



ESCUELA DE INGENIERÍA
FACULTAD DE INGENIERÍA

EDUCACIÓN
PROFESIONAL

Diplomado en Big Data y Ciencia de Datos
Ciencia de Datos y sus Aplicaciones

Clase 07: Segmentación

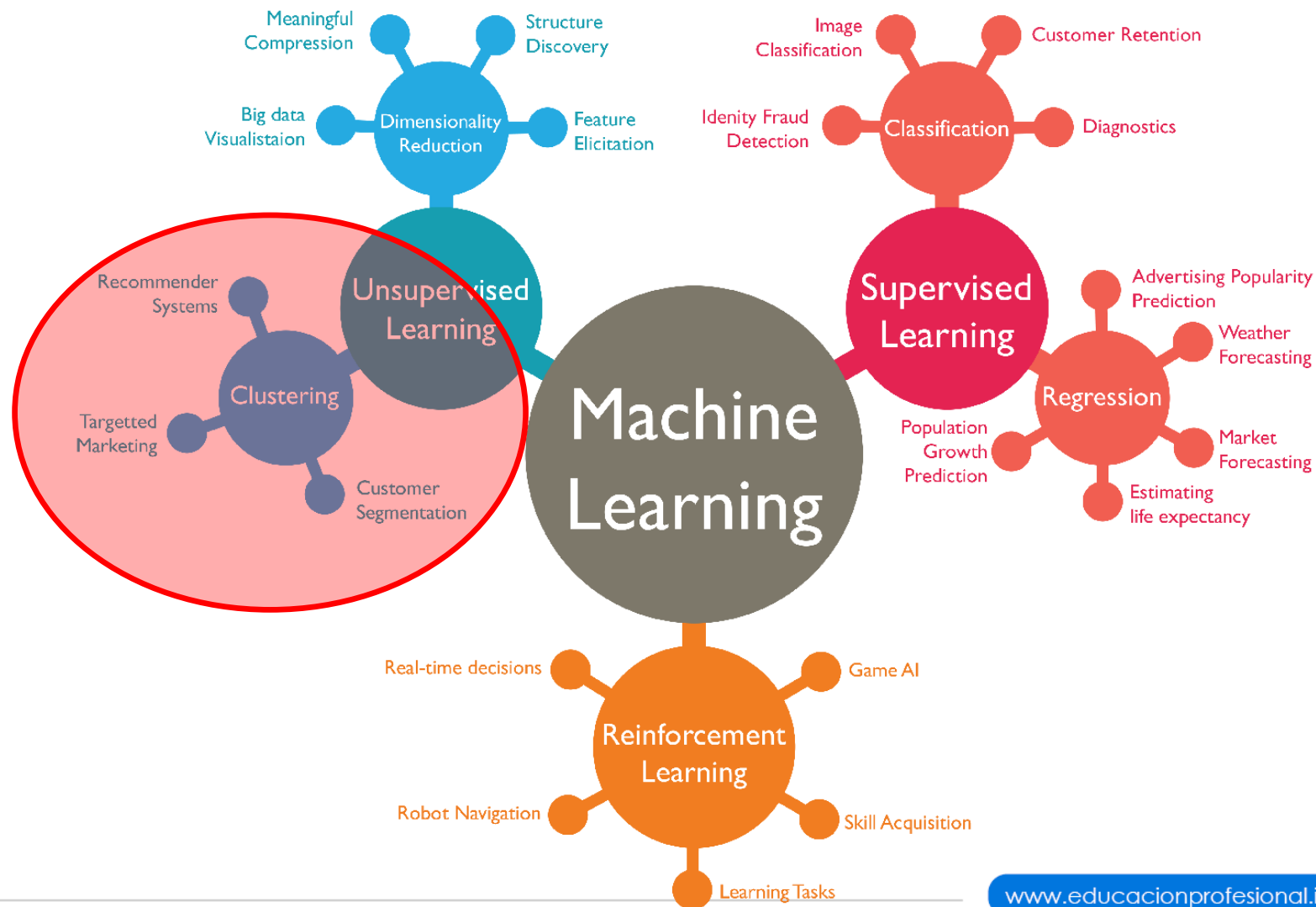
Roberto González



regonzar@uc.cl



Tipos Machine Learning

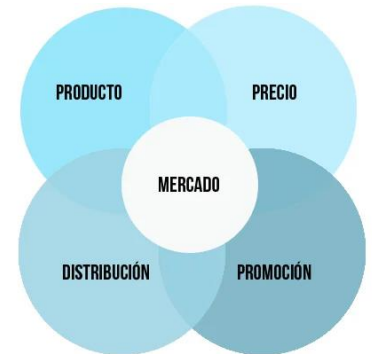




Clase 07: Segmentación

SEGMENTACIÓN EN MARKETING

Segmentación en Marketing



- Al decidir segmentar uno establece un “Marketing Mix” para cada segmento
- La segmentación no sólo es la base de una estrategia de marketing, también lo es de la planificación estratégica
- Las empresas usan distintos tipos de segmentación para distintas estrategias
- Una buena segmentación debe identificar a los grupos de valor, ser sencilla de implementar y comunicar , y sobre todo facilitar el éxito en las acciones



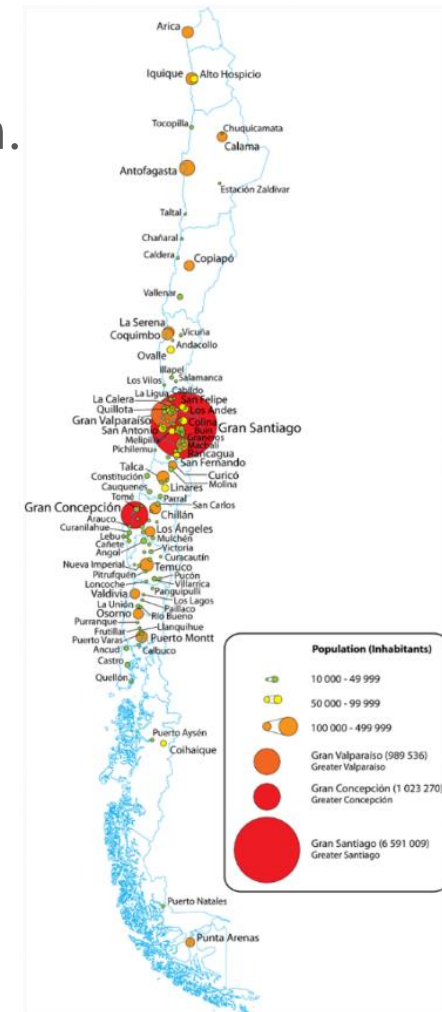
Tipos de Segmentación en Marketing

- Segmentación geográfica
- Segmentación demográfica
- Segmentación conductual
- Segmentación psicográfica



Segmentación Geográfica

- Fundamental en productos de difícil distribución.
- En Chile... Clima
- Distintos Niveles, Regional, Comunal, etc.



Segmentación Demográfica

- Son las más usadas
- Pueden ser:
 - Edad
 - Sexo
 - Ingreso o Grupos Socio Económicos
 - Ocupación o Profesión





Segmentación Conductual

- Ocasión de uso
 - Una persona puede pertenecer a distintos segmentos dependiendo de la ocasión
- Beneficio buscado
 - ¿Cuál es el beneficio que la persona busca en el producto?
- Tipo de usuario
 - Las personas pueden ser no usuarios, ex usuarios, usuario habitual o usuario nuevo



Segmentación Psicográfica

- Puede ser más poderosa ya que existen personas que pertenecen al mismo segmento demográfico y son muy distintos entre sí
- Incorporan estilos de vida y personalidad de los compradores
- Generalmente las empresas no disponen de este tipo de bases de datos, por lo que es difícil de usar
 - Es necesario sondear cada vez que se requiere



Clase 07: Segmentación

ENFOQUE Y MARCO DE TRABAJO

Enfoque segmentado

Un viaje de transformación hacia la orientación al cliente para racionalizar y priorizar mejor las ofertas en todas las líneas de negocio.

Aprovechar los datos y conocimientos para:



Refinar el modelo de segmentación existente



Completar una evaluación del perfil del cliente para identificar mejor los deseos, necesidades e inhibidores de los segmentos objetivo



Desarrollar un modelo de segmentación con el detalle adecuado para ser aprovechado en toda la organización



Definir mapas de viaje del cliente para apoyar los objetivos de la organización



Priorizar las capacidades para mejorar la experiencia, con aportes de valor real

Un enfoque segmentado requiere:

1

De un equipo interfuncional con un profundo conocimiento de la industria y de la organización, junto con experiencia en el diseño de estrategias de clientes.

2

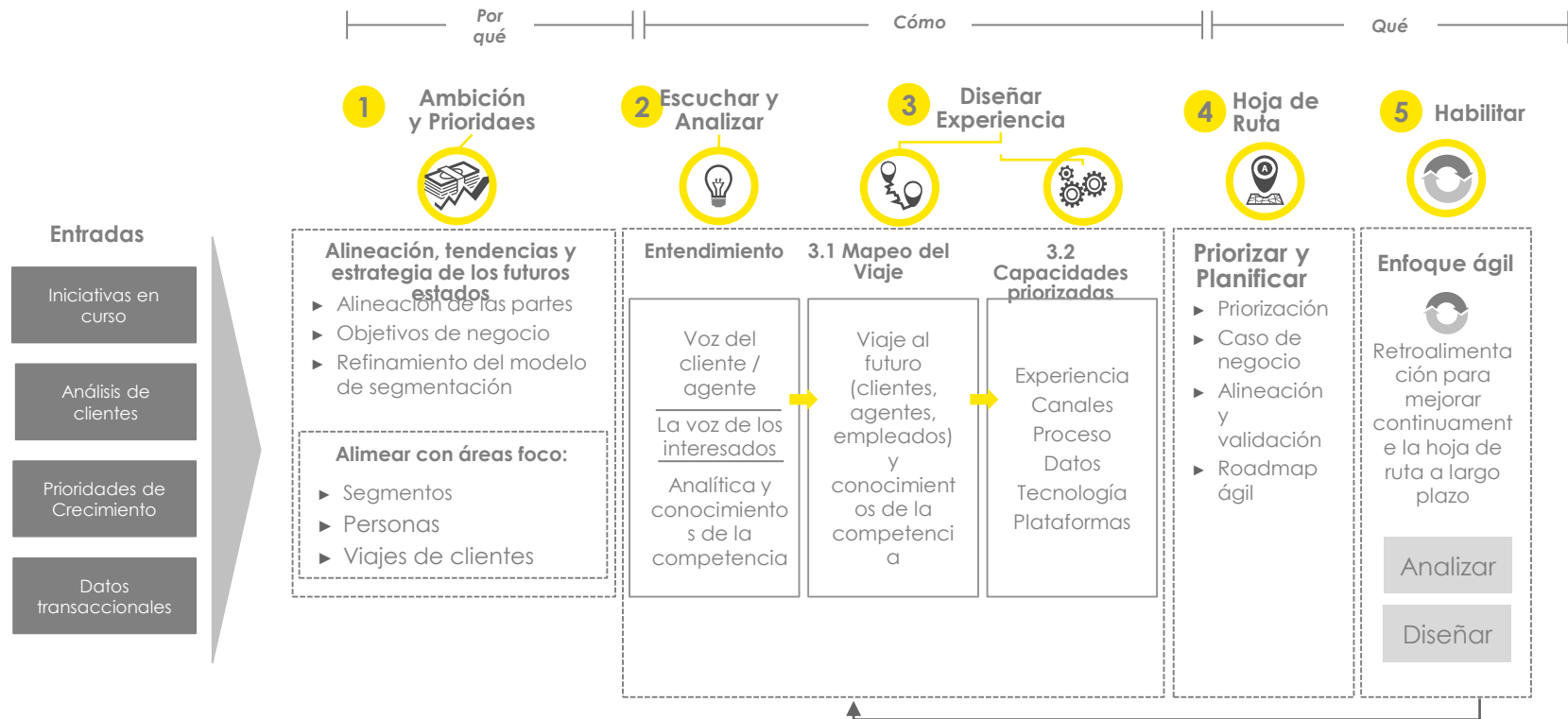
Un marco de trabajo que se adapte para fomentar la colaboración y ofrecer soluciones tangibles centradas en el cliente.

3

Mirada innovadora, pensamiento de diseño iterativo y rápidamente habilitante.

Mirada Metodológica

Metodología que acelera el viaje de organización hacia la centralidad en el cliente

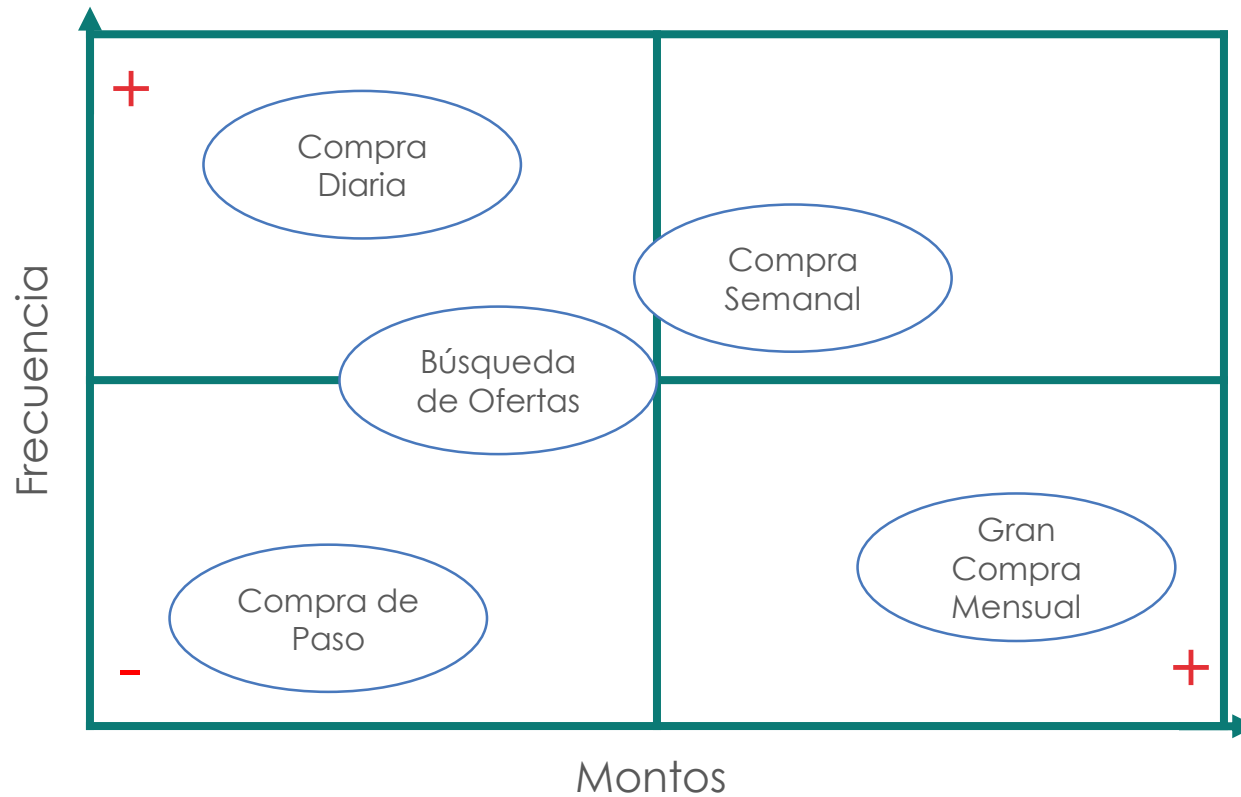




Clase 07: Segmentación

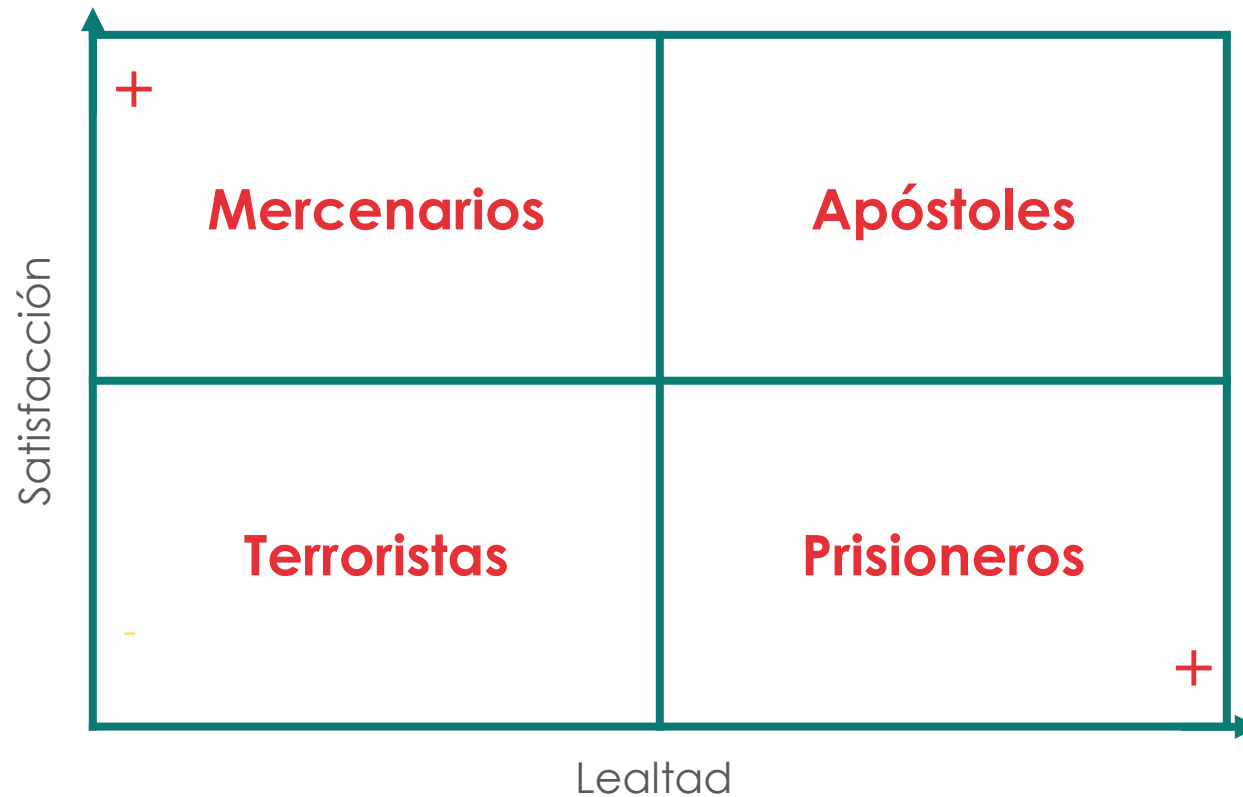
ALGUNAS SEGMENTACIONES

Por monto y frecuencia de compra

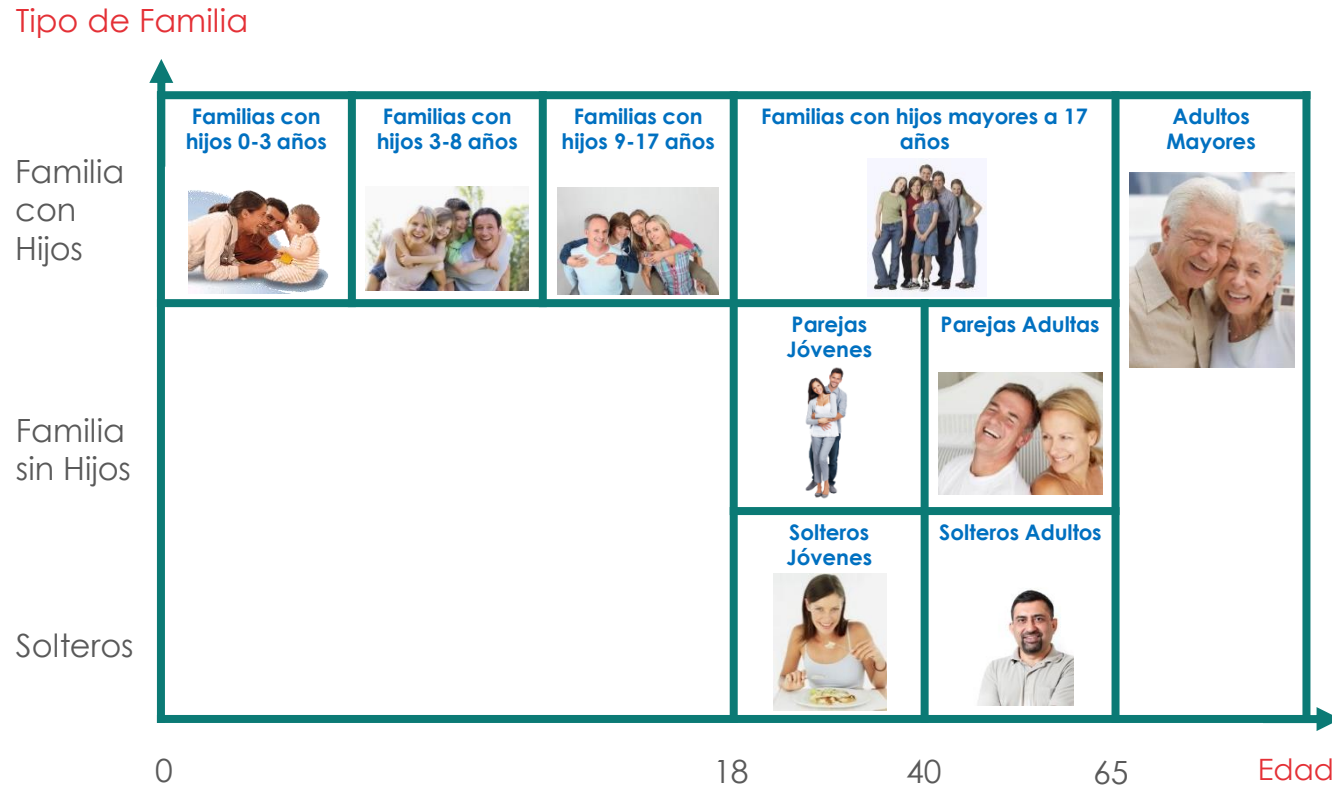




Por satisfacción y lealtad



Por ciclo de vida de las personas





Perfiles Psicográficos: “Los siete tipos de chilenos”

Estudio Brand Asset Valuator, Y&R

1. Los "mainstreamers"

- **Porcentaje de chilenos "mainstreamers": 33%**
- **Su meta:** la seguridad
- **Su motivación:** la pertenencia social
- **Su valor principal:** la aceptación social

2. Los exploradores

- **Porcentaje de Exploradores en Chile: 9%**
- **Su Meta:** el descubrimiento
- **Su Motivación:** la búsqueda
- **Su Valor:** la libertad

3. Los simuladores

- **Porcentaje de chilenos simuladores: 16%**
- **Su meta:** el estatus
- **Su motivación:** crear una impresión
- **Su valor:** la admiración

7. Los disconformes

- **Porcentaje de chilenos disconformes: 5%**
- **Su meta:** el escape
- **Su motivación:** rebelarse
- **Su valor principal:** la superación

4. Los resignados

- **Porcentaje de chilenos resignados: 5%**
- **Su meta:** la supervivencia
- **Su motivación:** el instinto
- **Su valor principal:** la subsistencia

5. Los reformadores

- **Porcentaje de los chilenos reformadores: 13%**
- **Su meta:** la autoexpresión
- **Su motivación:** la independencia
- **Su valor principal:** la individualidad

6. Los exitosos

- **Porcentaje de chilenos exitosos: 19%**
- **Su meta:** el control
- **Su motivación:** resolver desafíos
- **Su valor:** el reconocimiento

TIPOLOGÍA DE COMPRADORES CHILENOS



Atributo más destacado por perfil de consumidor



De
acuerdo
con



Consumidor Economicista

Antes de comprar,
siempre cotizo y
comparo precios



Consumidor Tradicional

Prefiero comprar marcas
que conozco y que me
han dado un buen
resultado en el pasado



Consumidor Hedonista

Acostumbro a
premiarme de vez en
cuando comprándome
algo



Consumidor Excluyente

Estoy dispuesto a pagar
más por marcas más
exclusivas



Consumidor Consciente

Elijo marcas o productos
que ayudan al cuidado
del medio ambiente

CONOCIENDO A LOS
COMPRADORES CHILENOS

chileOscopio
TENDENCIAS 2016 2016

Por consumo cerveza (México)

Segmentos y Características	Importancia del precio al momento de la compra	Toman otras bebidas alcohólicas	Susceptibilidad al cambio de marca	Enfocados al sabor	Enfocados en la imagen	Situación de consumo predominante
Buena Vida	Importante	Si	Si	No	Ser del grupo al tomar lo de moda	Salir a divertirse y amigos
Amiguero	Regular	Si	Si	No	Integrarme a un grupo	La familia y los amigos
Buscando destacar	Importante	Medio	Medio	No	Darme a notar con lo que tomo	Salir a divertirse y amigos en ambiente de energía
Enfocados al sabor	Regular	No	No	Si	No	Toma su cerveza en toda situación
Buscando escapar	Regular	No	No	Si	No	Toma su cerveza en casa y en toda situación
Arraigados	Regular	No	No	Si	No	Salir a divertirse y amigos en ambiente de conversación
No involucrados	Muy importante	No	Si	Medio	No	En casa y con la familia

Tabla 8. Segmentos y características de los consumidores de cerveza.

Fuente: Grupo Modelo

2016

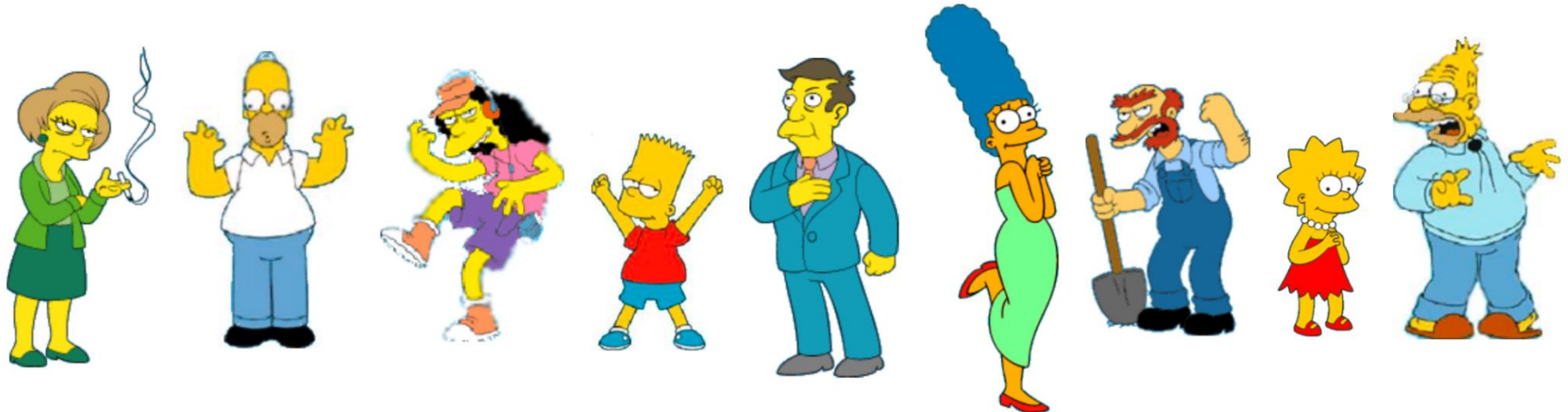


Clase 07: Segmentación

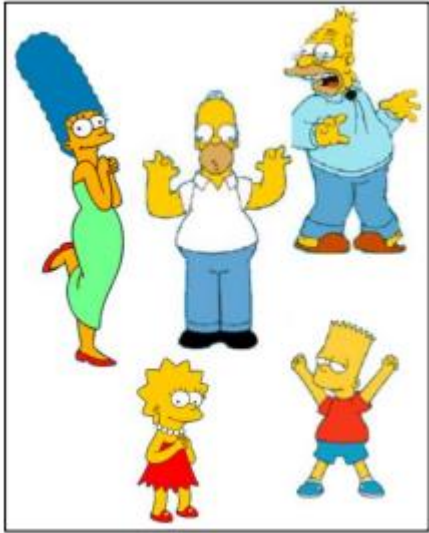
MÉTODOS DE CLUSTERING

¿Cómo segmentar?

Indique cómo agruparía a estos personajes



¿Cómo segmentar?



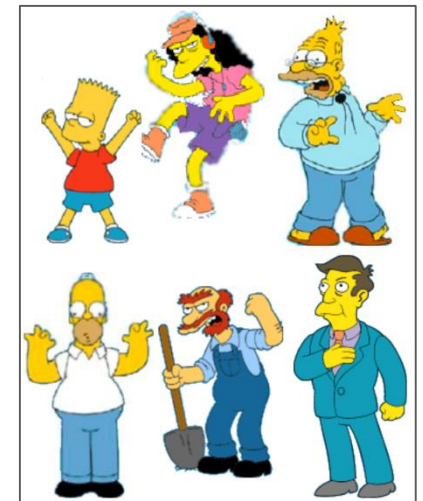
Los Simpson



Empleados del
colegio



Mujeres

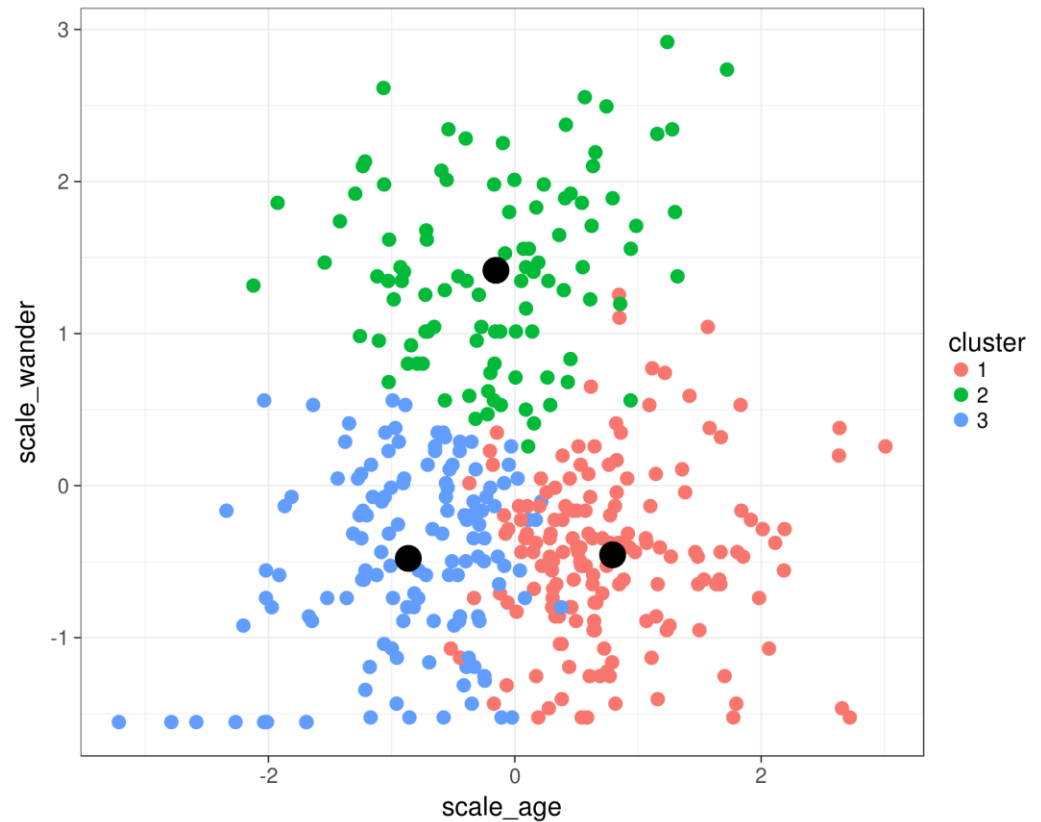


Hombres

Depende del objetivo
¿Cómo definimos similitud?

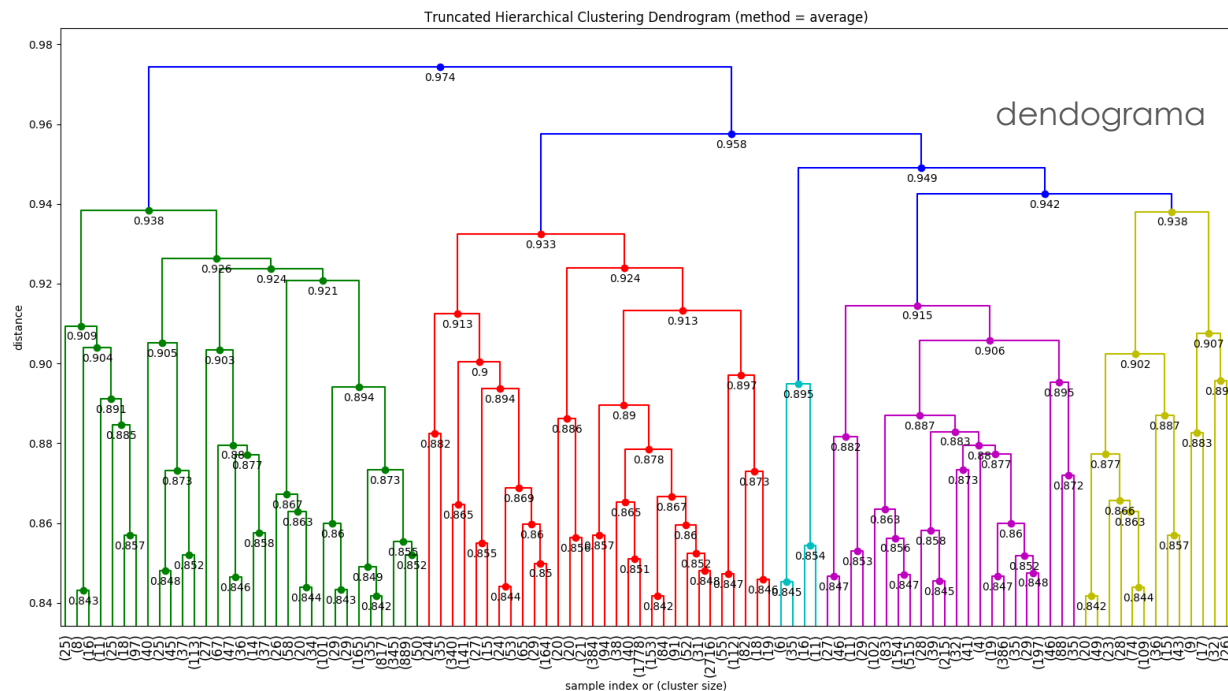
Principales métodos

- Hierarchical clustering
- K-means
- Gaussian mixture



Hierarchical clustering

- Partir con cada objeto en su propio cluster
- Encontrar el mejor par para crear nuevo cluster
- Repetir hasta fusionar todos los cluster
- Hay una familia de algoritmos para distintos tipos de uniones o *linkage*, Aglomerativo, Ward, Single linkage



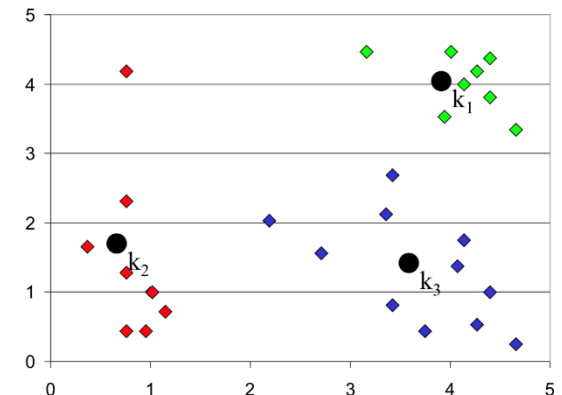
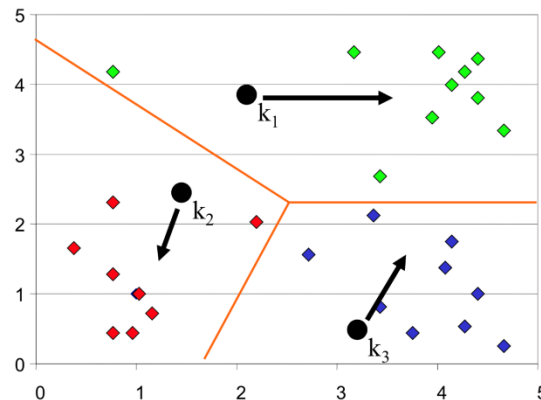
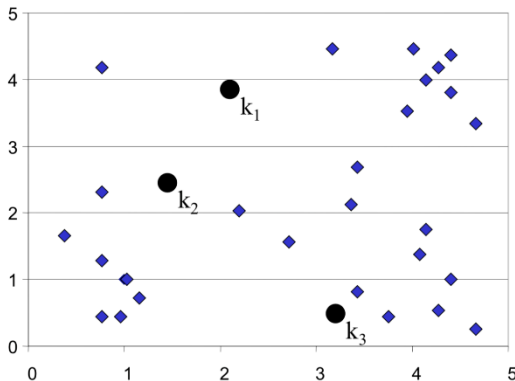


Hierarchical clustering

- Fortalezas
 - No necesita especificar N° clusters
 - Es intuitivo para ciertos problemas
- Debilidades
 - No escala bien. Complejidad $O(n^2)$
n: N de objetos
 - Suele converger en óptimo local
 - Interpretación de resultados es subjetiva

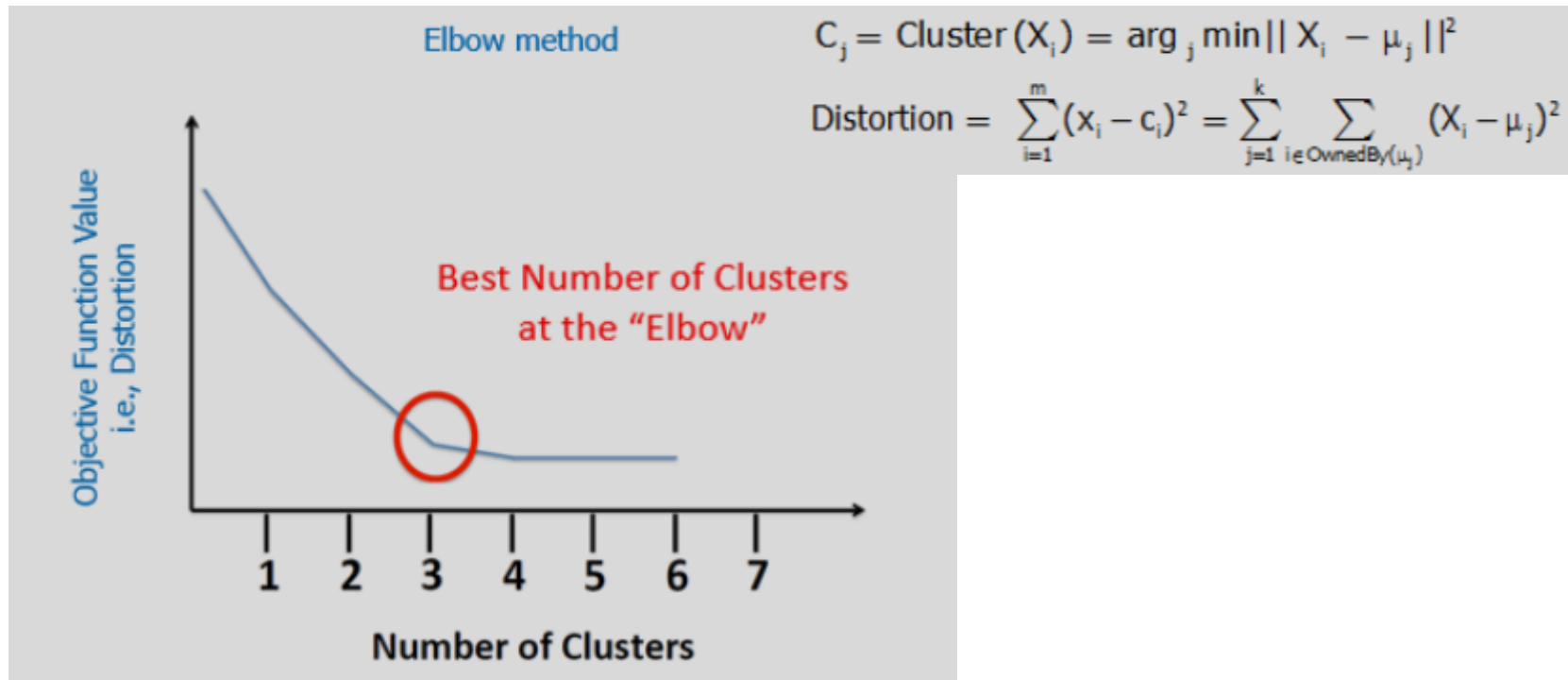
K-means

- Definir $K = N^{\circ}$ clusters
- Elegir de manera aleatoriamente los centros de K
- Asignar los objetos al centro más cercano
- Calcular el centro de K usando objetos asociados



K-means

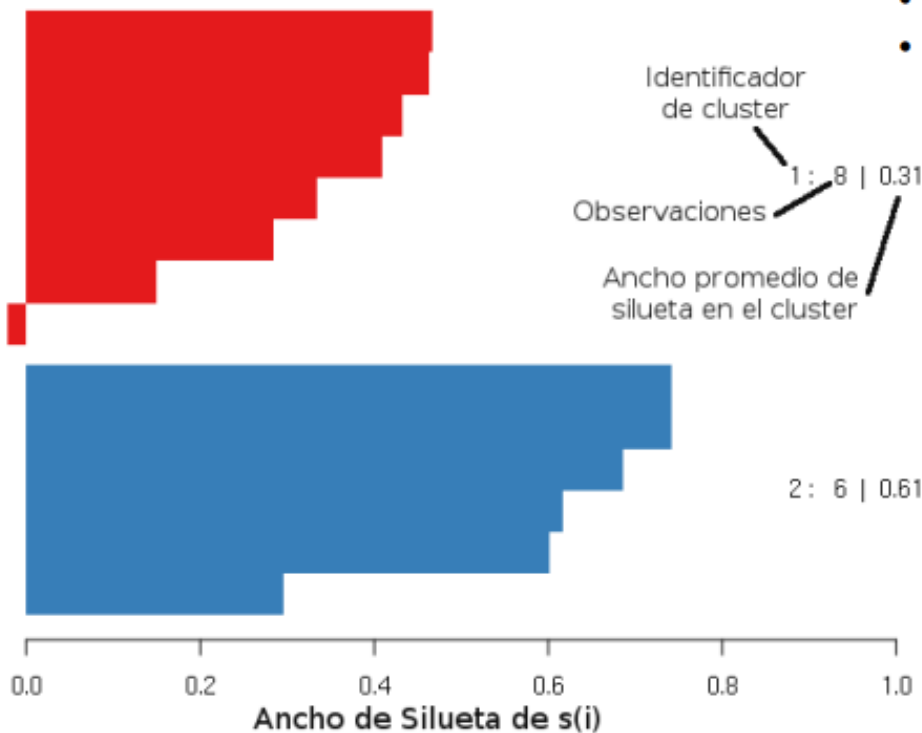
Como definir el numero de clúster? → ej. método codo



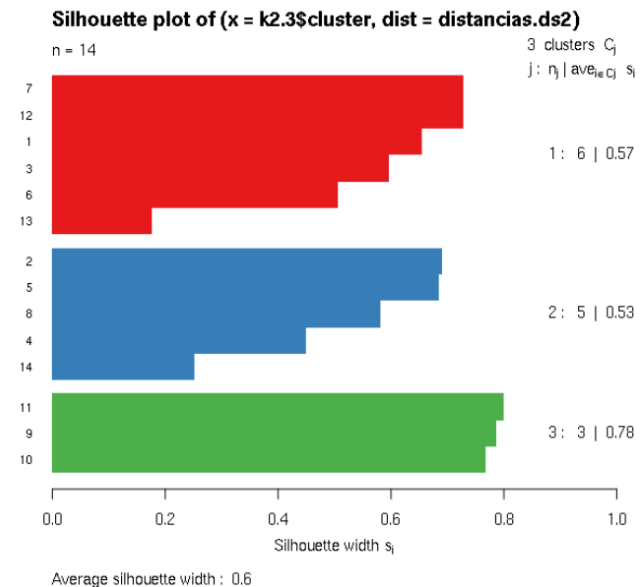
K-means

Como definir el numero de clúster? → silueta

- $s(i) \approx 1$, la observación i está bien asignada a su *cluster*
- $s(i) \approx 0$, la observación i está entre dos *cluster*
- $s(i) \approx -1$, la observación i está mal asignada a su *cluster*



```
> plot(coef.silueta.k2.3, col=brewer.pal(3,"Set1"),  
cex.names=0.7)
```





K-means

- Fortalezas
 - Simple
 - Función objetivo: Optimizar similitud intra-cluster
 - Relativamente eficiente. $O(tkn)$ y paralelizable (batches)
 $t=N$ iteraciones, $k=N$ clústeres, $n=N$ de objetos
- Debilidades
 - Aplicable cuando se puede calcular promedio
 - Suele converger en óptimo local. Inicialización!
 - Necesita especificar N° de clústeres
 - No funciona con datos ruidosos y outliers

Gaussian mixture

- Cada clúster es descrito por una distribución Gaussiana

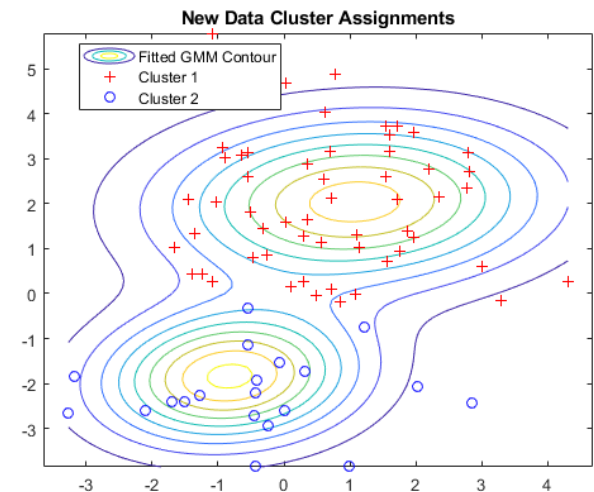
$$P(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\eta)^2}{2\sigma^2}}$$

- Para cada clúster definimos el peso w_i , media μ_i y dispersión σ_i .
- Calculamos la función de Likelihood.
 θ son lo parámetros de nuestro modelo
 i =clusters, j =datos

$$p(x_1 \cdots x_n | \theta) = \prod_{j=1}^n \left(\sum_{i=1}^k p(x_j | C=i) w_i \right)$$

Gaussian mixture

- Definir $K = N^\circ$ clústeres
- Inicializar parámetros de manera aleatoria para la función Gaussiana asociada a cada clúster
- Paso E: Asignar membresía probabilística para todos los objetos j
- Paso M: Re-estimar parámetros w_i , μ_i y σ de las funciones Gaussianas
- Repetir hasta que cambio de parámetros sea menor a threshold



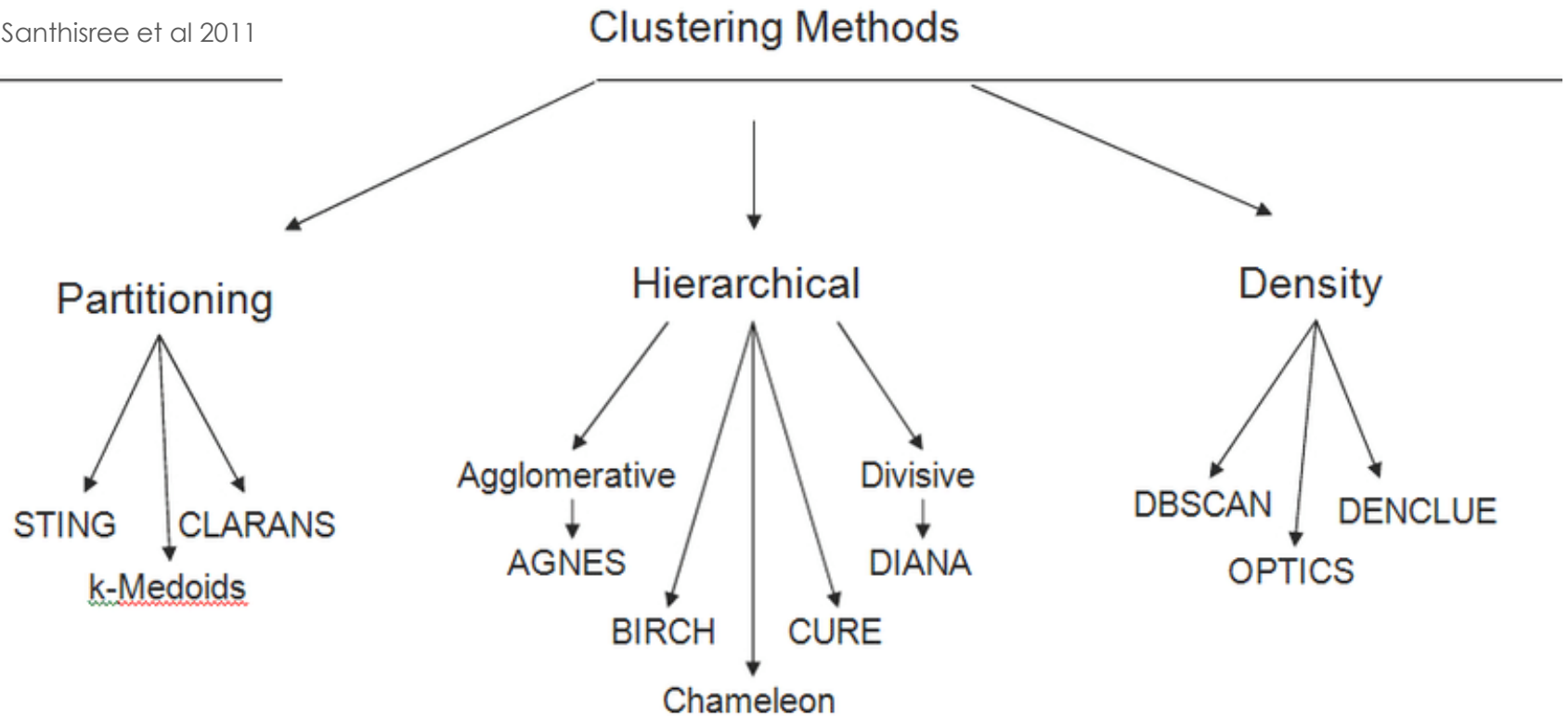


Gaussian mixture

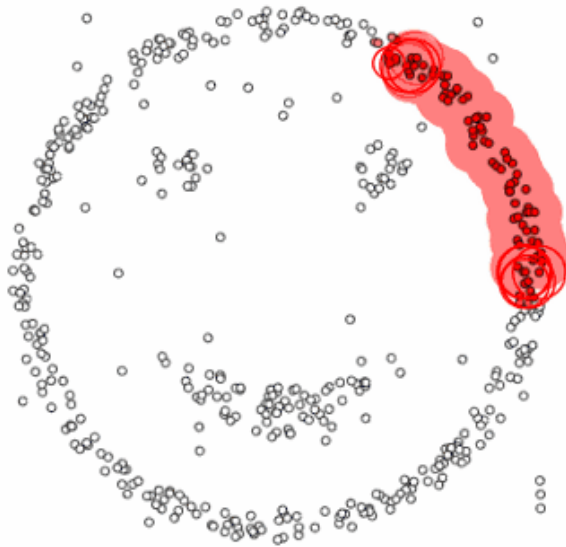
- Fortalezas
 - Interpretabilidad. Aprende modelo generativo a partir de cada clúster
 - Relativamente eficiente. $O(tkn)$
 $t=N$ iteraciones, $k=N$ clústeres, $n=N$ de objetos
 - Función objetivo: Optimizar likelihood
- Debilidades
 - Suele converger en optimo local. Inicialización!
 - Necesita especificar N° de clústeres

Otros Métodos

Santhisree et al 2011



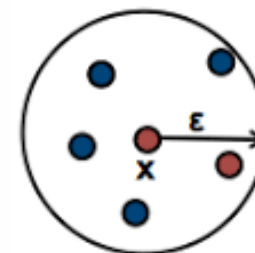
DBSCAN: Density-Based Spatial Clustering of Applications with noise



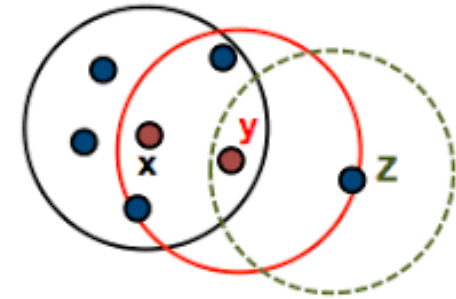
Restart



Pause



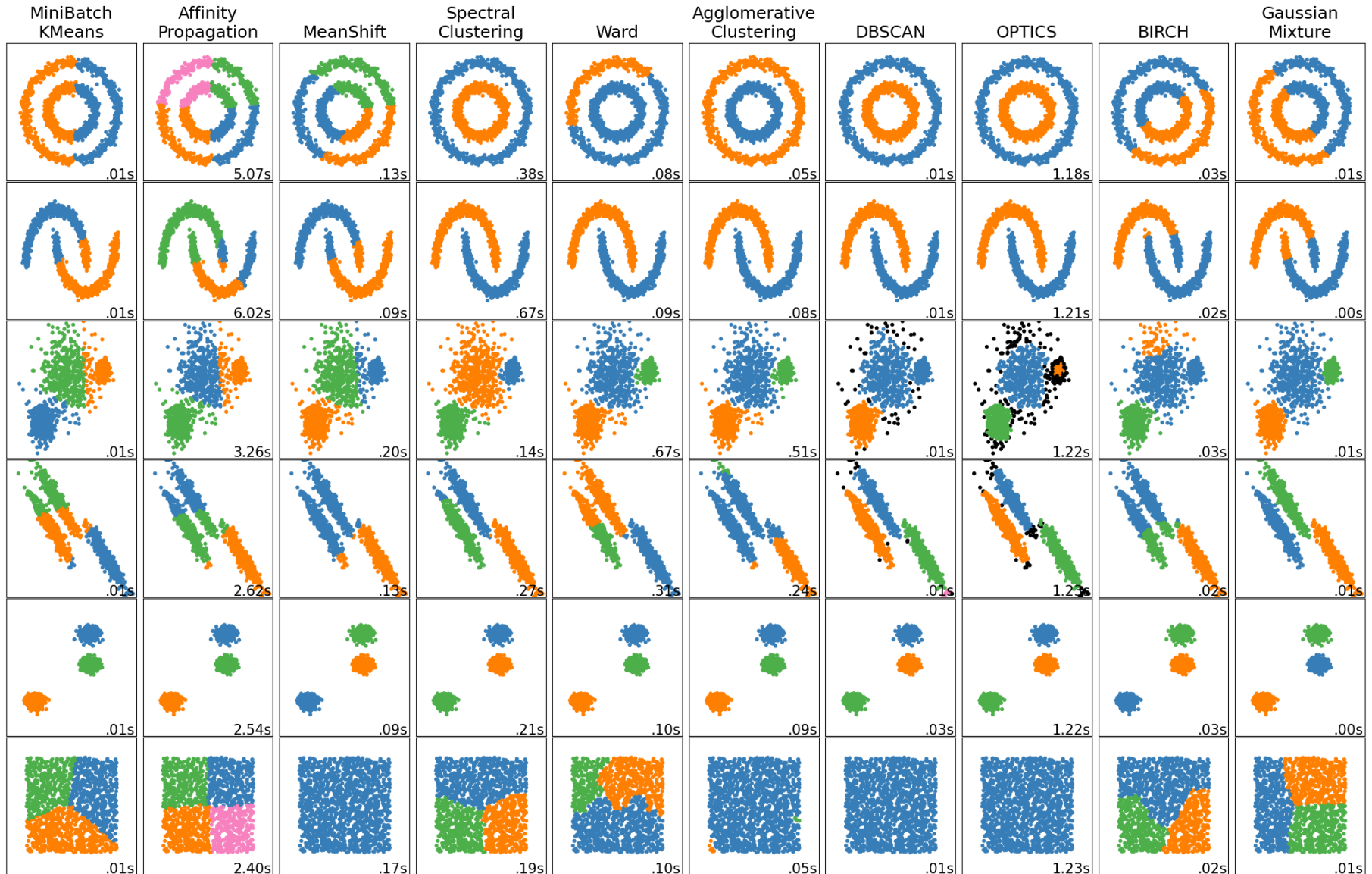
(a)



(b)



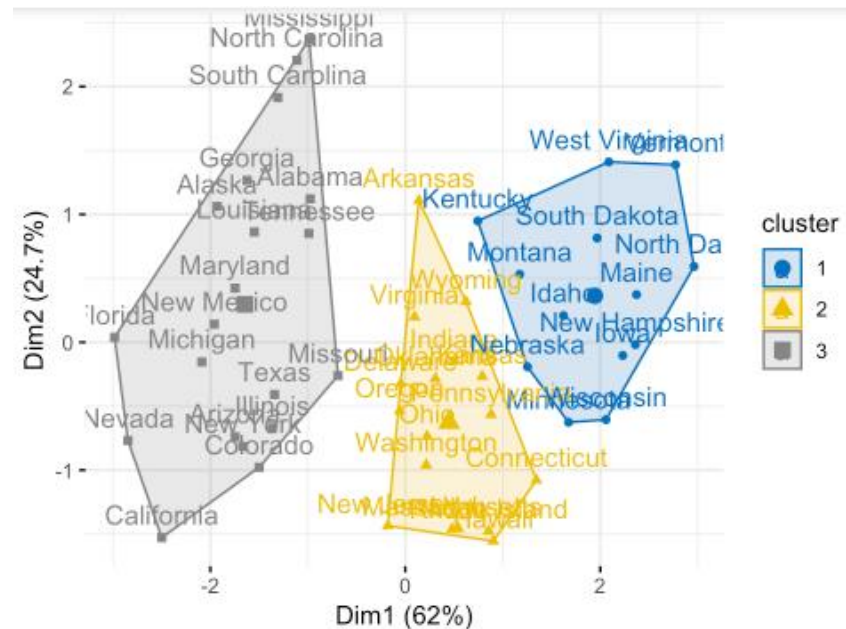
Otros Métodos



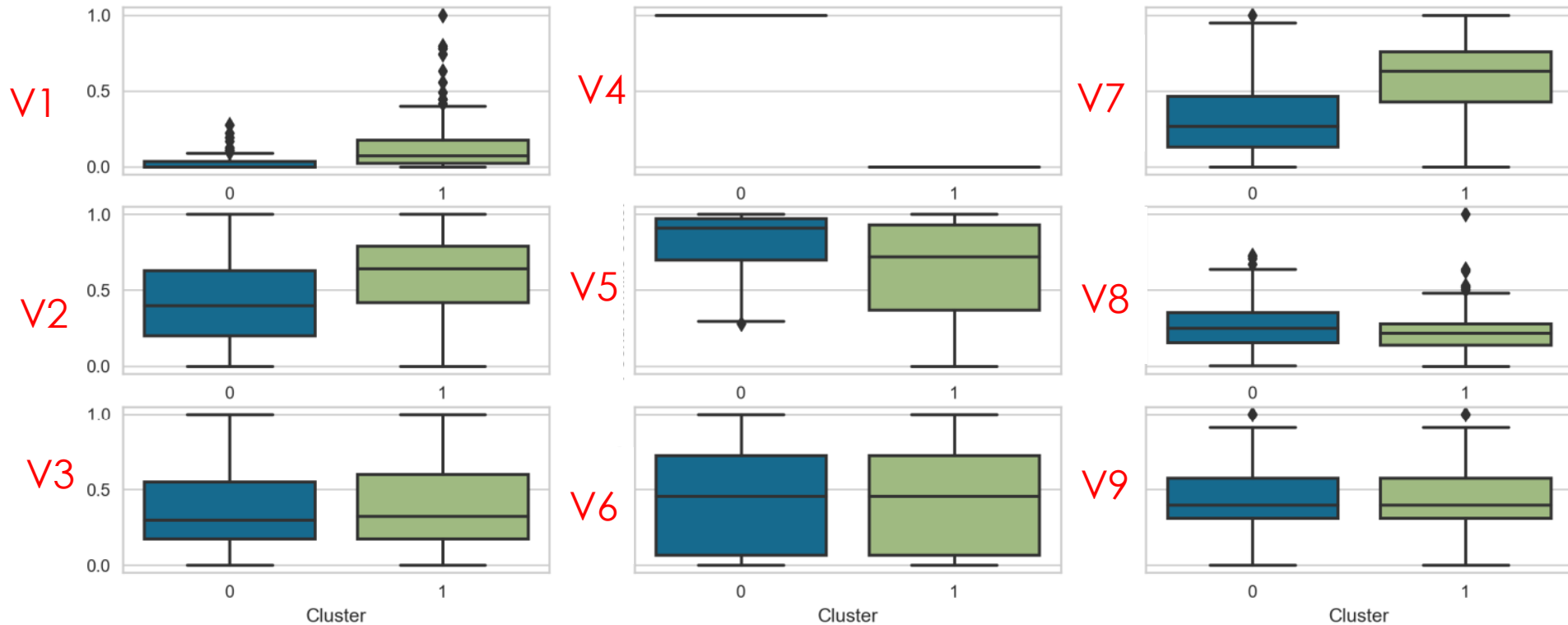
<https://www.datanovia.com/en/blog/types-of-clustering-methods-overview-and-quick-start-r-code/>

Compute and visualize k-means clustering

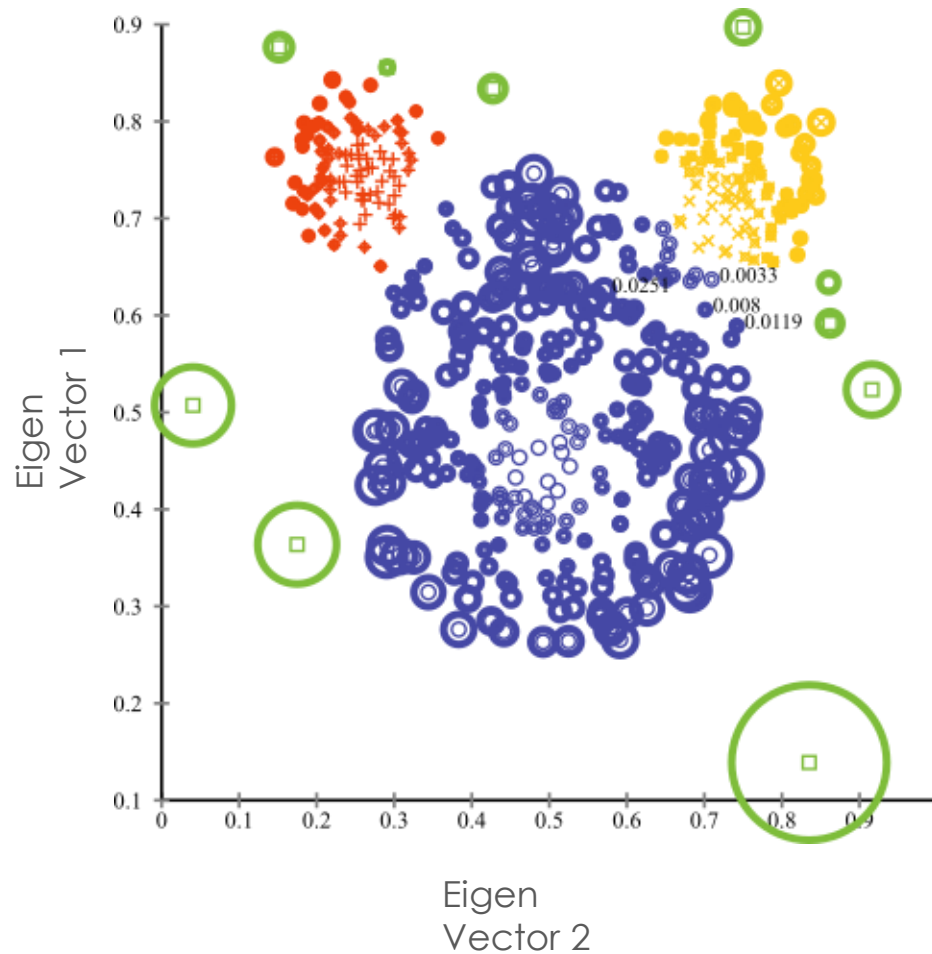
```
set.seed(123)
km.res <- kmeans(my_data, 3, nstart = 25)
# Visualize
library("factoextra")
fviz_cluster(km.res, data = my_data,
  ellipse.type = "convex",
  palette = "jco",
  ggtheme = theme_minimal())
```



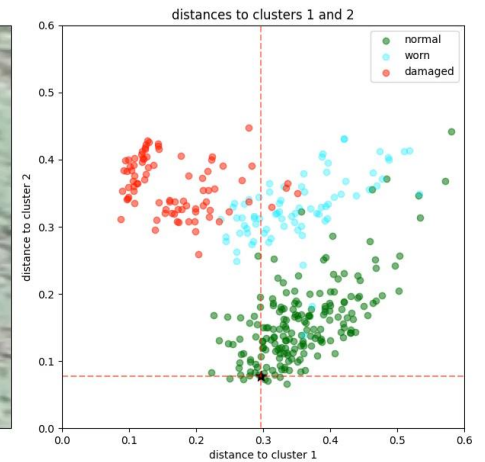
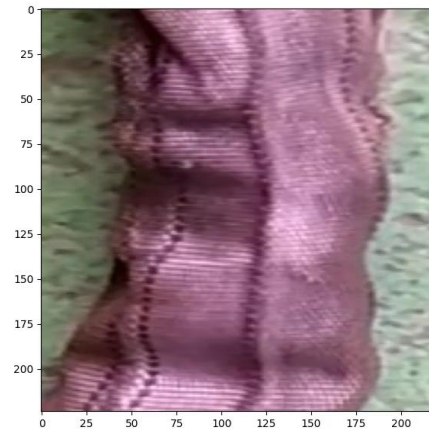
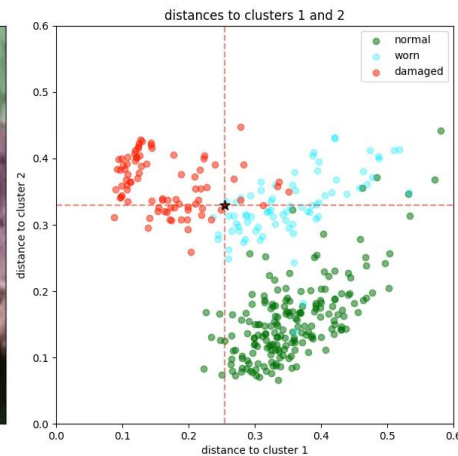
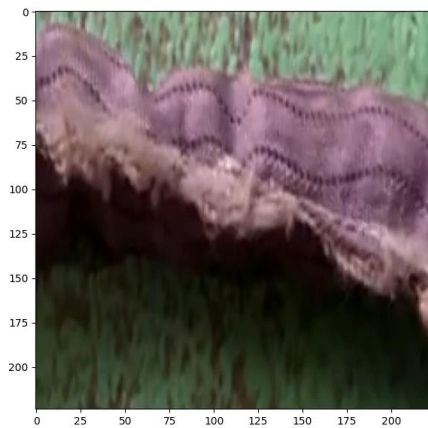
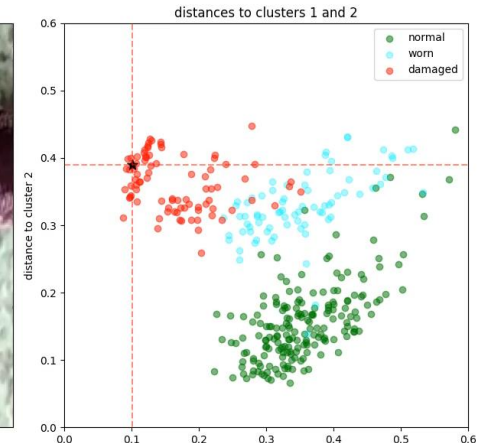
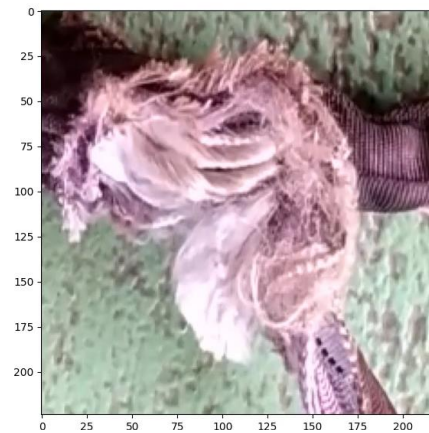
Ejemplo: comparar 2 clúster



Clustering para búsqueda de anomalías



EJEMPLO IDENTIFICACIÓN DE DESGASTE EN CINTAS DE CARGA (EY: CV+DBSCAN)





EJEMPLO CASO DE USO Y CÓDIGOS

<https://www.kaggle.com/datasets/vjchoudhary7/customer-segmentation-tutorial-in-python>

Mall Customer Segmentation Data

Market Basket Analysis

Data Card

Code (954)

Discussion (9)

About Dataset



Clase 07: Segmentación

PROYECTO



Proyecto

- Hacer análisis de segmentación de clientes bancarios usando lenguaje R y métodos de clustering.
- Elegir un objetivo de negocio, segmentar a los clientes y proponer estrategias de marketing para al menos uno de los segmentos.
- Recorrer el espacio de parámetros del algoritmo de clustering y anotar resultados.
- La base de datos contiene ~45.000 registros y campos tales como fecha de nacimiento, actividad y saldo en cuenta.



Evaluación

- Criterios de evaluación

Aspectos a ser Evaluados	Ponderación en la Nota Final
Informe: Estructura, orden y claridad	30%
Informe: Descripción del problema, metodología aplicada, análisis y sus resultados	70%

- Entrega de informe

Los informes(de hasta 15 páginas) deben enviarse hasta las 23:59 horas del lunes 23 de Octubre a regonzar@uc.cl



ESCUELA DE INGENIERÍA
FACULTAD DE INGENIERÍA

EDUCACIÓN
PROFESIONAL

Diplomado en Big Data y Ciencia de Datos
Ciencia de Datos y sus Aplicaciones

Clase 07:

Segmentación

Roberto González



regonzar@uc.cl

