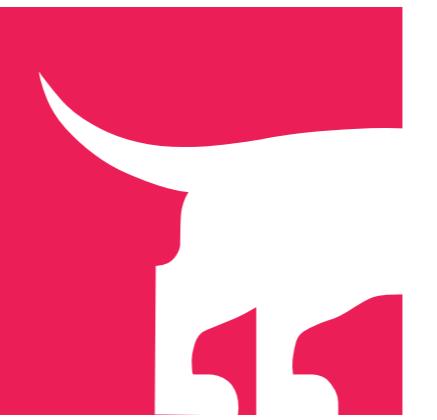


De la Calle al Hogar



Explorando Datos Estadísticos sobre la
Adopción de Animales Rescatados

Nombre del alumno_Gabriel Reyes

Número del comisión_60895

Nombre del Profesor_Germán Rodriguez

Nombre del Tutor_Ignacio Manuel Fernandez



transformar vidas animales

Temática



Cual es la situación de los perros rescatados?
Es sencillo encontrarles un hogar?

"**Del Refugio al Hogar**: Explorando Datos Estadísticos sobre la Adopción de Animales Rescatados" es un proyecto integral que busca arrojar luz sobre los patrones y factores que influyen en la transición exitosa de animales rescatados desde los refugios hacia nuevos hogares. A través de un enfoque centrado en datos extraídos del refugio Bloomington Animal Shelter, se analizarán estadísticas detalladas relacionadas con la adopción de mascotas, abordando aspectos como la edad, la raza y la duración de su estancia en el refugio.



Este proyecto no solo tiene como objetivo entender mejor los factores que afectan la adopción, sino también proporcionar información valiosa que podría utilizarse para mejorar estrategias de rescate y promoción de adopciones. Al explorar estas estadísticas, se busca fomentar la conexión entre animales rescatados y nuevos hogares, contribuyendo así al bienestar animal y promoviendo una cultura de adopción responsable. **"Del Refugio al Hogar"** aspira a ser una herramienta informada y comprensiva para aquellos comprometidos con la mejora de las vidas de los animales rescatados a través de la adopción.



que queremos comprobar?

Hipótesis



01. Estado de situación de rescate

El abandono es el principal estado de situación previo al rescate siendo los perros más jóvenes los mas perjudicados. Será el extravío o perdida la principal causa?

02. La edad de los perros afecta la velocidad de adopción

Los animales más jóvenes tienden a ser adoptados más rápidamente que los de mayor edad. También podemos analizar si existe algún patrón en relación al tiempo del animal en estado de adopción.

03. Influencia de la raza en la adopción

Algunas razas de perros son adoptadas con mayor facilidad que otras, debido a factores culturales, estereotipos o preferencias populares. ¿Existe el perjuicio de las razas “peligrosas”?



navegando por las estadísticas

Objetivo

El objetivo central del proyecto se propone sumergirse en los intrincados detalles que rodean la adopción de animales rescatados, analizando los factores que desempeñan un papel crucial en este proceso. Se buscará una comprensión profunda de cómo la edad, la raza y la duración de la estancia en el refugio afectan las probabilidades de adopción de un animal.

Edad: Investigar cómo la edad de los animales rescatados influye en su atractivo para los posibles adoptantes. Se explorará si existe una preferencia general por animales más jóvenes y en qué medida la edad afecta la rapidez con la que son adoptados.

Raza: Analizar el impacto de la raza en las tasas de adopción, investigando si ciertas razas son más o menos propensas a ser adoptadas. Esto puede incluir la percepción cultural de ciertas razas y cómo esto afecta las decisiones de adopción.



dataset elegido **Datos**

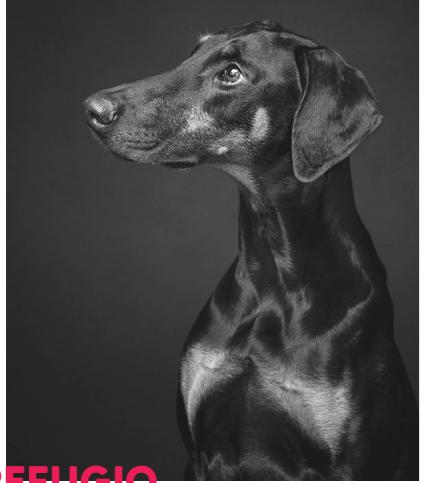
Para el desarrollo del proyectouento con un dataset extraído de la web de kaggle que consiste en la recopilación de datos del **Bloomington Animal Shelter**, en la ciudad de Bloomington, estado de Indiana, Estados Unidos, entre los años 2017 y 2020. A inicios del 2017 el refugio migró a un nuevo programa de gestión permitiéndole generar una base de datos con 7 tablas, con 23 columnas y **más de 2800 registros**. Se complemento con datos extraídos del American Kennel Club sobre las características específicas de las razas.

link:<https://www.kaggle.com/datasets/jinbonnie/animal-data>

soruce: <https://data.bloomington.in.gov/Public-Works/Animal-Shelter-Animals/e245-r9ub>

listado de columnas

Diccionario de Datos



REFUGIO

id_perro
Número de identificación perro

fecha_ingreso
Fecha de ingreso al refugio

id_ingreso
Número de identificación ingreso

fecha_traslado
Fecha de traslado a nuevo hogar

id_traslado
Número de identificación trasladado

fecha_retorno
Fecha de retorno a hogar

id_retorno
Número de identificación retorno

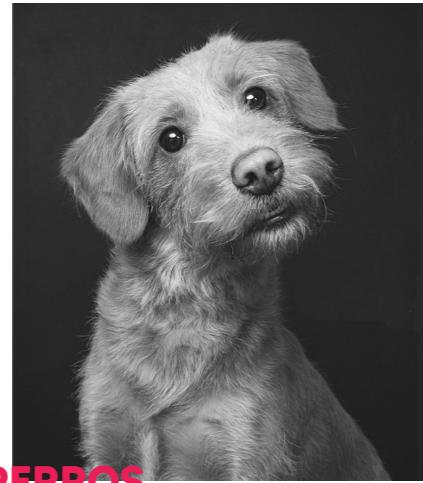
fecha_defunción
Fecha de fallecimiento

id_defución
Número de identificación defunción

estadía_en_refugio
Cantidad de días en refugio

estadía_en_hogar
Cantidad de días en hogar

nuevo_hogar
Exito de nuevo hogar



PERROS

id_perro
Número de identificación perro

id_raza
Número de identificación raza

color_base
Color de pelaje

nombre
Nombre

edad_detallada
Edad detallada

sexo
Sexo

edad
Edad resumida

edad_segmentada
Edad segmentada



RAZAS

id_raza
Número de identificación refugio

raza
Nombre de la raza

cantidad_perros
Cantidad de perros por raza

cantidad_reubicados
Cantidad de reubicados por raza

estadía_promedio
Estadía promedio

porcentaje_exito
Porcentaje de éxito de nuevo hogar

tamaño
tamaño de la raza (altura y peso)

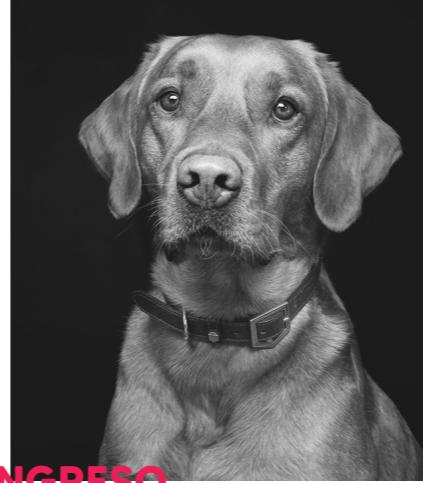
familia
Comportamiento con niños

entrenamiento
Facilidad de adiestramiento

relación
Relacionamiento con otros perros

energía
Nivel de actividad y energía

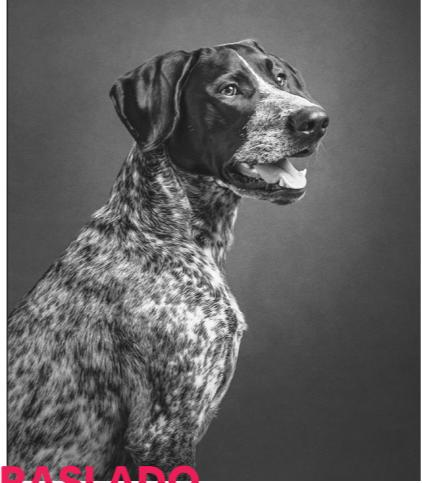
mantenimiento
Facilidad en el cuidado



INGRESO

id_ingreso
Número de identificación ingreso

estado_ingreso
Razón o situación en la que fue encontrado



TRASLADO

id_traslado
Número de identificación traslado

estado_traslado
Tipo de traslado



RETORNO

id_retorno
Número de identificación retorno

estado_retorno
Razón del retorno



DEFUNCIÓN

id_retorno
Número de identificación retorno

estado_retorno
Razón del retorno

Se extrajeron 7 tablas en formato .csv, que posteriormente se fusionaron para crear una sola base de datos conjunta. Luego se le realizaron las transformaciones que el proyecto necesite para su mejor rendimiento.

aprendizaje supervisado

Descripción de modelo

Debido a las características del dataset se plantea abordar el problema de predecir el "**porcentaje de éxito en la adopción**" de perros en un refugio.

Dado que la variable objetivo es continua, la **regresión** fue la opción adecuada para estimar el grado de éxito en la adopción en función de características clave como la **edad**, el **tamaño** y la **raza** del perro.

Además, se va a realizar una comparación entre tres algoritmos de regresión:

- **Linear Regression**
- **Random Forest Regression**
- **Gradient Boosting**



ciclo de data science

Etapas

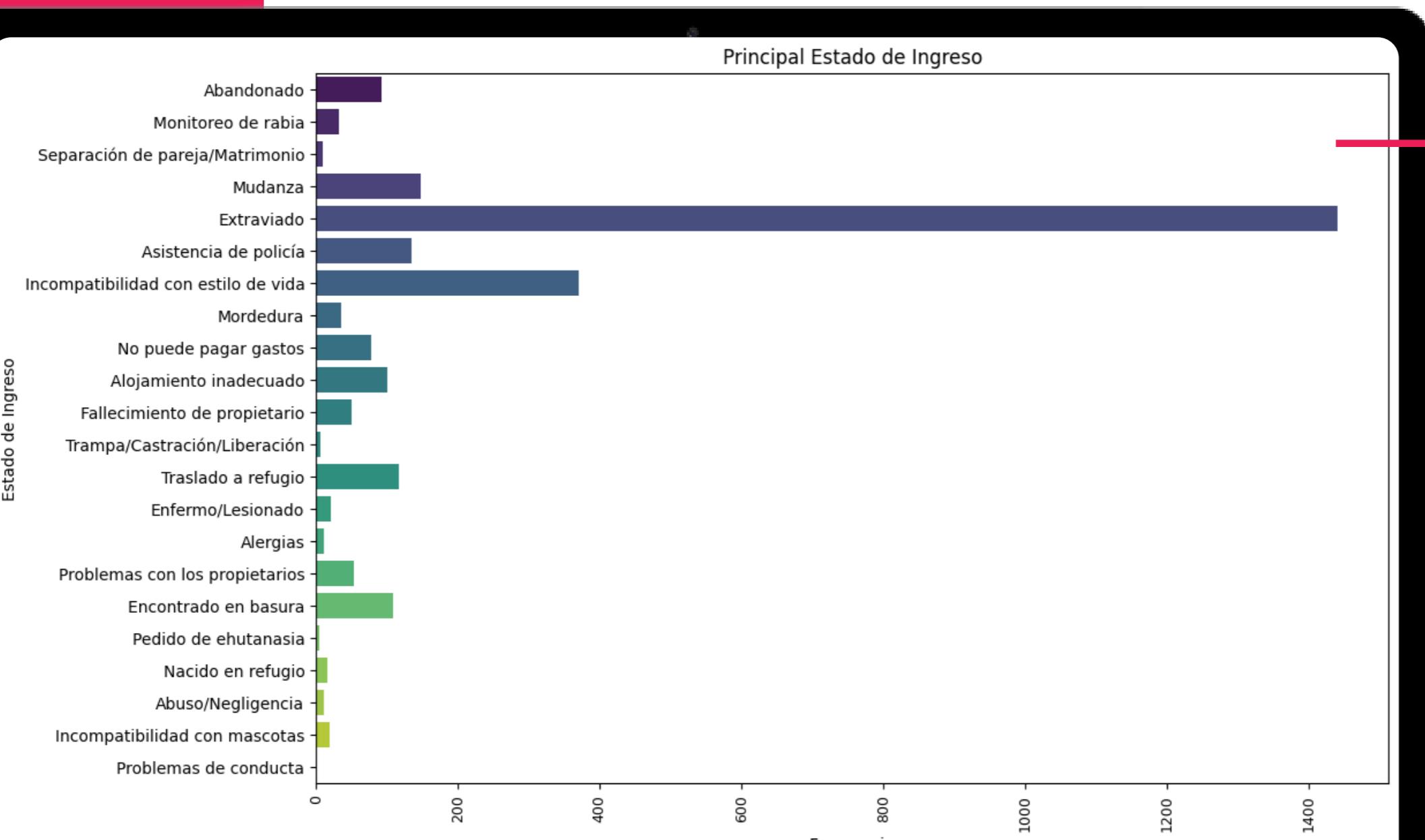
- .data acquisition**..... 01.
.lectura de datos.....
- .data wrangling 1**..... 02.
.merge de dataframes.....
- .data transformation**..... 03.
- .análisis exploratorio de datos**..... 04.
.análisis univariados.....
.análisis bivariados.....
- .data wrangling 2**..... 05.
- .análisis estadístico de datos**..... 06.
- .selección de modelo**..... 07.
- .preprocesamiento**..... 08.
- .entrenamiento**..... 09.
- .comparación de desempeño**..... 10.
- .validación del modelo**..... 11.
.k-fold cross-validation.....
- .optimización del modelo**..... 12.
.grid search y halving grid search.....
.diagnóstico y comparación.....
- .conclusión**..... 13.

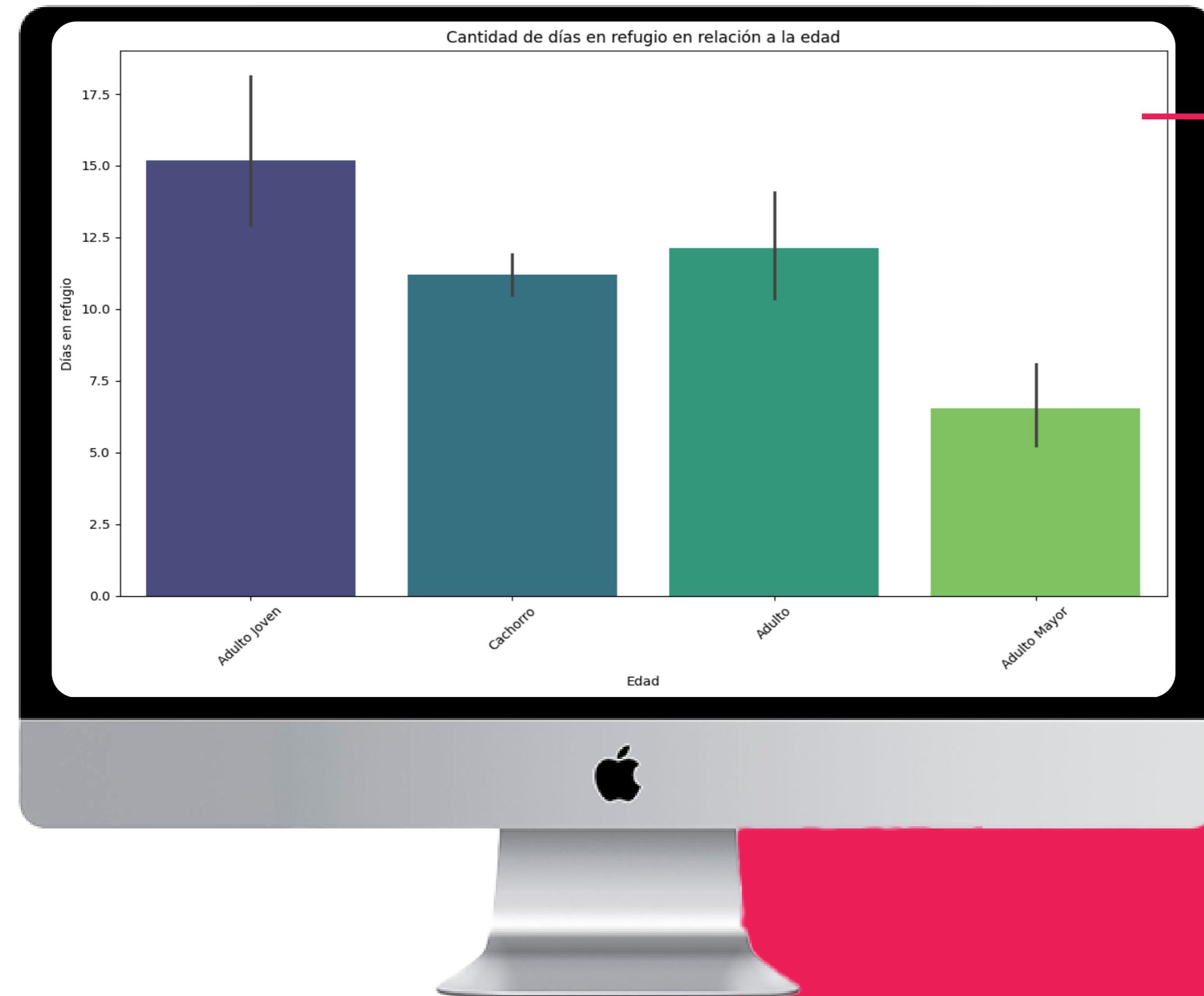


visualización de datos

Estado de ingreso

Este gráfico detalla cual es el estado de ingreso que representa a la mayor cantidad de perros.





visualización de datos

Días en el refugio

En este caso tenemos un gráfico que realciona la cantidad de días que pasan los perros en el refugio, con las edades de los animales. Las edades están segmentadas: cachorros (0 a 2 años), adulto joven (2 a 5 años), adulto (5 a 9 años) y adulto mayor (más de 9 años).

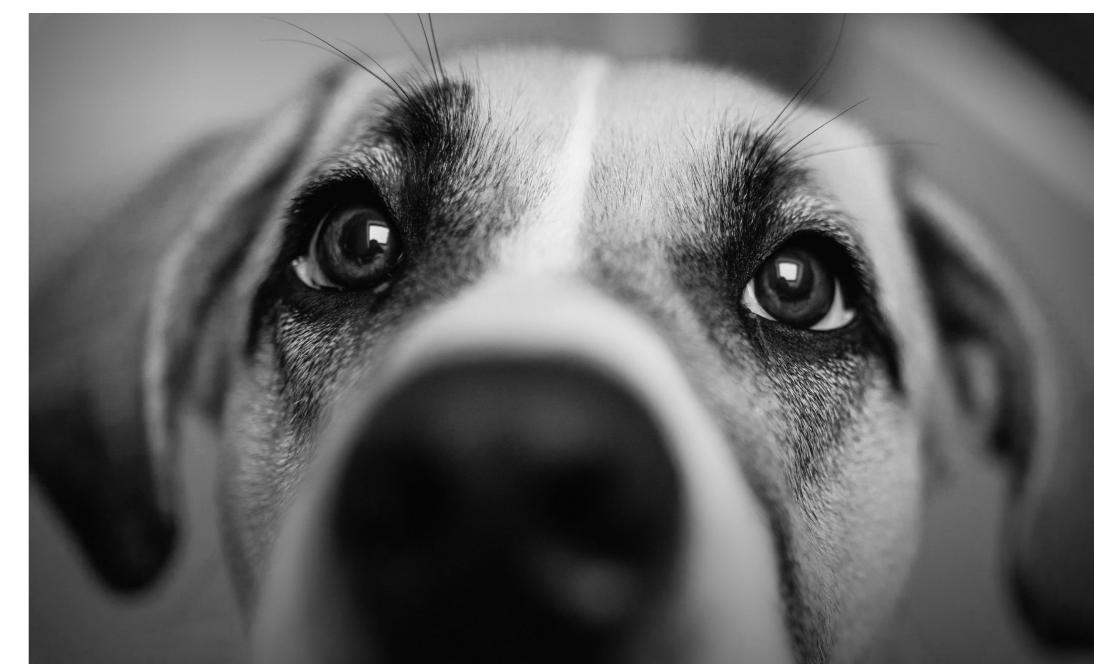




visualización de datos

Razas con mayor éxito

Este es un top 15 de las razas con mayor porcentaje de éxito. Útil para determinar los factores que pueden afectar a la hora de la adopción.



visualización de datos

Razas con menor éxito

Gráfico complementario del anterior, un top 15 de las razas con menor porcentaje de éxito. Necesario para crear un panorama más claro de la situación.



que nos dicen las datos?

Gráficas



La información generada mediante los datos que se manejaron nos permite alcanzar conclusiones que nos proporcionan una comprensión profunda de los factores que influyen en la adopción de animales rescatados.

En primer lugar, se comprobó que el **extravío** es la situación predominante que lleva a los animales a los refugios, y se observó que los perros más jóvenes (hasta 18 meses de vida) son los más afectados por esta situación, responsables de más del 50% de los ingresos.

Además, se encontró que la edad de los animales si influye en su atractivo para los posibles adoptantes ya que los **adultos mayores** tienden a ser adoptados más rápidamente pasando un **promedio de 6 días menos en el refugio** que los cachorros, además de tener una taza de éxito de reubicación mayor. Este hallazgo sugiere la importancia de promover la adopción de animales de todas las edades para evitar la superpoblación en los refugios y garantizar que todos los animales tengan la oportunidad de encontrar un hogar.

En cuanto a las razas, se confirmó que algunas de ellas son adoptadas con mayor facilidad que otras. Las razas como los **Golden Retriever, Yorkshire Terrier o los Ovejero Aleman**, tienden a encontrar hogar de manera mas sencilla con valores de **éxito de adopción superior al 90%**. Esto puede estar influenciado por factores culturales, estereotipos o preferencias populares.

Este resultado resalta la necesidad de educar al público sobre la diversidad de razas y desafiar los prejuicios asociados con ciertas razas consideradas "**peligrosas**", siendo las principales afectadas las integrantes del grupo **Bully Breeds** (Pit Bulls, Bull Terriers, Dogo Argentino, etc.) con **índices de éxito inferiores al 60%**.





como podemos mejorar

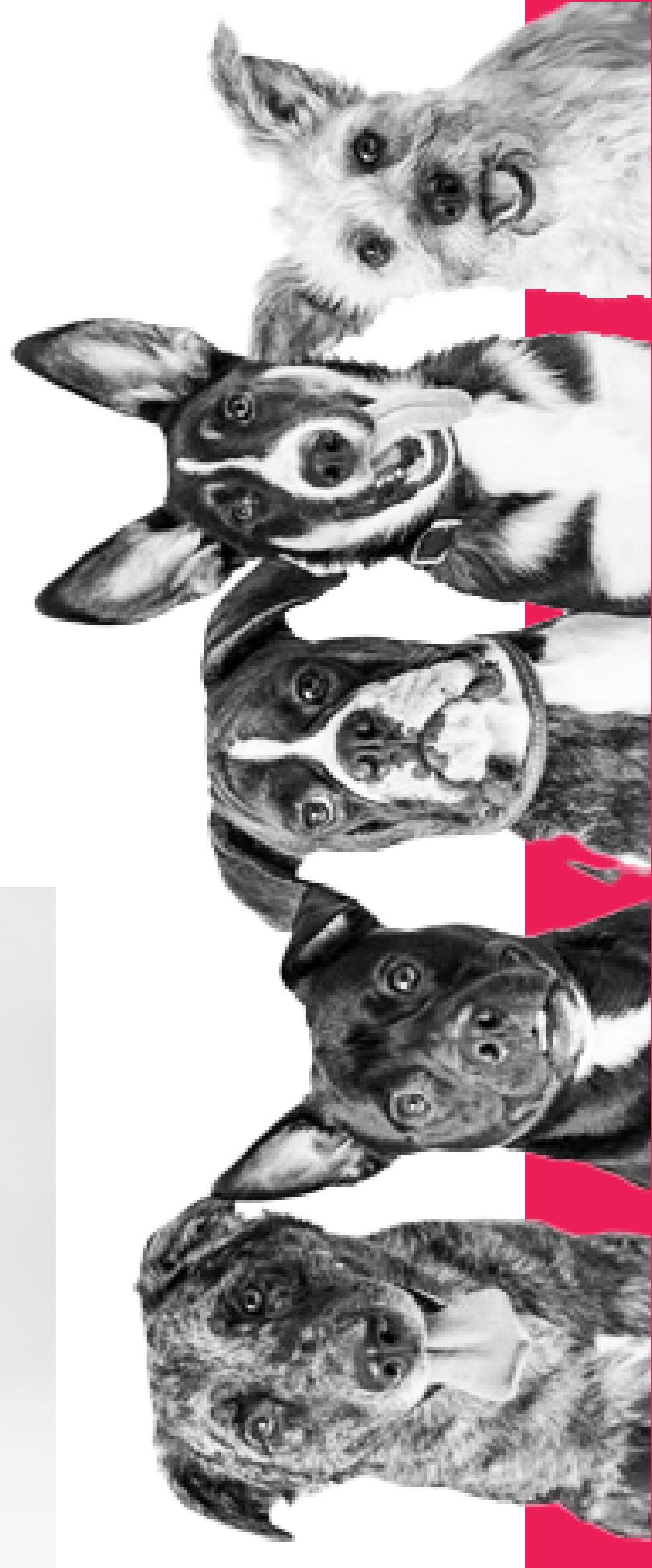
Conclusión

Tras la exploración y análisis de los datos del refugio, se optó por utilizar un modelo de **regresión lineal** optimizado mediante validación cruzada **K-Fold**, el cual demostró ser más robusto y consistente en la predicción del porcentaje de éxito en la adopción de perros. Este modelo se destacó con un **R² promedio de 0.79**, superando los resultados iniciales obtenidos sin validación cruzada.

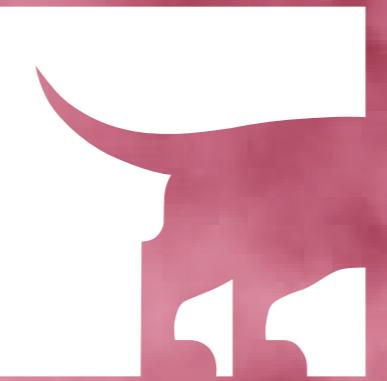
La decisión de utilizar este modelo se basó en su capacidad para ofrecer un equilibrio adecuado entre la precisión y la generalización. Aunque los intentos de ajuste de hiperparámetros no lograron mejorar las métricas, la validación cruzada demostró que el modelo elegido es confiable y menos propenso a sobreajustarse a los datos específicos de entrenamiento.

Este proyecto no solo logró identificar las características más relevantes que influyen en la adopción exitosa de los perros, sino que también ofreció un enfoque sistemático para predecir estas tasas de éxito de manera más eficiente. Las visualizaciones permitieron una mejor comprensión de cómo variables como **la edad, las razones del ingreso a un refugio, y el sexo** impactan en las probabilidades de que un perro encuentre un nuevo hogar.

La implementación de este modelo en el refugio puede guiar las estrategias de adopción, permitiendo personalizar campañas y tomar decisiones más informadas sobre cómo y cuándo promover la adopción de ciertos perros. A largo plazo, esta herramienta puede contribuir a aumentar las tasas de adopción y mejorar la calidad de vida de los perros, así como a optimizar los recursos del refugio al enfocarse en los casos con mayores probabilidades de éxito.



De la Calle al Hogar



Explorando Datos Estadísticos sobre la
Adopción de Animales Rescatados

