

## Reto 1: Codificación de Regiones Afectadas

### Objetivo:

Analizar y codificar las regiones en las que ocurrieron escapes de peces, con técnicas de codificación de variables categóricas.

### Instrucciones:

#### Paso 1: Carga y limpieza de datos

- Cargar correctamente el archivo CSV `escapes.csv`, utilizando el encabezado correcto.
- Eliminar filas vacías o incorrectas.
- Verificar si existen valores nulos en las columnas `REGIÓN` o `COMUNA`.

#### Paso 2: Codificación de variables

- Aplicar **Label Encoding** a la columna `REGIÓN`.
  - Aplicar **One-Hot Encoding** a la columna `COMUNA`.
  - Crear una tabla que muestre cuántas columnas nuevas se generaron con cada técnica.
- 

## Reto 2: Codificación de Especies Escapadas

### Objetivo:

Preparar los datos de especies escapadas para análisis automático utilizando codificación categórica.

### Instrucciones:

#### Paso 1: Carga y limpieza de datos

- Cargar el archivo `escapes.csv` y verificar los encabezados.
- Identificar cuántas especies diferentes existen en la columna `ESPECIE`.
- Verificar valores nulos y tratar correctamente la columna `ESPECIE`.

#### Paso 2: Codificación de variables

- Aplicar **Label Encoding** a la columna `ESPECIE`.
  - Aplicar `pandas.get_dummies()` para crear variables dummy de `ESPECIE`, sin eliminar categorías.
  - Comparar visualmente la codificación resultante.
- 

## Reto 3: Codificación del Titular/Operador

## **Objetivo:**

Codificar los titulares de los centros acuícolas que reportaron escapes.

## **Instrucciones:**

### **Paso 1: Carga y limpieza de datos**

- Leer el archivo y confirmar la estructura correcta del dataset.
- Verificar si hay duplicados o errores en la columna TITULAR/OPERADOR.

### **Paso 2: Codificación de variables**

- Aplicar **Label Encoding** a TITULAR/OPERADOR.
  - Aplicar **One-Hot Encoding** utilizando sklearn a la misma columna.
  - Comentar las ventajas y desventajas de cada codificación en este caso.
- 

## **Reto 4: Análisis del Año y Región**

## **Objetivo:**

Codificar las variables temporales y espaciales para análisis posteriores.

## **Instrucciones:**

### **Paso 1: Carga y limpieza de datos**

- Leer el archivo escapes.csv con el encabezado correcto.
- Verificar que la columna AÑO esté en formato numérico.
- Comprobar si hay regiones no válidas o combinadas con otras.

### **Paso 2: Codificación de variables**

- Aplicar **Label Encoding** a AÑO.
- Aplicar **One-Hot Encoding** a la combinación AÑO + REGIÓN.
- Mostrar la cantidad total de columnas generadas.