

Practica 2 - Modulo 2_Curso R Básico - Exploración y Preprocesamiento de Datos en R

Contexto: Este ejercicio tiene como objetivo la exploración inicial y preparación del dataset *PescaPacífico Palangre 2005-2013*, el cual contiene información sobre capturas de especies marinas con palangre en el Pacífico. Realizaremos tareas de inspección, limpieza de datos, y procesamiento con estructuras de datos clave en R como vectores, matrices y dataframes.

1. Preparación del conjunto de datos

1.1 Cargar el archivo de datos

- Carga en R el fichero de datos 'PescaPacífico_Palangre_2005-13.csv' utilizando la función `read.csv()` .
 - Revisa su estructura con `str()` .
-

1.2 Identificación de duplicados

- Verifica si existen filas duplicadas en el set de datos.
 - En caso de haberlas, revísalas y elimínalas.
-

1.3 Identificar valores vacíos o nulos

- Detecta las columnas con valores vacíos o nulos en cada una de las columnas.
 - No olvides columnas enteramente de 0's.
 - Elimina las columnas con más de 50 % de valores nulos o vacíos.
-

1.4 Renombrar columnas para simplificar

- Renombra el nombre de las columnas del dataset para simplificar el procesamiento y análisis posteriori.

- Acorta nombre
 - Elimina espacios
 - Elimina caracteres especiales
-

2. Normalización de variables categóricas

2.1 Corregir tipos de datos erróneos

- Revisa el tipo de datos con el que R ha interpretado cada variable, y sustitúyelo allí donde sea necesario.
-

2.2 Creación de columna

- Genera una nueva columna llamada "ClasificacionAnos" que contenga el valor '2005 - 2008', '2009 - 2010' y '> 2010' en relación a que años incluye.
-

3. Detección de valores erróneos o atípicos

4.1 Verificar valores extremos

- Verifica si existen valores extremos o atípicos (*outliers*) en las columnas numéricas.
 - Si los hay guárdalos en un dataframe independiente para analizarlos posteriormente.
-

3. Trabajo con vectores

3.1 Crear vectores

Extrae el peso capturado del set de datos y las especies en dos vectores independientes.

3.2 Operaciones con vectores

- Muestra las estadísticas básicas del peso.
 - Identifica el boquerón usando `grep1` en el vector `Especie`. Si no aparece, indica cual hay.
-

4. Trabajo con matrices

4.1. Crear matrix numérica

- Genera una matriz con las columnas numéricas del DataFrame (utiliza peso descartado y peso capturado únicamente).
 - Hazla sobre la agrupación de ambos pesos sobre los siguiente niveles de agrupación: Zona.
-

4.2 Operaciones con matrices

- Muestra la media por columna.
 - Muestra la media por fila.
-

5. Guardar datos procesados

Utiliza la función `write.csv` para guardar los datos preprocesados.