

# *Clustering*

## Estadística Multivariante

Sofía Almeida Bruno  
Daniel Bolaños Martínez  
José María Borrás Serrano  
Fernando de la Hoz Moreno  
Pedro Manuel Flores Crespo  
María Victoria Granados Pozo

16 de enero de 2020

# Clustering

- Objetivo: agrupar objetos similares.
- Dadas  $x_1, \dots, x_n$  medidas de  $p$  variables en  $n$  objetos considerados *heterogéneos*. El objetivo del análisis clúster es agrupar estos objetos en  $k$  clases *homogéneas*, donde  $k$  es también desconocido.

# Clustering

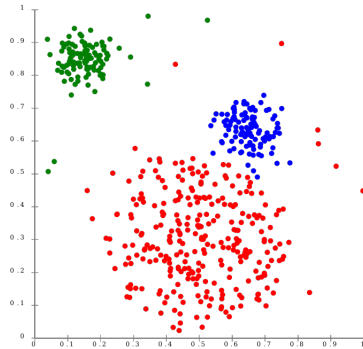


Figura: Ejemplo de *clustering*. [Chi11]

## Ejemplos de *Clustering*

- Biología: determinación de especies.
- *Marketing*: descubrimiento de grupos de clientes.

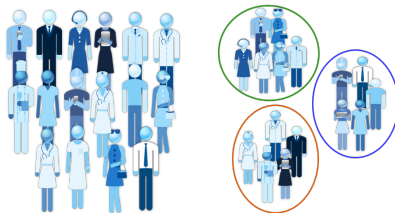


Figura: Ejemplo de *clustering*. [noa]

- Psicología: encontrar tipos de personalidad.
- Arqueología: datar objetos encontrados.
- Planificación urbana: identificar grupos de viviendas.

# *Clustering*

Para realizar un análisis clúster hay que:

- Elegir una medida de similitud.
- Elegir un algoritmo para construir los grupos.
  - ▶ Particionamiento.
  - ▶ Jerárquicos.

## Referencias I



Chire, *Cluster analysis with optics on a density-based data set.*,  
[https://commons.wikimedia.org/wiki/File:  
OPTICS-Gaussian-data.svg](https://commons.wikimedia.org/wiki/File:OPTICS-Gaussian-data.svg), October 2011.



*Understanding data mining clustering methods.*