



Práctica. Descripción

- Una cadena de supermercados desea **identificar áreas** en las que abrir nuevos centros.
- Utilizando **información censal** (archivo census2000_conLongitudLatitud.csv) el objetivo es **agrupar zonas censales cuyos perfiles de clientes se adapten a diferentes tipologías de centros**: boutiques, supermercados de presupuesto medio, grandes superficies, etc.
- Las variables de las que dispone para ello son:
 - MEANHHSIZE.- Tamaño medio de la unidad familiar (HH = Household).
 - MEDHHINC.- Nivel de ingresos mediano de la unidad familiar.
 - REGDENS.- Densidad de población de la región.
 - REGPOP.- Número de habitantes de la región.
- Las variables LocX,LocY, no deben entrar en la segmentación, si bien se pueden utilizar para representar las zonas en un mapa y colorearlas en función del identificador de cluster (opcional).
- Se deberá **entregar un archivo en formato .PDF** en el que:
 - Se justifiquen las decisiones tomadas para determinar el número de clusters establecido: basar la decisión en un método jerárquico y realizar la segmentación final con un algoritmo tipo k-means.
 - Se presenten las características de las regiones agrupadas (a través de los centroides) y se asocie un nombre a los mismos.