

# Documentación Técnica del Dataset: Optimización de Espacios y Eficiencia Energética en la UAB

## DESCRIPCIÓN

Este conjunto de datos contiene información detallada sobre la ocupación de espacios y aulas en la Universidad Autònoma de Barcelona (UAB). Incluye datos sobre grupos de estudiantes, características de los espacios, horarios de uso y otros atributos relevantes.

**Los participantes del reto deberán realizar el preprocesamiento y limpieza de los datos crudos proporcionados** para analizar y optimizar el uso de los espacios universitarios, con el objetivo de mejorar la eficiencia energética y la sostenibilidad.

## ESTRUCTURA

Los datasets se proporcionan en la siguiente carpeta:

- **datasets/**: Contiene los datasets originales sin procesar. **Parte del reto consiste en realizar el preprocesamiento y limpieza de estos datos** para su posterior análisis.

### Contenido de la Carpeta **datasets/**

- grupos.csv
- ubicaciones.csv
- características.csv
- recursos\_características.csv
- franjas\_media\_hora.csv
- franjas\_cuarto\_hora.csv
- calendario\_grupos.csv

## DATASETS

### 1. grupos.csv

Contiene información detallada sobre los grupos de estudiantes.

Campo	Tipo	Descripción	Unidades	Ejemplo
ID_GRUPO	string	Identificador del grupo		"2024-MA T1"
IND_ALUMNOS_ GRUPO_PREV	integer	Número de alumnos previstos en el grupo	estudiantes	50
IND_ALUMNOS_ GRUPO_REAL	integer	Número real de alumnos en el grupo	estudiantes	45
ID_CURSO_ACA DEMICO	integer	Año académico		2024
TIPO_ACTIVIDA D	string	Tipo de actividad (ver diccionario)		"Teoría"

## 2. ubicaciones.csv

Contiene datos sobre las ubicaciones disponibles (aulas y otros espacios).

Campo	Tipo	Descripción	Unidades	Ejemplo
ID_UBICACION	string	Identificador único de la ubicación (aula)		"A101"
DS_UBICACION	string	Descripción de la ubicación		"Aula 101"
ID_TECNICA	integer	Identificador del tipo de espacio		16
CAPACIDAD	integer	Capacidad máxima del aula	estudiantes	60
SUPERFICIE	float	Superficie del aula	metros cuadrados (m <sup>2</sup> )	85.5
TIENE_CLIMATIZACION	boolean	Indica si el aula tiene climatización		True
ID_CENTRO	integer	Identificador del centro o facultad		102

### 3. **caracteristicas.csv**

Contiene las características de los espacios, definiendo qué instalaciones o recursos adicionales tiene cada aula.

Campo	Tipo	Descripción	Unidades	Ejemplo
ID_CARACTERISTICA	integer	Identificador de la característica		7
DS_CARACTERISTICA	string	Descripción de la característica		"Proyector"
TIPOLOGIA	string	Tipología del espacio (ver diccionario)		"Aula"

### 4. **recursos\_caracteristicas.csv**

Relaciona los recursos disponibles con sus características específicas.

Campo	Tipo	Descripción	Unidades	Ejemplo
ID_RECURS	integer	Identificador del recurso		503
ID_CARACTERISTICA	integer	Identificador de la característica asociada		7
VALOR	string	Valor específico de la característica		"Sí"

## 5. franjas\_media\_hora.csv

Define las franjas horarias de media hora disponibles para la asignación de horas de clase.

Campo	Tipo	Descripción	Unidades	Ejemplo
ID_MITJA_HORA	integer	Identificador de la franja de media hora		1
DS_MITJA_HORA	string	Descripción de la franja horaria		"08:00-08:30"

## 6. franjas\_cuarto\_hora.csv

Define las franjas horarias de cuarto de hora disponibles para la asignación de horas de clase.

Campo	Tipo	Descripción	Unidades	Ejemplo
ID_QUART_HORA	integer	Identificador de la franja de cuarto de hora		1
DS_QUART_HORA	string	Descripción de la franja horaria		"08:00-08:15"
ID_MITJA_HORA	integer	Identificador de la franja de media hora asociada		1
IND_HORAS_POSIBLES	float	Horas posibles en la franja	horas	0.25

## 7. calendario\_grupos.csv

Asigna grupos a horarios y espacios, detallando las horas asignadas y los espacios utilizados.

Campo	Tipo	Descripción	Unidades	Ejemplo
ID_GRUPO	string	Identificador del grupo		"2024-MAT1"
ID_UBICACION	string	Identificador de la ubicación (aula)		"A101"
ID_QUART_HORA	integer	Identificador de la franja de cuarto de hora		32
ID_FECHA_GRUPO	date	Fecha de la clase		"2024-10-25"

## DICCIONARIOS

### 1. TIPOLOGIA en características.csv

Valor	Descripción
"Aula"	Espacio destinado a clases teóricas
"Laboratorio"	Espacio destinado a prácticas y experimentos
"SalaReuniones"	Espacio para reuniones y conferencias

## 2. TIENE\_CLIMATIZACION en ubicaciones.csv

Valor	Descripción
True	El espacio dispone de sistema de climatización
False	El espacio no dispone de sistema de climatización

## 3. TIPO\_ACTIVIDAD en grupos.csv

Valor	Descripción
"Teoría"	Clase teórica
"Práctica"	Clase práctica
"Seminario"	Sesión de seminario o taller
"Laboratorio"	Sesión en laboratorio

## UNIDADES

- **Estudiantes:** Cantidad de alumnos (número entero).
- **Superficie:** Metros cuadrados (m<sup>2</sup>).
- **Tiempo:** Horas (h) y minutos (min).
- **Consumo Energético:** Kilovatios hora (kWh) (si se considera en el análisis).

# TAREAS

## 1. LIMPIEZA DE DATOS

- **Gestión de Valores Nulos:**
  - Identificar campos con valores nulos o faltantes.
  - Decidir si imputar valores, eliminar registros incompletos o manejar nulos en el análisis.
- **Estandarización de Formatos:**
  - Asegurar que los tipos de datos sean consistentes (e.g., fechas en formato correcto, campos numéricos sin caracteres no numéricos).
  - Unificar formatos de texto (e.g., nombres de campos, valores categóricos).
- **Validación de Coherencia:**
  - Verificar que los identificadores clave permiten unir los datasets correctamente.
  - Comprobar que los datos relacionados son consistentes entre sí (e.g., el número de estudiantes no excede la capacidad del aula asignada).

## 2. PROCESAMIENTO

- **Unión de Datasets:**
  - Integrar los diferentes datasets ([grupos.csv](#), [ubicaciones.csv](#), [calendario\\_grupos.csv](#), etc.) en un dataset consolidado que permita el análisis conjunto.
- **Cálculo de Ocupación:**
  - Calcular la ocupación real vs. la capacidad de las aulas.
  - Determinar la eficiencia de uso de los espacios (e.g., porcentaje de ocupación).
- **Análisis Temporal:**
  - Analizar patrones de ocupación a lo largo del tiempo.
  - Identificar franjas horarias y días con mayor o menor uso de espacios.

## 3. OPTIMIZACIÓN

- **Propuestas de Reasignación:**
  - Desarrollar algoritmos o modelos para reasignar grupos a aulas de manera más eficiente.
  - Considerar restricciones como capacidad, características necesarias y horarios disponibles.
- **Eficiencia Energética (Opcional):**
  - Podéis considerar aspectos relacionados con el consumo energético.
  - Al no disponer de parámetros energéticos validados, es válido realizar estimaciones o asumir valores razonables basados en investigación propia.
- **Validación de Restricciones:**
  - Asegurar que no se producen solapamientos horarios en las aulas.
  - Respetar la capacidad máxima y las características requeridas para cada grupo.



# CONSIDERACIONES

## Restricciones Importantes

1. **Capacidad de las Aulas**
  - No asignar grupos a aulas con capacidad inferior al número de estudiantes.
  - Superficie mínima por estudiante: **1.5 m<sup>2</sup>**.
2. **Características de las Aulas**
  - Las aulas deben cumplir con las características necesarias para cada tipo de actividad (e.g., proyector, laboratorio equipado).
3. **Horarios y Solapamientos**
  - No se permite que dos grupos ocupen el mismo espacio en la misma franja horaria.
  - Respetar las franjas horarias definidas en [franja\\_media\\_hora.csv](#) y [franja\\_cuarto\\_hora.csv](#). **Para simplificar puedes utilizar [franja\\_media\\_hora.csv](#)**
4. **Eficiencia Energética\***
  - Considerar el impacto energético de mantener aulas ocupadas vs. vacías.
  - Buscar minimizar el número de aulas en uso sin comprometer la calidad educativa.

## \*Consideración de Parámetros Energéticos (Opcional)

Al no disponer de parámetros energéticos validados, la consideración de variables relacionadas con el consumo energético es opcional y queda a criterio de cada grupo.

### Algunas ideas:

- Podéis investigar valores típicos de consumo energético en aulas y edificios educativos.
- Fuentes como informes oficiales, estudios académicos o datos públicos pueden ser útiles.
- Asumir valores estándar para el consumo de iluminación, climatización, etc. Por ejemplo, estimar un consumo promedio por metro cuadrado.
- Incorporar variables como el tipo de iluminación, presencia de climatización, etc., si se considera relevante.
- Analizar cómo la ocupación de las aulas puede influir en el consumo energético.
- Estimar el consumo energético de las aulas en función de las variables consideradas.
- Comparar el consumo antes y después de las propuestas de optimización.
- Además de optimizar la asignación de espacios, puedes plantear objetivos relacionados con la reducción del consumo energético.
- Es importante documentar claramente las asunciones realizadas y las fuentes de información utilizadas.

### Variables adicionales:

- Puedes considerar variables adicionales que puedan afectar al análisis y la optimización.
  - Ejemplos de variables a considerar:
    - **Orientación del Aula:** Norte, Sur, Este, Oeste.
    - **Tipo de Iluminación:** LED, Fluorescente, Incandescente.
    - **Antigüedad del Edificio o Aula.**
    - **Características Específicas:** Accesibilidad, equipamiento especial, etc.
- Evaluar cómo estas variables pueden influir en el uso y eficiencia de los espacios.
- Añadir estas variables al dataset si se dispone de la información o si se pueden inferir razonablemente.

### Posibilidades y sugerencias:

- **Análisis Exploratorio:**
  - Visualizar la distribución de tamaños de grupos y capacidades de aulas.
  - Identificar aulas subutilizadas o sobreutilizadas.
- **Modelos de Optimización:**
  - Aplicar técnicas de programación lineal o entera para asignar grupos a aulas.
  - Utilizar algoritmos heurísticos o metaheurísticos si el problema es muy grande.
- **Visualización de Resultados:**
  - Crear mapas de calor para mostrar la ocupación de aulas en diferentes horarios.
  - Generar informes que destaquen las mejoras logradas con las propuestas.
- **Documentación de Procesos:**
  - Documentar los procesos, asunciones y decisiones.
  - La claridad en la presentación de resultados y metodología será valorada.
- **Eficiencia Computacional:**
  - Dado el tamaño del dataset, es importante considerar la eficiencia computacional de los algoritmos implementados.

## NOTAS FINALES

- Asegúrese de manejar correctamente los formatos de fechas y horas al unir y analizar los datos.
- Los campos **ID\_GRUPO**, **ID\_UBICACION**, **ID\_QUART\_HORA**, etc., son clave para relacionar los datasets.
- Aunque los datos han sido anonimizados, trate toda la información con confidencialidad.

Si tiene preguntas técnicas sobre los datos o las tareas, puede contactar con el equipo organizador, a través del canal de Discord.