

EVALUACIÓN FINAL: ANÁLISIS DE DATOS DE MIGRACIÓN CON NUMPY Y PANDAS

Un instituto de investigación internacional está analizando datos de migración del siglo XXI para identificar tendencias y factores socioeconómicos que afectan los movimientos de población. Se ha entregado un dataset en formato CSV con información sobre migraciones entre distintos países, incluyendo cantidad de migrantes, razones de migración y variables económicas.

El instituto necesita que analices, transformes y prepares los datos para extraer información útil que ayude en la toma de decisiones.

Objetivo

Evaluar la capacidad del estudiante para manejar datos utilizando NumPy y Pandas, aplicando técnicas de limpieza, transformación, agrupamiento y combinación de datos en un contexto realista.

INSTRUCCIONES

La actividad se divide en 5 secciones, cada una con su propio puntaje. Debes realizar cada tarea en Python usando NumPy y Pandas.

1. Limpieza y Transformación de Datos (3 puntos)

- Carga el dataset en un DataFrame de Pandas.
- Identifica y trata valores perdidos en el dataset.
- Detecta y filtra outliers usando el método del rango intercuartílico (IQR).
- Reemplaza los valores de la columna "Razon_Migracion" usando mapeo de valores (ejemplo: "Económica" → "Trabajo", "Conflicto" → "Guerra").

2. Análisis Exploratorio (2 puntos)

- Muestra las 5 primeras filas del dataset.
- Obtén información general del dataset con `.info()` y `.describe()`.
- Calcula estadísticas clave:
 - Media y mediana de la cantidad de migrantes.
 - PIB promedio de los países de origen y destino: Usa `.value_counts()` para contar cuántos movimientos de migración ocurrieron por cada razón.

3. Agrupamiento y Sumarización de Datos (2 puntos)

- Agrupa los datos por "Razon_Migracion" y calcula la suma total de migrantes para cada categoría.
- Obtén la media del IDH de los países de origen por cada tipo de migración.
- Ordena el DataFrame de mayor a menor cantidad de migrantes.

4. Filtros y Selección de Datos (2 puntos)

- Filtra y muestra solo las migraciones por conflicto.
- Selecciona y muestra las filas donde el IDH del país de destino sea mayor a 0.90.
- Crea una nueva columna "Diferencia_IDH" que calcule la diferencia de IDH entre país de origen y destino.

5. Exportación de Datos (1 punto)

- Guarda el DataFrame final en un nuevo archivo CSV llamado "Migracion_Limpio.csv", sin el índice.

INSTRUCCIONES ADICIONALES:

- Puntos totales = 10.
- Descarga el material complementario migracion.csv
- Comprimir el archivo completo en formato .zip o .rar.
- Sube el archivo a la plataforma.