

## Agrupación basada en centro representativo

Agrupamos los perfiles diarios de consumo de cada edificio en k grupos de manera independiente para cada edificio.

Luego, para cada edificio se selecciona el centro que incorpora un mayor número de perfiles diarios de consumo, al cual llamaremos el centro representativo del edificio. Utilizando dicho centro se busca el número óptimo de agrupaciones (con los mismos criterios que el caso anterior).

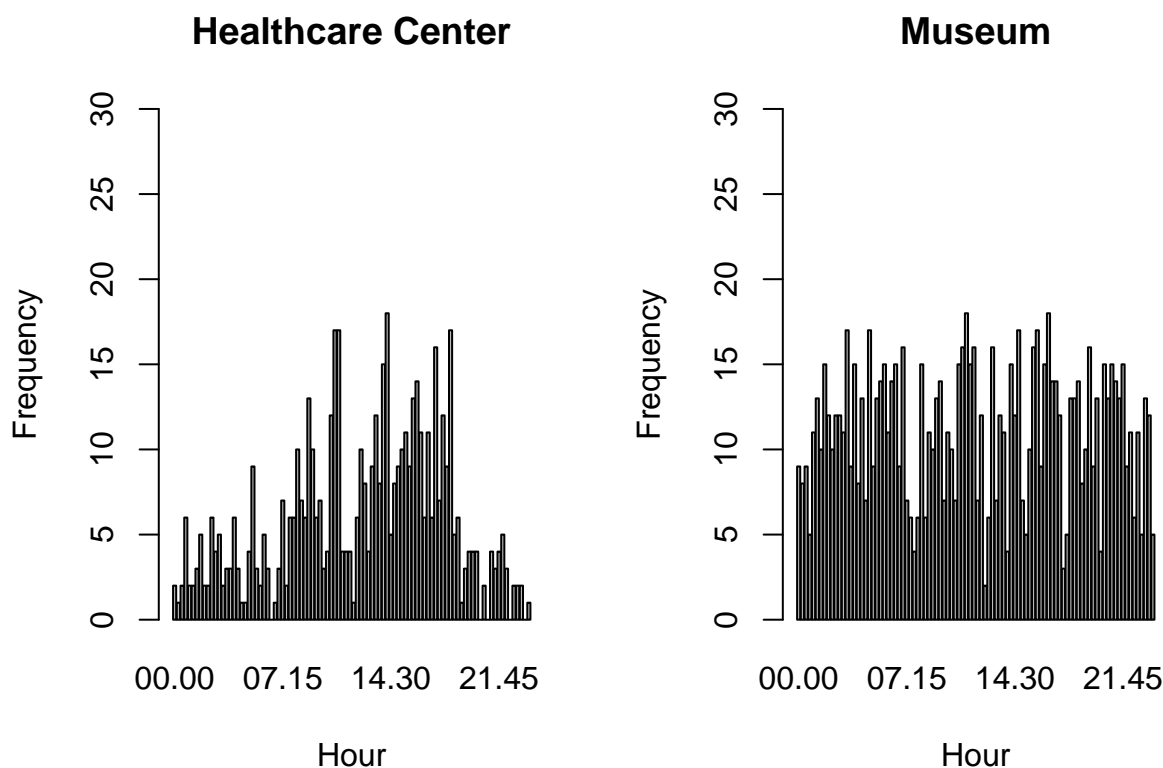
A continuación se muestran los resultados:

##	Score	Method	Clusters
## Connectivity	3.4900794	kmeans	2
## Dunn	0.4197119	hierarchichal	4
## Silhouette	0.7492431	kmeans	2

## Agrupación basada en atributos representativos

Utilizando la técnica propuesta para encontrar el subconjunto de atributos mínimo sin pérdida de información se plantea sintetizar la información de un edificio. Para ello se ejecuta 30 veces de manera independiente el algoritmo para cada edificio, obteniendo 30 subconjuntos de atributos. A partir de estas soluciones se construye el histograma de frecuencia de aparición de cada atributo en las soluciones.

Por ejemplo, a continuación se muestra el histograma de frecuencia de los atributos escogidos para dos edificios en particular (*Centro de Salud Arroyo de la Miel* y *Museo Picasso*).



Utilizando los criterios previamente definidos buscamos el tamaño óptimo de agrupaciones. Los resultados obtenidos se muestran a continuación.

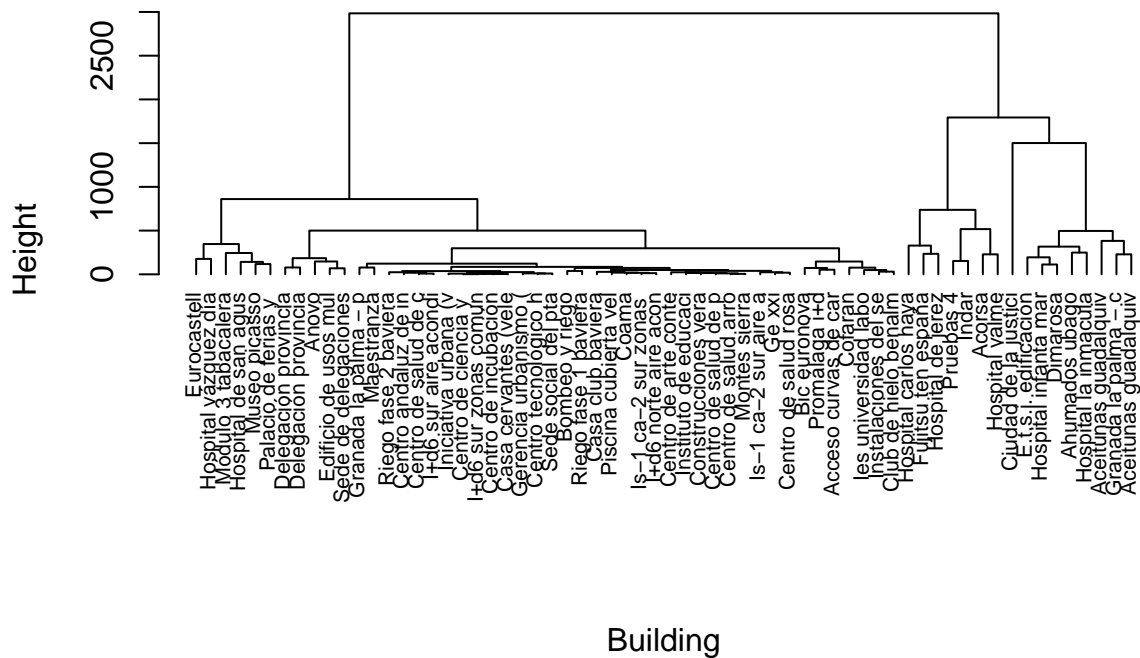
##	Score	Method	Clusters
## Connectivity	12.912302	kmeans	2
## Dunn	0.624596	kmeans	8
## Silhouette	0.291140	som	2

## Comparación de resultados

El método propuesto basado en el subconjunto de atributos presenta los mejores resultados según la métrica de **Dunn**. En cambio, el método propuesto de utilizar el centro “representativo” obtiene el mejor resultado para las métricas **Connectivity** y **Silhouette**.

A continuación se muestran los dendrogramas para las dos mejores propuestas. A simple vista se puede observar una gran diferencia en el balanceo entre el primero y el segundo dendrograma. Por otra parte, es posible apreciar que el segundo dendrograma (frecuencia de atributos) presenta una mejor agrupación en términos “semánticos”, por ejemplo la mayoría de los hospitales se encuentran a poca distancia, lo cual también se verifica para los centros de salud, y para otros edificios de características similares.

### Representative Center Dendrogram



Attribute Freq Dendrogram

