

Actividad 2

Agrupamiento o Clustering

Usar el dataset 'divorce_data.csv', correspondiente a datos de divorcio y consecuencias en adultos jóvenes, que es un estudio centrado en examinar las consecuencias del divorcio reciente de los padres para jóvenes de 18 a 23 años, que se divorciaron en los 15 meses siguientes a la primera oleada del estudio (1990-91). La muestra consistía en 257 encuestados con padres recién divorciados. El archivo 'divorce_data.csv' es un subconjunto con 47 encuestados y 7 variables, las cuales se detallan en la siguiente tabla.

Variable	Descripción
DIVYEAR	Año en el que los padres se divorciaron. Variable dicotómica 1989 y 1990
Momint	Relación con la madre (1 = muy próxima; 2 = próxima; 3 = normal; 4 = lejana)
Dadint	Relación con el padre (1 = muy próxima; 2 = próxima; 3 = normal; 4 = lejana)
Momclose	Cuánto de cerca está el niño con la madre (1 = extremadamente cerca; 2 = bastante cerca; 3 = algo cerca; 4 = nada cerca en absoluto)
Depression	Sentimiento de depresión en las últimas 4 semanas (1 = a menudo; 2 = a veces; 3 = casi nunca; 4 = nunca)
Livewith	1 = vive solo con la madre; 2 = vive sólo con el padre; 3 = vive con ambos padres
Gethitched	Plan de matrimonio del encuestado (1 = casarse bastante pronto; 2 = casarse algún día; 3 = no casarse nunca; 4 = no sabe/no contesta)

Realizar en primer lugar un *clustering* jerárquico para obtener una aproximación del número de grupos, para poder llevar a cabo un agrupamiento *k-means*.

¿Cuáles son las características que representan mayoritariamente los distintos grupos?