

# **BEDU | Procesamiento de datos con Python**



## **Huellas que Encuentran Hogar: Sistema de Adopción Responsable e Inteligente**



### **Equipa 1**

- **Montserrat Flores**
- **Elena Medina**
- **Aurora Pinzón**
- **Montserrat Jiménez**
- **Natalia Esquivel**

## **Introducción**

En México, el abandono de mascotas es un problema serio. Cada año se abandonan alrededor de 500,000 animales, y aproximadamente 27 millones viven en situación de calle. Esta situación genera una presión enorme sobre los albergues y refugios, que muchas veces enfrentan sobrepoblación y altas tasas de devolución de mascotas adoptadas.

Este proyecto busca ayudar a mejorar el proceso de adopción mediante el análisis de datos sobre perros disponibles para adopción. El objetivo es identificar qué perros tienen más dificultades para encontrar hogar y proporcionar información útil para los albergues.

## **Objetivos del Proyecto**

Los objetivos principales fueron:

- Organizar y explorar la información sobre perros disponibles para adopción (edad, peso, raza, color, género)
- Apoyar a los albergues con herramientas para conocer mejor su población de perros
- Identificar tendencias sobre qué mascotas son más vulnerables o tienen más dificultades para ser adoptadas
- Promover la adopción consciente y reducir las devoluciones

## **Metodología**

### **Herramientas utilizadas**

Para este proyecto se usó Python con la biblioteca Pandas, que es muy útil para trabajar con datos organizados en tablas. El archivo de datos era un CSV llamado `dogs_dataset.csv` con información de 3,000 perros.

### **Proceso de la realización del proyecto**

El trabajo se dividió en cuatro etapas principales:

#### **1. Carga y exploración inicial**

Primero se cargó el archivo de datos y se revisó su contenido para entender qué información teníamos. El archivo contenía 5 columnas: raza del perro, edad en años, peso en kilogramos, color y género.

#### **2. Limpieza de datos**

En esta etapa se hizo lo siguiente:

- Se encontraron y eliminaron 5 registros duplicados

- Se verificó que no hubiera información faltante (afortunadamente no había valores nulos)
- Se estandarizó el texto en las columnas de raza, color y género para que estuviera consistente
- Se convirtieron los datos a los tipos apropiados (números para edad y peso, categorías para texto)

Al final de la limpieza quedaron 2,995 registros válidos.

### **3. Transformación de datos**

Esta fue la parte más importante. Se crearon nuevas variables que ayudan a entender mejor la situación de cada perro:

- Grupo de edad: Se clasificó a los perros en tres categorías: Cachorro (hasta 2 años), Adulto (3 a 7 años) y Senior (más de 7 años)
- Tamaño: Según el peso se clasificaron como Pequeño (hasta 15 kg), Mediano (16 a 25 kg) o Grande (más de 25 kg)
- Riesgo de salud: Se identificaron razas con problemas conocidos. Por ejemplo, las razas braquicefálicas (como bulldogs o pugs) tienen problemas respiratorios, mientras que razas grandes (como pastores alemanes o labradores) suelen tener problemas articulares
- Cuidados especiales: Se marcó si el perro necesita cuidados especiales por su edad avanzada o por tener problemas de salud
- Índice de adoptabilidad: Se creó un sistema de puntuación para identificar qué perros tienen más dificultades para ser adoptados. Este índice considera la edad (los senior suman más), el tamaño (los grandes suman más) y los problemas de salud (suman puntos si tienen problemas respiratorios o articulares)

Con base en este índice, cada perro se clasificó en una de tres categorías:

- Alta adoptabilidad: perros que generalmente encuentran hogar más fácil
- Media adoptabilidad: perros con posibilidades normales
- Baja adoptabilidad: perros que enfrentan dificultades serias

### **4. Análisis de resultados**

Se ordenaron los datos poniendo primero a los perros con mayor riesgo de no ser adoptados. También se hicieron análisis cruzados para ver cómo se relacionan diferentes características y se obtuvieron los siguientes resultados.

#### **Resultados**

##### **Distribución general**

La base de datos mostró la siguiente distribución:

Por edad:

- 50% son perros senior
- 36% son adultos
- 14% son cachorros

Por tamaño:

- 63% son perros grandes
- 20% son pequeños
- 17% son medianos

Por salud:

- 85.5% no tienen riesgos específicos conocidos
- 7.4% tienen riesgo respiratorio
- 7.1% tienen riesgo articular

## **Índices de adoptabilidad**

Los resultados del análisis de adoptabilidad fueron:

- Alta adoptabilidad: 979 perros (32.7%)
- Media adoptabilidad: 1,817 perros (60.7%)
- Baja adoptabilidad: 199 perros (6.6%)

Esto significa que aproximadamente 1 de cada 15 perros en la base de datos enfrenta serias dificultades para ser adoptado.

## **Casos prioritarios**

Los casos más críticos identificados fueron:

- 61 perros senior grandes con problemas articulares
- 74 perros senior grandes con problemas respiratorios

Estos grupos requieren atención especial y estrategias específicas para su adopción.

## **Verificación de hipótesis**

Se confirmaron tres hipótesis planteadas al inicio:

1. Los perros senior y de raza grande efectivamente tienen menor adoptabilidad. El 83.7% de los perros senior quedaron en categorías de adoptabilidad media o baja.

2. Las razas con problemas de salud específicos (braquicefálicas con problemas respiratorios y razas grandes con problemas articulares) enfrentan mayores dificultades cuando además son de gran tamaño.
3. Los perros pequeños y cachorros tienen mayor demanda. El 100% de los cachorros pequeños quedaron en la categoría de alta adoptabilidad.

## **Conclusiones**

Este proyecto logró crear un sistema funcional para identificar perros en riesgo de no ser adoptados. El análisis de datos permitió:

- Cuantificar qué características hacen que un perro tenga más dificultades para encontrar hogar
- Identificar casos prioritarios que requieren atención especial
- Proporcionar información que puede ayudar a los albergues en la toma de decisiones

Con esta información, los albergues pueden desarrollar estrategias específicas para promover la adopción de perros que históricamente tienen más dificultades. Por ejemplo, hacer campañas especiales para perros senior o buscar adoptantes específicamente interesados en perros grandes.

## **Limitaciones y trabajo futuro**

Es importante mencionar algunas limitaciones del proyecto:

- Los datos no provienen de albergues reales, además de que el dataset es de Estados Unidos, posiblemente en México encontraríamos mayores casos de baja adoptabilidad
- No se tienen datos reales sobre adopciones exitosas o devoluciones
- Faltan más variables importantes como el comportamiento del perro

Para mejorar este sistema en el futuro se podría:

- Integrar datos reales de refugios mexicanos
- Incluir información sobre el comportamiento y temperamento de cada perro
- Hacer seguimiento de adopciones para validar si las predicciones son correctas
- Combinar o filtrar por información sobre las preferencias de los adoptantes

A pesar de estas limitaciones, el proyecto demuestra que es posible usar análisis de datos para apoyar procesos de adopción más conscientes y reducir el abandono de mascotas.