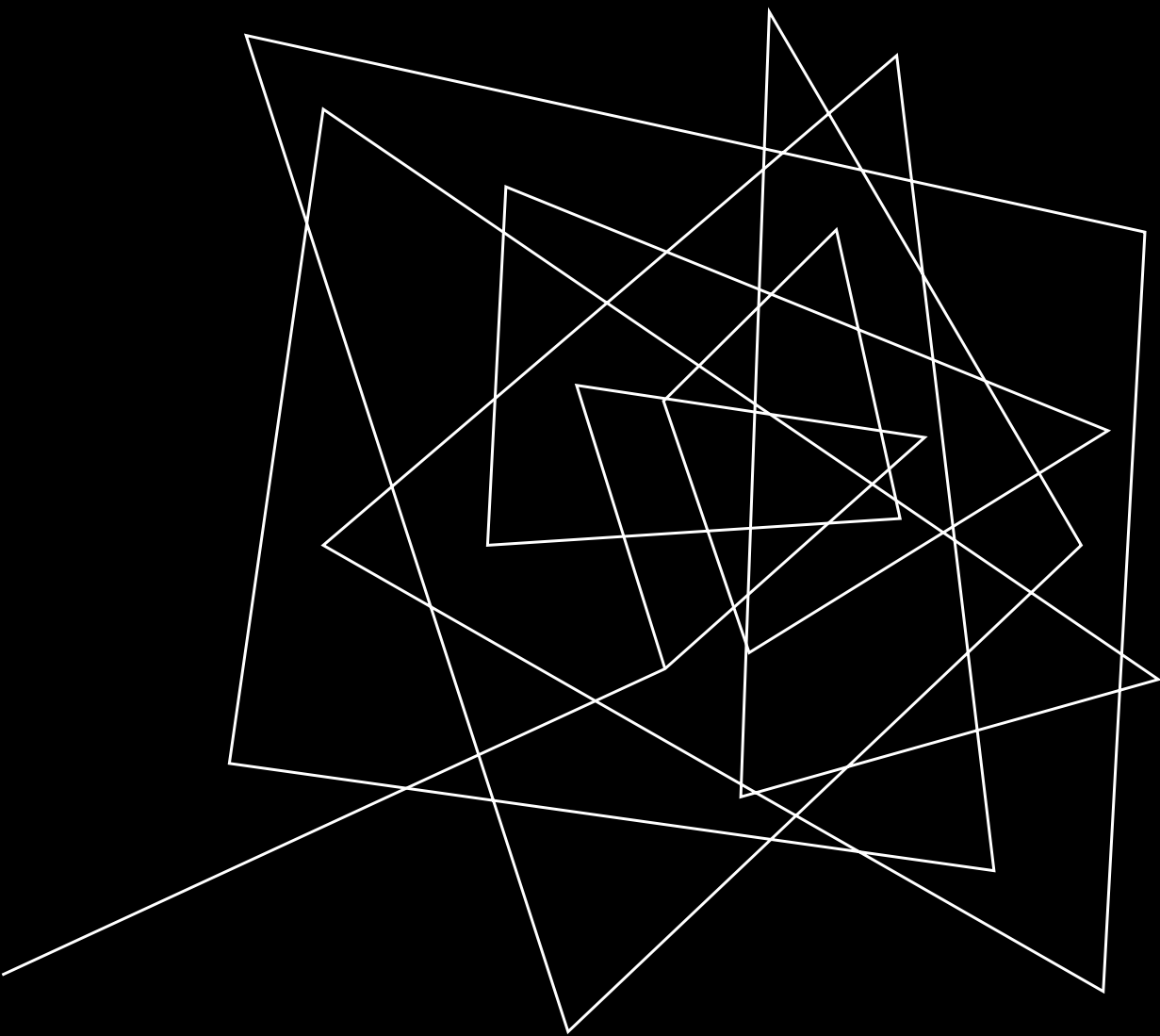
8| Análisis Exploratorio (EDA)

11| Metodología de Modelos

14| Conclusiones

18| Referencias





# INTRODUCCIÓN

# CONTEXTO Y AUDIENCIA

Este estudio se enfoca en el análisis de los robos de autos en Argentina durante los años 2022 y 2023.

La motivación es proporcionar insights clave para entidades como autoridades de seguridad, compañías de seguros y propietarios de vehículos, con el objetivo de desarrollar estrategias efectivas para prevenir y responder a estos incidentes.

La audiencia principal incluye analistas de datos, funcionarios de seguridad pública y profesionales del sector asegurador en Argentina.

# PREGUNTAS DE INTERÉS Y OBJETIVO DE INVESTIGACIÓN

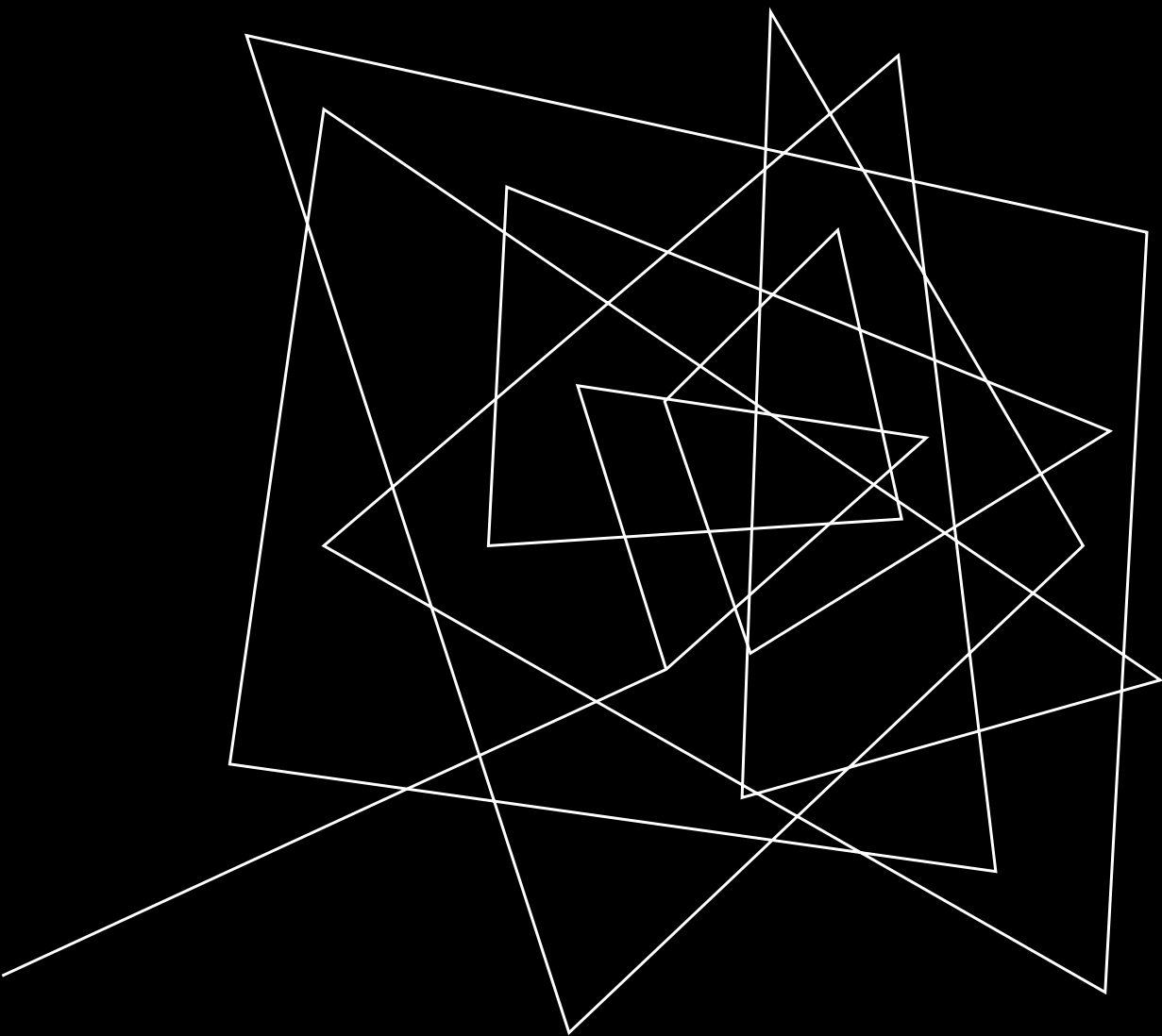
¿Qué factores contribuyen más significativamente a los robos de autos en Argentina?

¿Existen patrones geográficos en los robos de vehículos?

¿Cómo influyen las características del vehículo y del titular en la probabilidad de robo?

## **Objetivo Principal:**

El objetivo principal de este estudio es identificar patrones y tendencias en los robos de autos en Argentina. Utilizando técnicas avanzadas de análisis de datos y machine learning, buscamos comprender mejor este fenómeno y sus dinámicas subyacentes, para contribuir significativamente a la creación y optimización de medidas preventivas más efectivas.

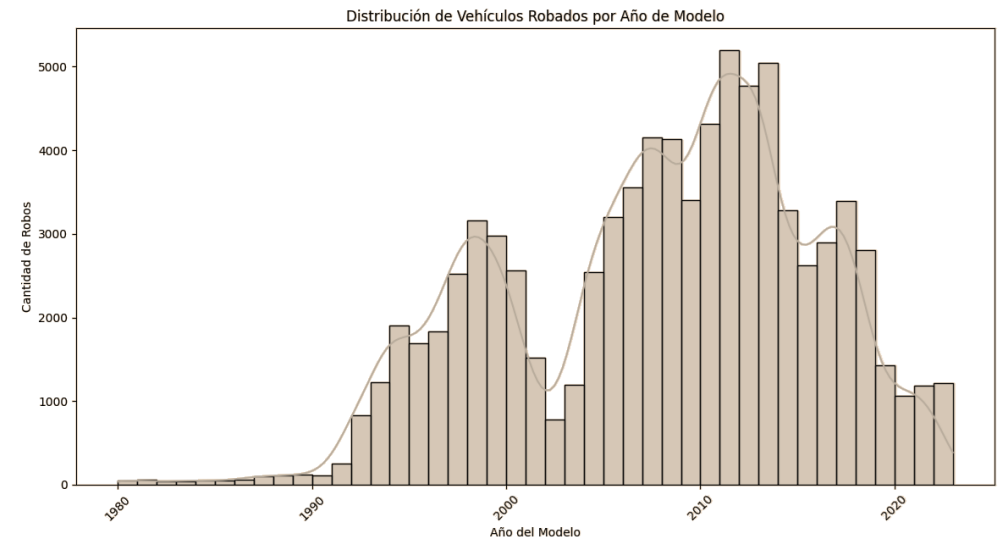


# EXPLORACIÓN DE DATOS

# RESUMEN METADATA

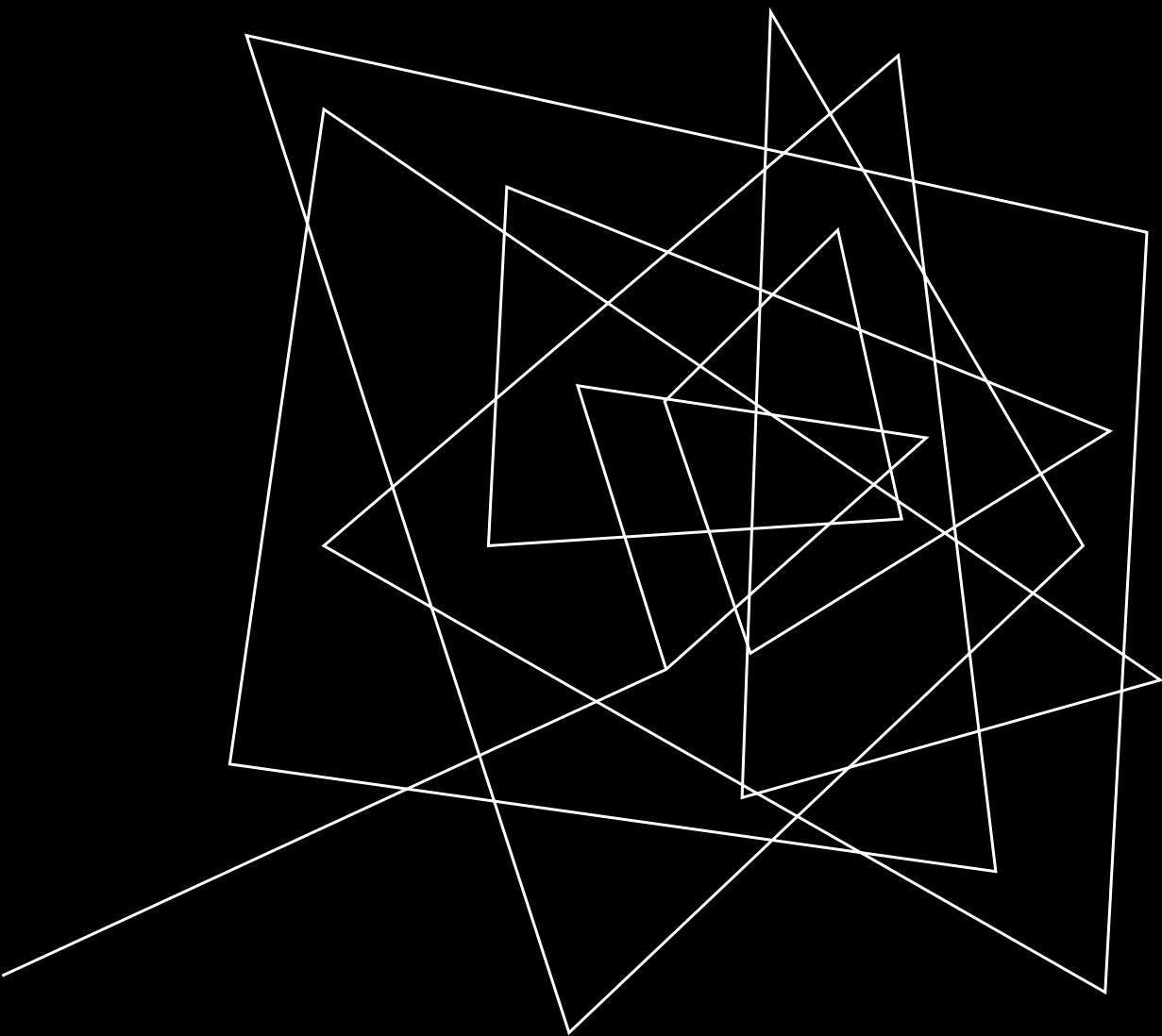
- **Cantidad de Registros Analizados:** 83.834
- **Rango Temporal:** Datos comprendidos entre enero de 2022 y diciembre de 2023.
- **Fuente:** Dirección Nacional de Registros Nacionales de la Propiedad Automotor y Créditos Prendarios del Ministerio de Justicia y Derechos Humanos de Argentina.
- **Características incluidas en el dataset:**
  - `tramite\_tipo`, `tramite\_fecha`: Información sobre el tipo y la fecha del trámite.
  - `fecha\_inscripcion\_inicial`: Fecha original de inscripción del vehículo.
  - `registro\_seccional\_codigo`, `registro\_seccional\_descripcion`, `registro\_seccional\_provincia`: Datos del registro del vehículo.
  - `automotor\_origen`, `automotor\_anio\_modelo`, `automotor\_marca\_codigo`, `automotor\_marca\_descripcion`, `automotor\_modelo\_codigo`, `automotor\_modelo\_descripcion`: Detalles específicos del automotor.
  - `automotor\_tipo\_codigo`, `automotor\_tipo\_descripcion`: Clasificación del tipo de vehículo.
  - `automotor\_uso\_codigo`, `automotor\_uso\_descripcion`: Uso del vehículo según registro.
  - `titular\_tipo\_persona`, `titular\_domicilio\_localidad`, `titular\_domicilio\_provincia`, `titular\_genero`, `titular\_anio\_nacimiento`: Información sobre el titular del vehículo.
  - `titular\_pais\_nacimiento`, `titular\_porcentaje\_titularidad`: Nacionalidad del titular y porcentaje de titularidad.
  - `titular\_domicilio\_provincia\_id`, `titular\_pais\_nacimiento\_id`: IDs codificados de la provincia y país de nacimiento del titular.

## DISTRIBUCIÓN DE LOS VEHICULOS ROBADOS POR AÑO DE MODELO (2022 – 2023)



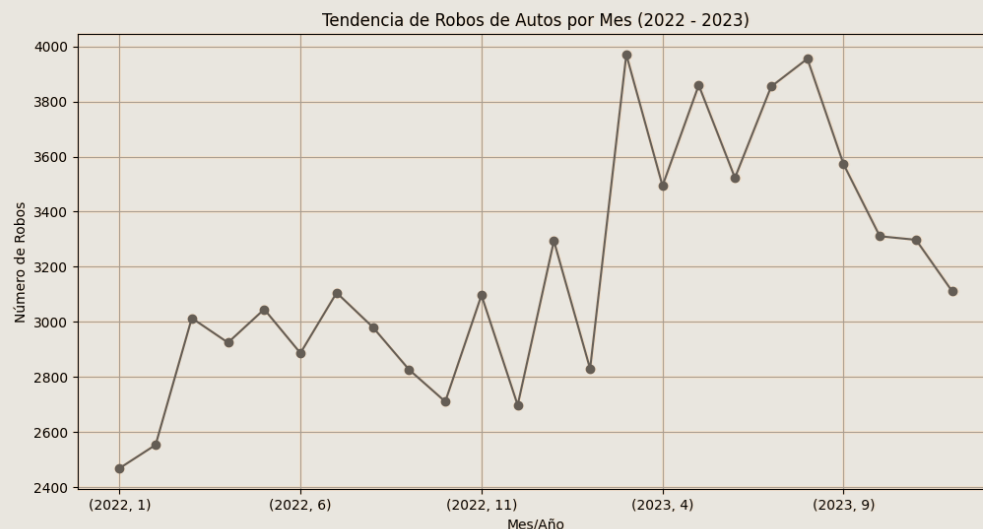
Este histograma ilustra la distribución de los vehículos robados en Argentina, resaltando los años de modelo más afectados por esta problemática entre 2022 y 2023.

Los datos muestran una concentración elevada de robos en vehículos modelos entre 2010 y 2015.



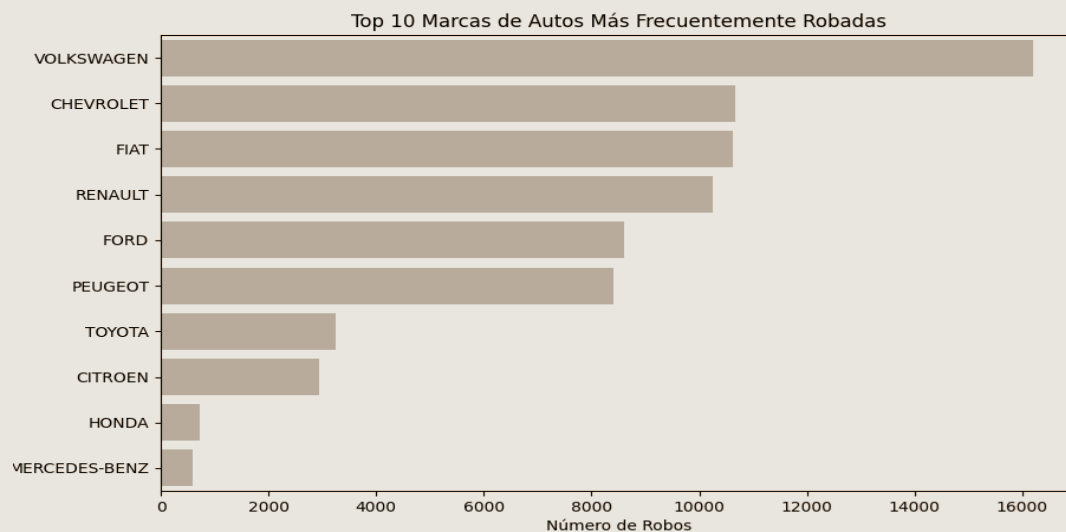
# ANÁLISIS EXPLORATORIO (EDA)





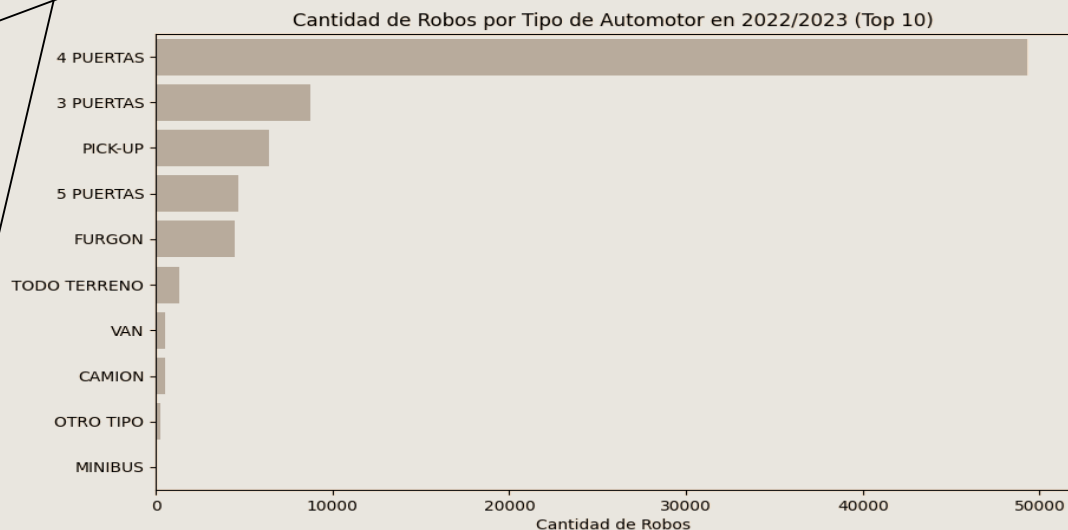
## TENDENCIA DE ROBOS DE AUTOS POR MES (2022-2023)

El gráfico en cuestión muestra la variabilidad mensual en los robos de autos en Argentina a lo largo de 2022 y 2023, evidenciando que ciertos periodos del año, como vacaciones y eventos festivos, podrían incrementar la incidencia de robos. Este patrón sugiere una correlación entre las fluctuaciones de los robos y los cambios estacionales en los hábitos sociales y económicos, afectando la seguridad vehicular.



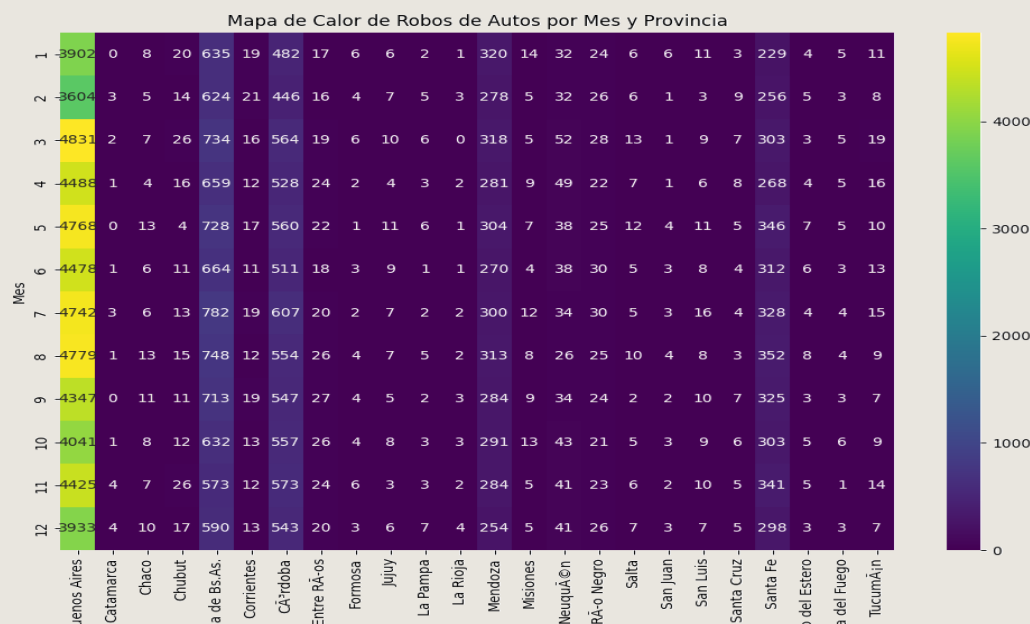
## RANKING DE MARCAS DE AUTOS MÁS ROBADAS (2022-2023)

Este gráfico destaca las marcas de autos más frecuentemente robadas, revelando patrones específicos que reflejan las preferencias de los ladrones, influenciadas por la popularidad de las marcas, el valor de reventa, la demanda de repuestos, y las vulnerabilidades de seguridad. Las diferencias en la incidencia de robos entre las marcas sugieren que factores como la disponibilidad y visibilidad de los vehículos pueden aumentar su riesgo de ser objetivos, mientras que el atractivo económico de ciertos autos, debido a su valor de reventa y la demanda de sus partes, los hace especialmente vulnerables a los robos.



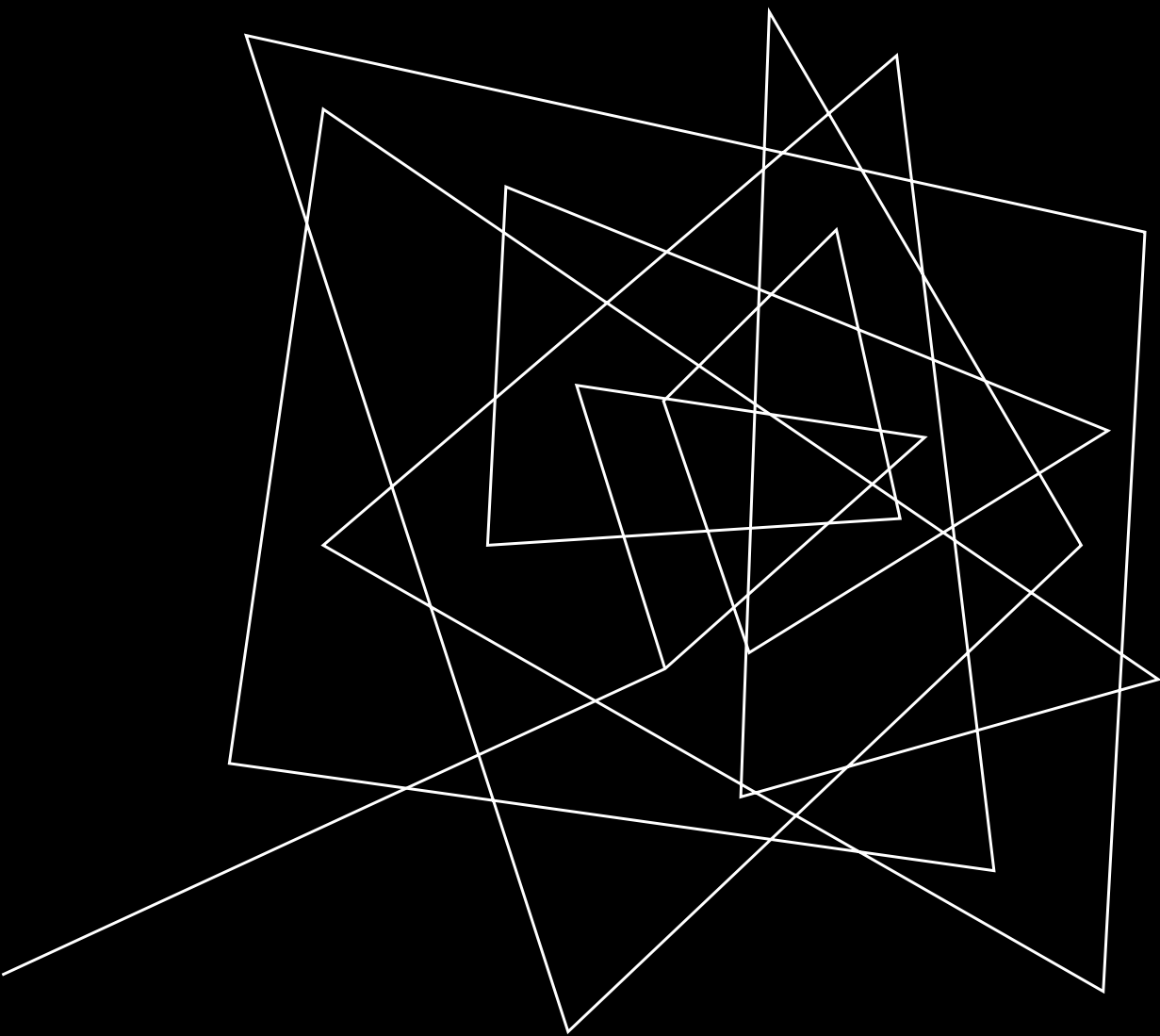
## RANKING DE TIPO DE AUTOMOTORES MÁS ROBADOS (2022-2023)

El gráfico de barras refleja que determinados tipos de automotores en Argentina son más propensos al robo, con preferencias marcadas por los delincuentes hacia vehículos que presentan mayor facilidad de robo, alta demanda en el mercado negro, o valor de reventa elevado. Esta tendencia no solo revela las vulnerabilidades de ciertos vehículos, posiblemente debido a deficiencias en seguridad o su prevalencia en las calles, sino también las inclinaciones de los ladrones hacia automotores que pueden ser rápidamente convertidos en ganancia.



## DISTRIBUCIÓN DE ROBOS DE AUTOS POR MES Y PROVINCIA (2022-2023)

El mapa de calor demuestra que los robos de autos en Argentina tienen una distribución temporal y regional definida, con Buenos Aires registrando los números más altos, seguido por CABA, y luego por Córdoba, Mendoza y Santa Fe, en menor medida. Los picos de robos reflejan aspectos como la densidad poblacional, la concentración de actividades y la circulación vehicular. Las variaciones en la frecuencia de robos sugieren una influencia de factores urbanos, como zonas de menor vigilancia y la facilidad de movilidad para los delincuentes, especialmente en áreas con alta actividad económica y social.



# METODOLOGÍA DE MODELOS

# ENTRENAMIENTO DE MODELOS

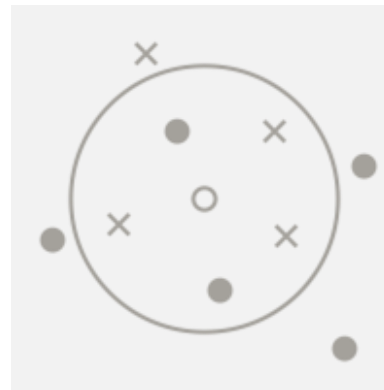
Para abordar el problema de la predicción de robos de vehículos, se utilizaron varios algoritmos de machine learning. Cada uno de estos modelos fue seleccionado por sus capacidades únicas para manejar datos de alta dimensionalidad y su eficacia en la clasificación.

A continuación, se describen los modelos entrenados y las técnicas de optimización empleadas para mejorar su rendimiento.



**RANDOM FOREST**

Seleccionado por su robustez y capacidad para manejar grandes cantidades de datos con múltiples características. Es menos propenso al sobreajuste y puede evaluar la importancia de cada característica en la predicción.



**KNN (K-Nearest  
Neighbors)**

Elegido por su simplicidad y eficacia en clasificaciones basadas en proximidad. Es útil cuando la relación entre las características no es lineal y puede detectar patrones locales en los datos.




**SVM (Support Vector  
Machines)**


Utilizado por su eficiencia en espacios de alta dimensión y su capacidad para maximizar el margen de separación entre clases. Es ideal para nuestro dataset, que incluye múltiples características transformadas y sintéticas.


# RESULTADOS OPTIMIZACIÓN Y SELECCIÓN DE MODELO

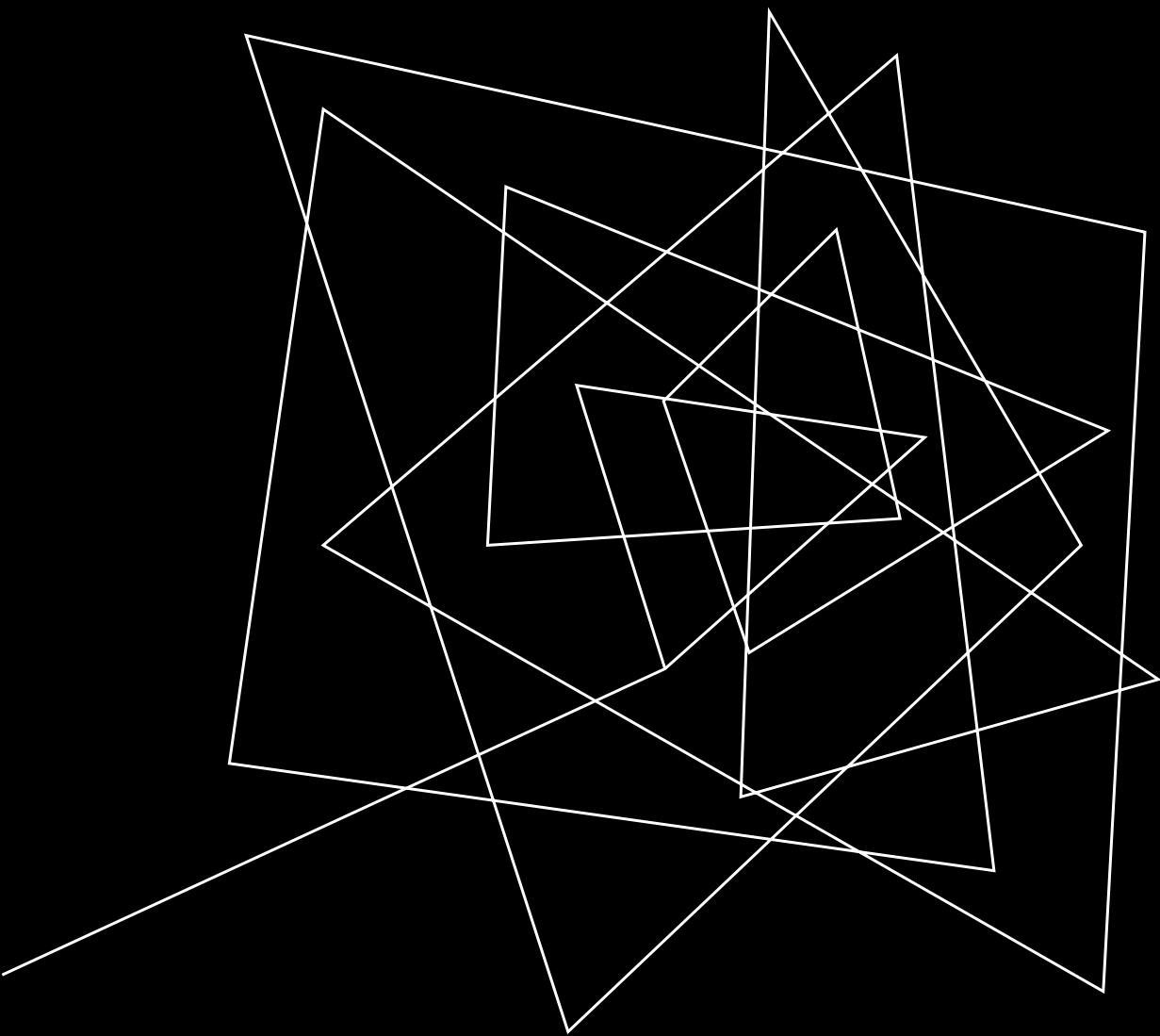
**Random Forest optimizado con HalvingRandomSearchCV** es el modelo que elegimos como el más adecuado para predecir robos de vehículos debido a los siguientes motivos:

- Alta Precisión:**  
Alcanzó una precisión de 0.8506, garantizando predicciones confiables.
- Robustez:**  
Su capacidad para manejar grandes datasets y evitar el sobreajuste asegura predicciones fiables tanto en datos de entrenamiento como en nuevos datos.
- Importancia de Características:**  
Proporciona una evaluación clara de las características más influyentes, ayudando a identificar factores clave en los robos de vehículos.
- Eficiencia en Tiempo de Ejecución:**  
Con HalvingRandomSearchCV, se optimizó en 152.52 segundos, balanceando precisión y eficiencia.
- Equilibrio entre Rendimiento y Complejidad:**  
Comparado con otros modelos como KNN y SVM, Random Forest ofrece una alta precisión sin una complejidad excesiva.

MODELO	TÉCNICA	PRECISIÓN	TIEMPO DE EJECUCIÓN (seg)
	RANDOM FOREST		
	GridSearchCV	0.8506	427.83
	RandomizedSearchCV	0.8506	419.31
	HalvingGridSearchCV	0.8497	164.87
HalvingRandomSearchCV		0.8506	152.52

MODELO	TÉCNICA	PRECISIÓN	TIEMPO DE EJECUCIÓN (seg)
	KNN		
	GridSearchCV	0.7565	10.53
	RandomizedSearchCV	0.7565	10.29
	HalvingGridSearchCV	0.7565	15.94
HalvingRandomSearchCV		0.7565	16.17

MODELO	TÉCNICA	PRECISIÓN	TIEMPO DE EJECUCIÓN (seg)
	SVM		
	GridSearchCV	0.6603	3047.64
	RandomizedSearchCV	0.6603	3003.92
	HalvingGridSearchCV	0.6520	755.82
HalvingRandomSearchCV		0.6499	456.18



CONCLUSIONES

# CONCLUSIONES

## 1. ANÁLISIS EXPLORATORIO DE DATOS (EDA)

Se realizó el EDA sobre el dataset, obteniendo una descripción general de las variables que lo componen, identificando patrones clave y correlaciones.

## 2. ESTUDIO DE FACTORES DE RIESGO

Con el dataset, se pudo estudiar los factores que influyen en los robos de vehículos, considerando tanto características del automotor como demográficas del titular.

## 3. VARIABLE OBJETIVO

La variable objetivo, tipo de automotor, se analizó para entender cómo diferentes características influyen en su clasificación

# CONCLUSIONES

## 4. MODELO SELECCIONADO – RANDOM FOREST

El modelo de Random Forest, optimizado con HalvingRandomSearchCV, fue seleccionado por su robustez y alta precisión. Este modelo permite predecir el tipo de vehículo robado y entender los factores más influyentes en los robos.

## 5. PRECISIÓN Y EFICIENCIA DEL MODELO

El modelo de Random Forest alcanzó una precisión de 85,06% y un tiempo de ejecución de 152.52 segundos, demostrando un equilibrio óptimo entre precisión y eficiencia.

## 6. IMPORTANCIA DE LAS CARACTERÍSTICAS

El modelo proporciona una evaluación clara de la importancia de las características, permitiendo identificar los factores más críticos asociados con los robos de vehículos, como la marca, modelo y ubicación geográfica.





# CONCLUSIONES

## 7. IMPACTO Y APLICACIONES

Los insights obtenidos del modelo pueden ayudar a las autoridades y compañías de seguros a diseñar estrategias de prevención más efectivas y personalizadas, enfocándose en los factores de mayor riesgo identificados.

## 8. RECOMENDACIONES

Basado en los hallazgos, se recomienda implementar medidas específicas en las regiones y para los tipos de vehículos más vulnerables, además de fomentar la colaboración entre autoridades y aseguradoras para compartir datos y estrategias preventivas.



# REFERENCIAS

- Datos Argentina / Dataset / Robos y recuperos de autos. DNRPA. Robos y recuperos de autos - 2022/2023. Ministerio de Justicia y Derechos Humanos. Subsecretaría de Asuntos Registrales. Dirección Nacional de Registros Nacionales de la Propiedad Automotor y Créditos Prendarios. Trámites de robos y recuperos de automotores inscriptos en los Registros Seccionales de la Propiedad del Automotor y Créditos Prendarios.

[Datos Argentina - Robos y recuperos de autos](#)