



Informes de Mercado para el Sector Automotriz mediante Ciencia de Datos

Proyecto desarrollado por Erick López, Christian Nüesch y Débora Zurita.
Dirigido por los docentes Nahuel Pratta y Marcos Ugarte.

El objetivo es crear un sistema automatizado de informes de mercado para optimizar decisiones en el sector automotriz.



Resumen del Proyecto

1

Ciencia de Datos e IA

Uso de machine learning para análisis predictivo en el sector automotriz.

2

Automatización

Scripts en Python para generar informes automáticamente.

3

Análisis Predictivo

Identificación de tendencias y prevención de fraudes en la industria.

4

Visualización

Desarrollo de dashboards y gráficos interactivos para presentar datos.

Problemáticas

Competencias y Metas

Problemáticas

Falta de datos actualizados y análisis predictivo en el sector automotriz.

Necesidad de herramientas de decisión en tiempo real para concesionarios y aseguradoras.

Metas

Generar 100 informes automatizados en tres meses.

Reducir tiempo de decisiones estratégicas en un 50%.

Disminuir fraudes en un 20% en el primer año.

→ *Limpieza y preparación de datos*

Introducción y Objetivos

1

Análisis de Dataset

Características de automóviles obtenidas por web scraping en Argentina (enero 2023).

2

Objetivos Generales

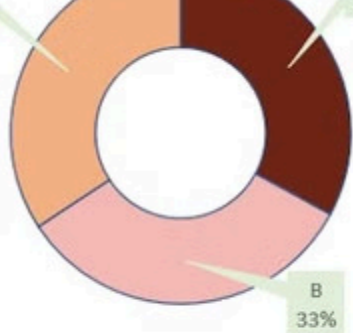
Desarrollar habilidades de organización y comunicación de hallazgos de datos.

3

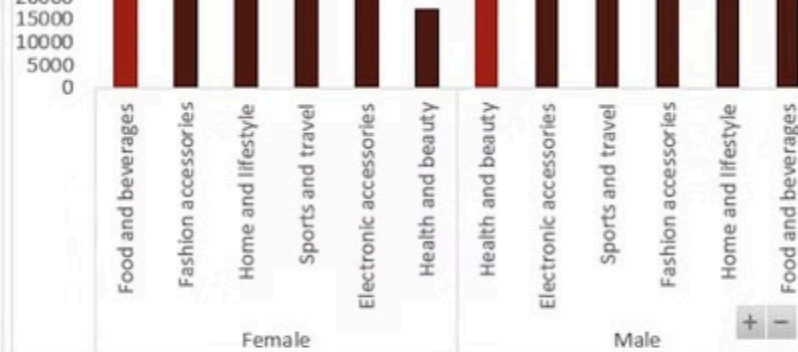
Objetivos Específicos

Limpieza de datos y análisis gráfico para relacionar variables relevantes.

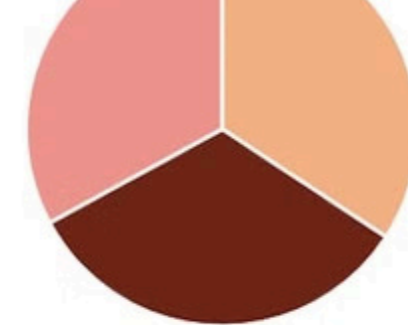




Sales Trend



Payment Method Per Branch



Gross Income Per Product Line

Desarrollo y Procesamiento de Datos

1

Origen de Datos

Dataset de Kaggle con características de autos como precio y marca.

2

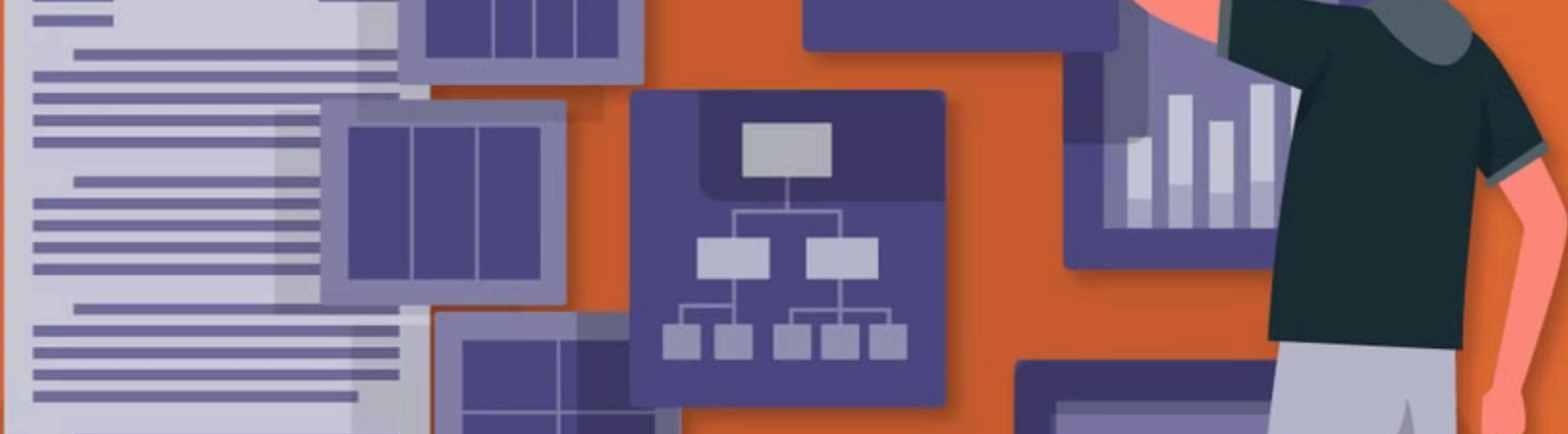
Limpieza

Tratamiento de valores ausentes y corrección de formatos incorrectos.

3

Análisis Estadístico

Extracción de estadísticas y visualización mediante gráficos diversos.



Conclusiones y Resultados Finales

Habilidades Desarrolladas

Limpieza de datos, detección de outliers y análisis exploratorio.

Insights Obtenidos

Identificación de tendencias clave en el mercado de autos en Argentina.

Resultados

Informe completo y códigos reproducibles en Python para análisis de datos.



→ *Regresión Lineal Simple*

Objetivo y Preparación de Datos

1

Objetivo

Explorar relaciones entre precio, kilometraje, año y potencia de vehículos en Argentina.

2

Carga de Datos

Importación de dataset desde GitHub y conversión de precios a dólares.

3

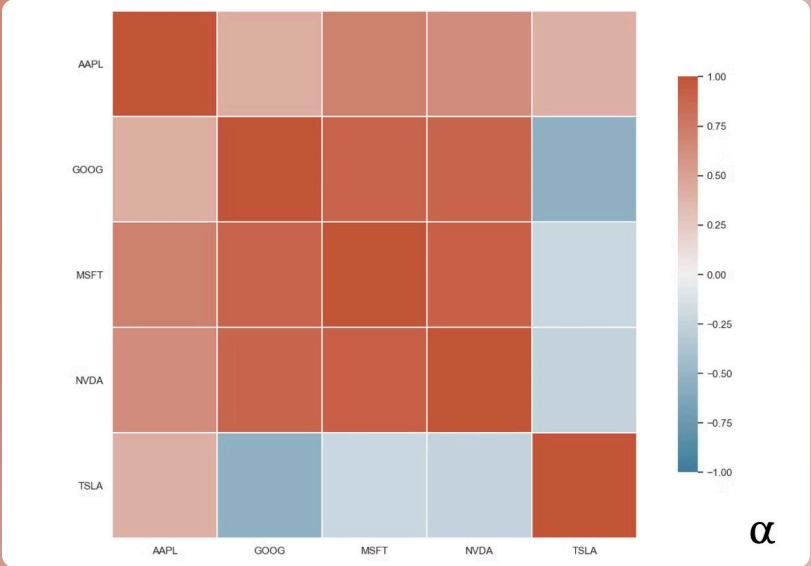
Herramientas

Uso de Python con librerías como Pandas, NumPy y Scikit-learn.

Análisis de Correlación y Modelo

Variable	Correlación con Precio
Kilometraje	-0.27
Año de Fabricación	0.24
Motor	0.56

Modelo de regresión lineal: $R^2 = 0.468$, RMSE = 33,964 km.



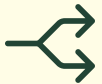
→ Regresión Lineal Múltiple

Regresión Lineal Múltiple y Análisis de Correlación



Preprocesamiento

Transformación de columnas categóricas y numéricas.



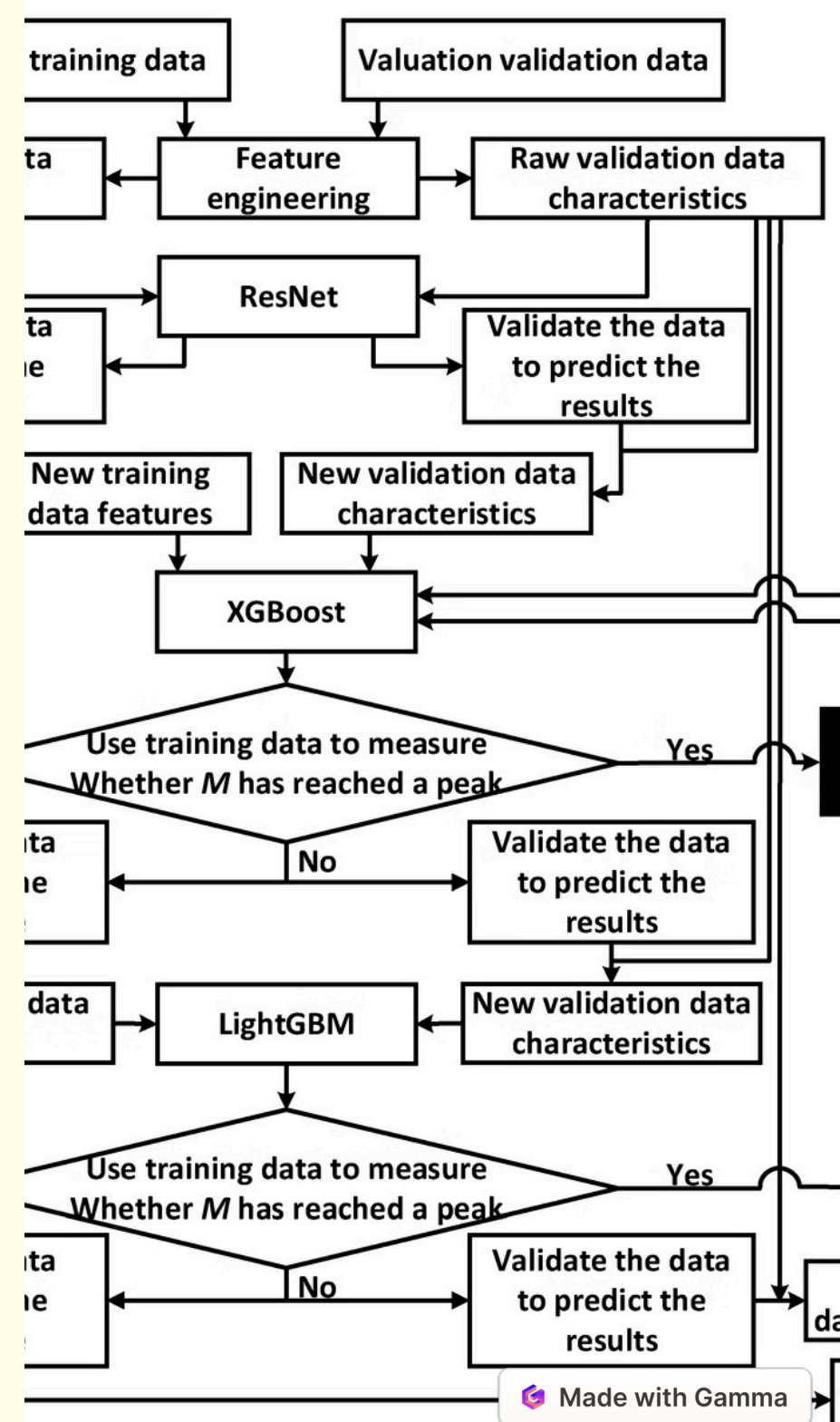
División del Dataset

80% entrenamiento, 20% prueba para evaluar el modelo.

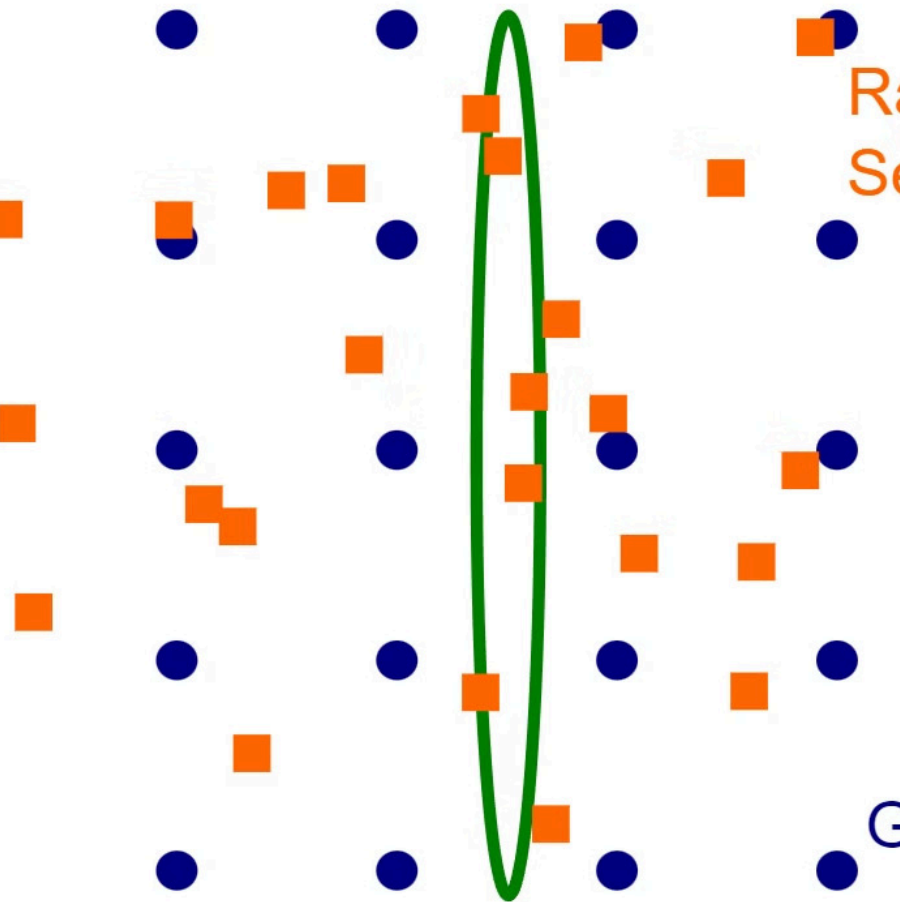


Análisis de Correlación

Evaluación de relaciones entre variables mediante matriz y heatmap.



Hyperparameter 1 (important hyperparameter)



Region of good
hyperparameters

Ajuste del Modelo y Resultados

1

Construcción del Modelo

Uso de Statsmodels para
regresión lineal múltiple y
escalado de variables.

2

Selección de Variables

Métodos forward y backward
para identificar las mejores
predictoras.

3

Resultados

Alto R-cuadrado ajustado y buen equilibrio entre precisión y
simplicidad.