

Retos Integradores: Limpieza, Discretización y Análisis de Datos

Reto 1: Clasificación de Temperaturas Diarias

Archivo: reto1_temperatura_ampliado.csv

Columnas:

- ID
- Ciudad
- Temperatura_Maxima
- Humedad_Relativa

Objetivos:

1. Detectar y rellenar valores nulos en Temperatura_Maxima con la media.
 2. Discretizar con `pd.cut()` en categorías:
 - o Frío (≤ 15)
 - o Templado (15-25)
 - o Caluroso (> 25)Resultado: Categoria_Temp
 3. Analizar diferencias de temperatura por ciudad.
 4. Investigar relación entre Temperatura_Maxima y Humedad_Relativa.
-

Reto 2: Cuartiles de Calificaciones Académicas

Archivo: reto2_calificaciones_ampliado.csv

Columnas:

- ID
- Genero
- Materia
- Calificacion

Objetivos:

1. Eliminar duplicados y detectar valores nulos en Calificacion.
 2. Imputar nulos con la **media por materia**.
 3. Aplicar `pd.qcut()` para categorizar en: Q1, Q2, Q3, Q4.
Resultado: Nivel_Calificacion
 4. Analizar diferencias por género y materia.
-

Reto 3: Segmentación por Años de Experiencia

Archivo: reto3_experiencia_ampliado.csv

Columnas:

- ID
- Edad
- Experiencia
- Area

Objetivos:

1. Detectar nulos en `Experiencia` y rellenar con la mediana por Area.
 2. Detectar y tratar outliers.
 3. Clasificar con `pd.cut()` en:
 - o Principiante (0-2)
 - o Intermedio (3-5)
 - o Avanzado (6-10)
 - o Experto (>10)Resultado: `Grupo_Experiencia`
 4. Analizar relación entre `Edad` y `Experiencia`.
-

Reto 4: Clasificación Socioeconómica por Ingreso

Archivo: `reto4_ingresos_ampliado.csv`

Columnas:

- ID
- Ingreso_Mensual
- Escolaridad
- Contrato

Objetivos:

1. Detectar y eliminar outliers en `Ingreso_Mensual` usando IQR.
 2. Rellenar valores nulos con la media por `Escolaridad`.
 3. Agrupar con `pd.qcut()` en:
 - o Muy Bajo
 - o Bajo
 - o Medio
 - o AltoResultado: `Segmento_Ingreso`
-

Reto 5: Categorización de Satisfacción del Usuario

Archivo (sugerido): `reto5_satisfaccion.csv`

Columnas:

- ID
- Satisfaccion (0-100)

Objetivos:

1. Verificar si hay valores fuera de rango (0-100).
 2. Rellenar valores nulos con la **moda**.
 3. Clasificar con `pd.cut()` en:
 - o Insatisfecho (0-50)
 - o Neutral (51-75)
 - o Satisfecho (76-100)Resultado: `Categoria_Satisfaccion`
-

Reto 6: Análisis de Precios de Productos

Archivo (sugerido): reto6_precios_productos.csv

Columnas:

- ID
- Producto
- Precio

Objetivos:

1. Validar que `Precio` no tenga texto y esté dentro de un rango realista.
2. Rellenar valores nulos con la **mediana**.
3. Aplicar `pd.qcut()` para categorizar en:
 - o Económico
 - o Intermedio Bajo
 - o Intermedio Alto
 - o PremiumResultado: `Grupo_Precio`