

## EJERCICIO DE CLASE

### ANÁLISIS DE CLUSTER

El fichero BARRIOS contiene información socio-económica de algunos barrios de Madrid. Para reducir el número de variables e intentar encontrar relaciones, tanto entre variables como entre provincias, realizar los siguientes apartados.

1. Para explorar el conjunto de datos, representar el mapa de calor interactivo de los valores de las variables en los barrios.
2. Calcular la matriz de distancias entre los barrios con las variables sin estandarizar y estandarizadas.
  - a. Comparar los gráficos que representan dichas distancias tanto con la función `fviz_dist` como `heatmaply`. ¿Se pueden observar grupos de barrios?
3. Realizar un análisis Jerárquico de clusters para determinar si existen grupos de barrios con comportamiento similar.
  - a. Realizar una agrupación jerárquica con los datos sin estandarizar y otra con los datos estandarizados representando ambos dendrogramas. Comentar las diferencias. ¿Cuántos clusters recomendarías?
  - b. Utilizando los datos estandarizados, representar los individuos en los planos de las primeras Componentes, agrupados según el número de clusters elegido.
  - c. ¿Qué número óptimo de clusters nos indican los criterios Silhoutte y de Elbow?
  - d. Con el número de clusters decidido en el apartado anterior realizar un agrupamiento no jerárquico.
    - i. Representar los clusters formados.
    - ii. Evaluar la calidad de los clusters
4. Mostrar una tabla con los barrios que forman cada uno de los clusters y comentar cuales son las características socioeconómicas de cada cluster