



EVALUACIÓN FINAL- OBTENCIÓN Y PREPARACIÓN DE DATOS

Objetivo

Aplicar técnicas de obtención, limpieza y preparación de datos, utilizando criterios de imputación y manipulando las estructuras de datos a conveniencia para satisfacer las necesidades de información acorde al lenguaje python

Situación:

Una organización sin fines de lucro te ha solicitado apoyar en el análisis de datos de encuestas de satisfacción sobre sus servicios comunitarios. Los datos provienen de distintas fuentes (archivos CSV, Excel y una API pública), presentan problemas de formato, valores faltantes y necesitan ser transformados para facilitar su análisis.

Tu tarea será obtener los datos, prepararlos y dejarlos listos para un análisis posterior, utilizando Python y librerías como numpy, pandas y otras herramientas utilitarias.

Requerimientos de la Actividad:**1. Obtención de Datos**

- Extrae datasets desde al menos dos fuentes distintas (CSV, Excel, JSON o API REST usando requests).
- Carga los datos utilizando pandas, verificando su correcta lectura y estructura.

2. Exploración Inicial

- Utiliza métodos básicos (head(), info(), describe(), value_counts(), etc.) para explorar el set de datos y detectar posibles problemas.
- Aplica filtros y selecciones con loc[] o iloc[] para identificar registros específicos.

3. Limpieza y Tratamiento de Valores Faltantes

- Identifica valores nulos o inconsistentes y aplica criterios de imputación (por ejemplo: eliminar, reemplazar por media/mediana/moda o crear indicadores).
- Justifica brevemente la estrategia utilizada.

4. Manipulación de Datos con Numpy y Pandas

- Realiza transformaciones utilizando numpy para resolver un problema simple (por ejemplo: normalización, operaciones vectorizadas).
- Realiza transformaciones estructurales usando pandas: renombrar columnas, cambiar tipos de datos, crear nuevas variables a partir de otras, etc.

5. Data Wrangling y Enriquecimiento

- Aplica técnicas de data wrangling para enriquecer y transformar el dataset (unificar columnas, reestructurar datos, combinar campos).
- Utiliza `merge()`, `groupby()`, `pivot_table()` o `melt()` para reorganizar los datos según una necesidad específica.

Entregables:

1. Notebook en Jupyter o archivo .py que contenga:
 - Código funcional con comentarios explicativos.
 - Limpieza y preparación completa del dataset.
 - Aplicación de técnicas de manipulación, exploración y transformación.
 - Justificación breve (comentarios en el código) de las decisiones tomadas en cada paso.
2. Informe resumen (máx. 1 página):
 - Descripción general del proceso aplicado.
 - Problemas detectados y cómo se abordaron.
 - Preparación final del dataset y posibles usos posteriores.