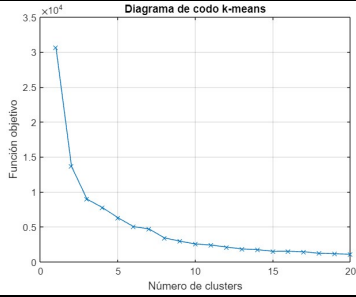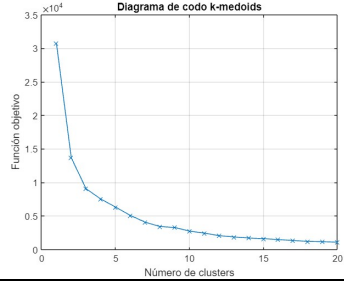# ESCUELA POLITÉCNICA NACIONAL

# REDES NEURONALES

# TAREA 9
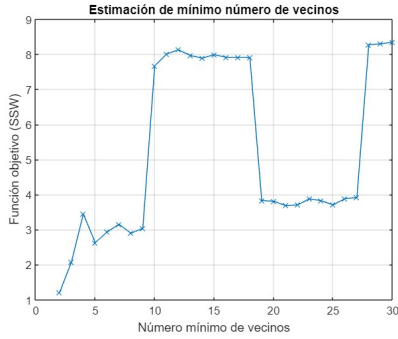# (PARTE 1: Datasets 1 a 5)

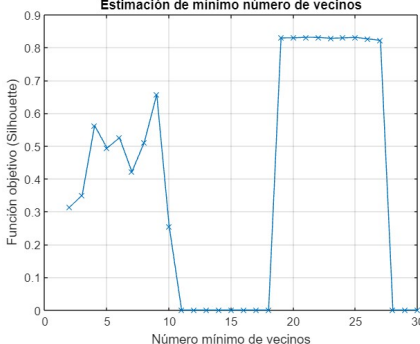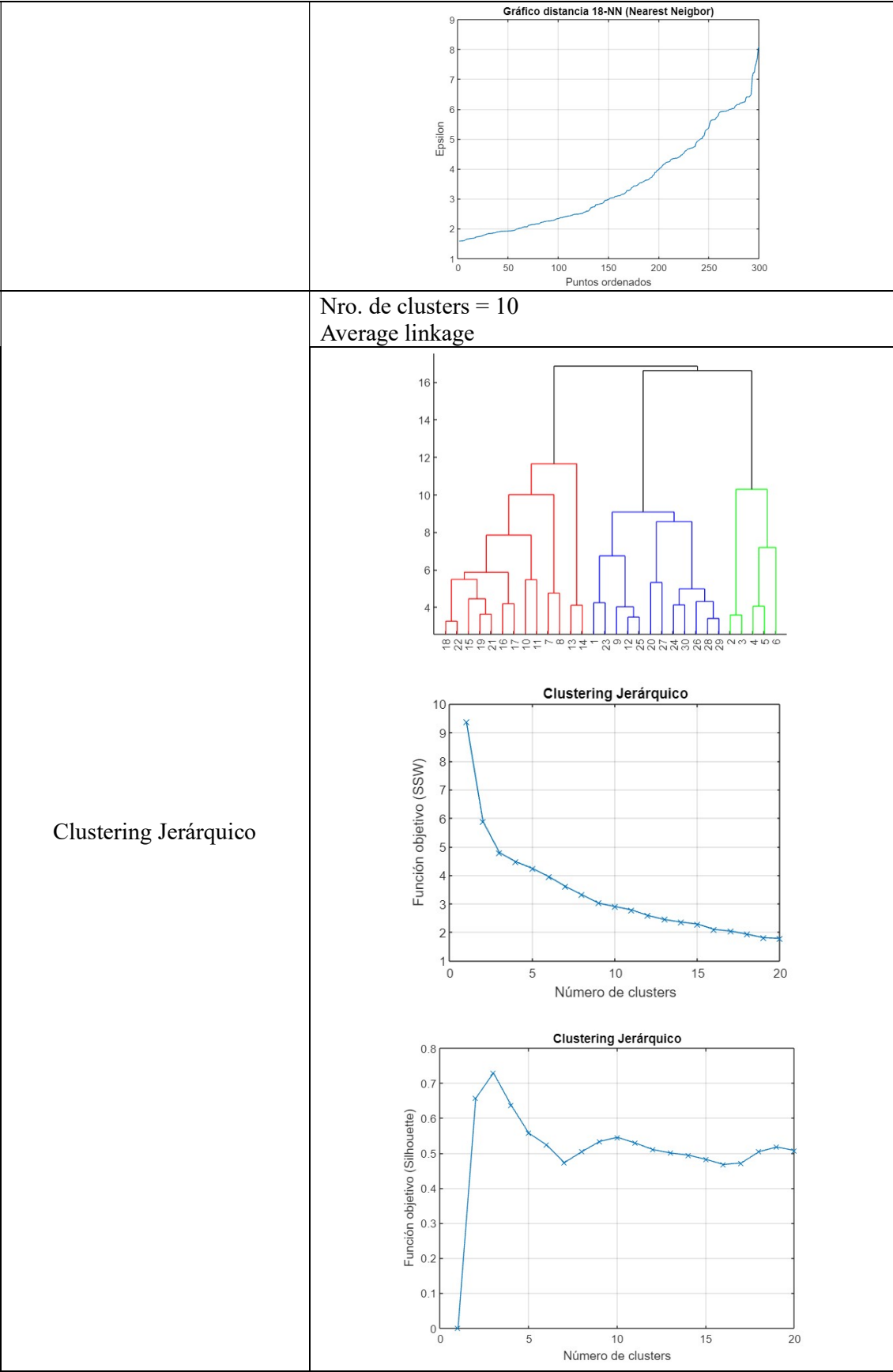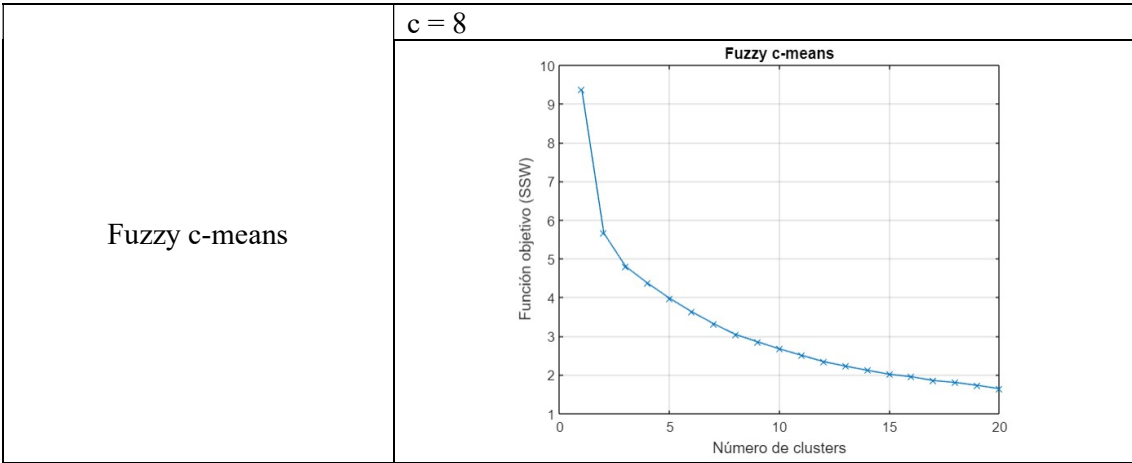

David Fabián Cevallos Salas

2023-08-29

# APRENDIZAJE NO SUPERVISADO

## 1. Dataset Path-Based1
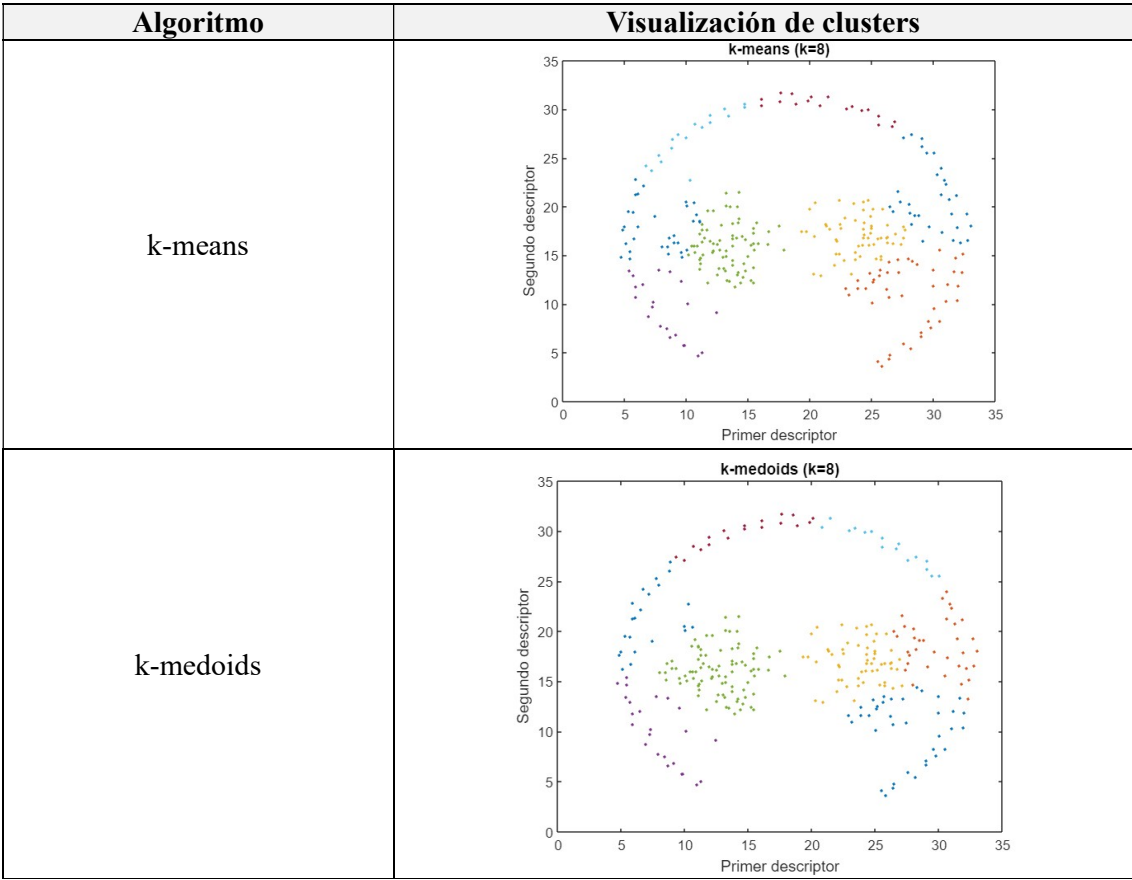
### 1.1 Estimación de parámetros
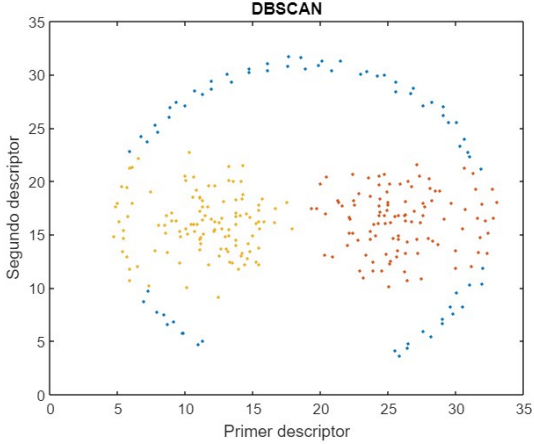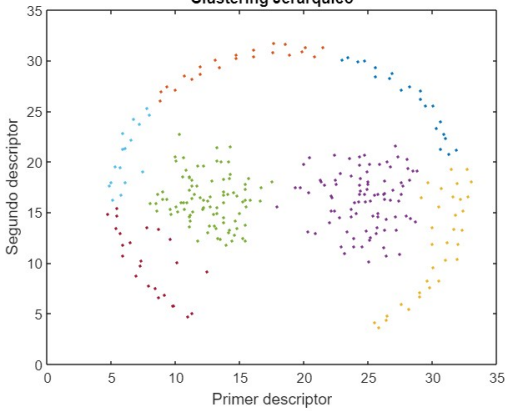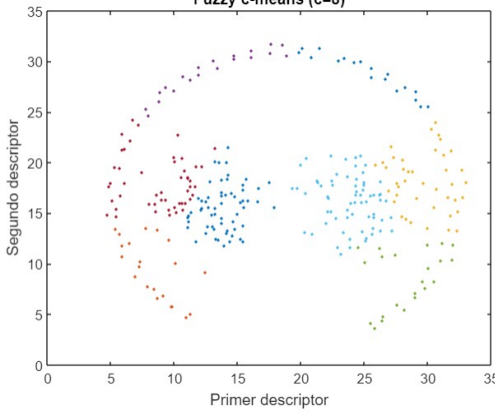
| Algoritmo | Estimación de parámetros |
|---|---|
| k-means | k = 8 <br>  |
| k-medoids | k = 8 <br>  |
| DBSCAN | Mínimo Nro. de vecinos = 19 <br> epsilon = 3,4482 <br>  |

| | |
|---|---|
| | **Gráfico distancia 18-NN (Nearest Neigbor)**<br> |
| | Nro. de clusters = 10<br>Average linkage |
| Clustering Jerárquico | <br><br>**Clustering Jerárquico**<br><br><br>**Clustering Jerárquico**<br> |

| | c = 8 |
|---|---|
| Fuzzy c-means |  |

## 1.2 Visualización de clusters

| Algoritmo | Visualización de clusters |
|---|---|
| k-means |  |
| k-medoids |  |

| | DBSCAN |
|---|---|
| DBSCAN | <br><br>(2 clústeres en total más un cluster de anomalías)<br>(71 observaciones anómalas) |
| Clustering Jerárquico |  |
| Fuzzy c-means |  |

## 1.3 Métricas
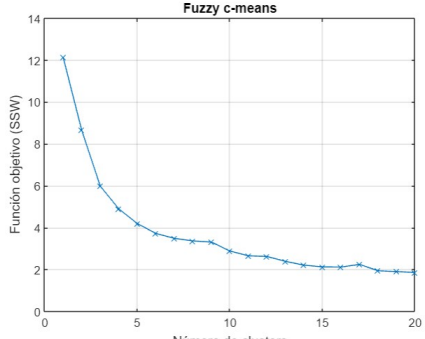
| | k-means | k-medoids | DBSCAN | Clustering Jerárquico | Fuzzy c-means |
|---|---|---|---|---|---|
| **SSW** | 3,0668 | 3,0575 | 3,8258 | 2,8998 | 3,0472 |
| **SSB** | 8,9472 | 8,9552 | 7,3168 | 8,9096 | 8,9276 |
| **WB-index** | 2,7422 | 2,7314 | 1,0458 | 3,2547 | 2,7306 |
| **Silhouette** | 0,5409 | 0,5684 | 0,8296 | 0,5401 | 0,5356 |

# 2. Dataset Aggregation
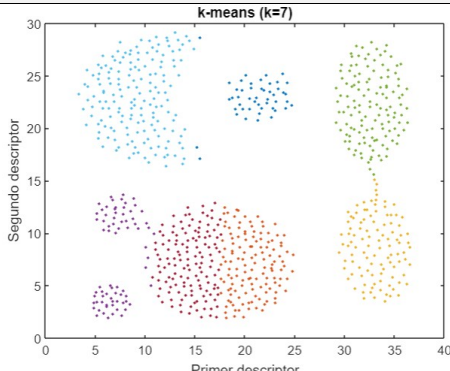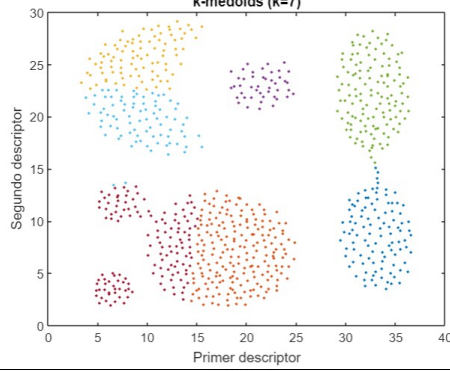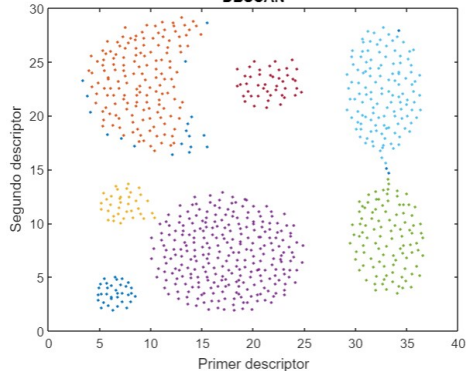
## 2.1 Estimación de parámetros

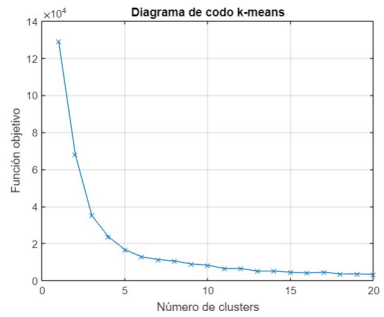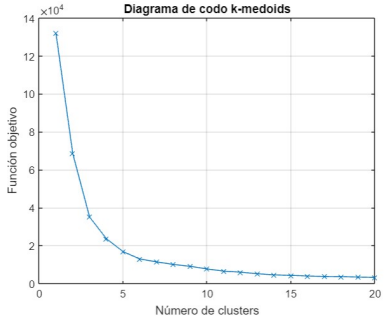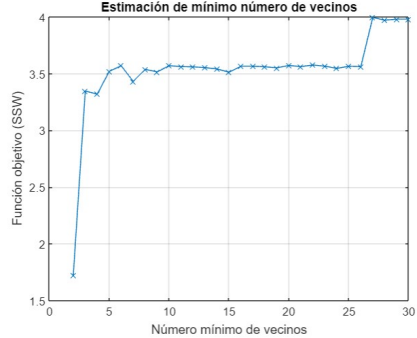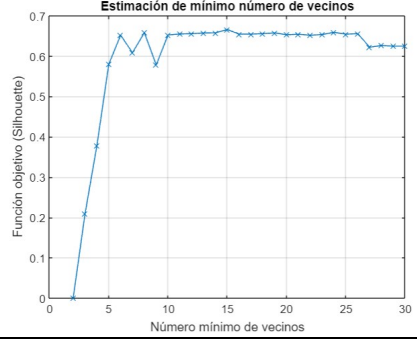| Algoritmo | Estimación de parámetros |
|---|---|
| k-means | k = 7 <br>  |
| k-medoids | k = 7 <br>  |
| DBSCAN | Mínimo Nro. de vecinos = 15 <br> epsilon = 1,8200 <br>  <br>  |

| | |
|---|---|
| | Gráfico distancia 14-NN (Nearest Neigbor) |
| Clustering Jerárquico | Nro. de clusters = 7<br>Average linkage<br><br><br><br>Clustering Jerárquico<br><br>Clustering Jerárquico |

| | c = 7 |
|---|---|
| Fuzzy c-means |  |

## 2.2 Visualización de clusters

| Algoritmo | Visualización de clusters |
|---|---|
| k-means |  |
| k-medoids |  |
| DBSCAN | <br>(7 clústeres en total más un cluster de anomalías)<br>(21 observaciones anómalas) |

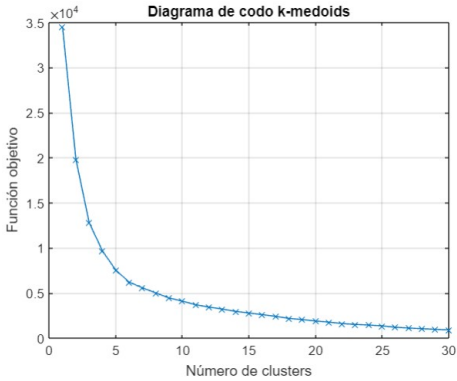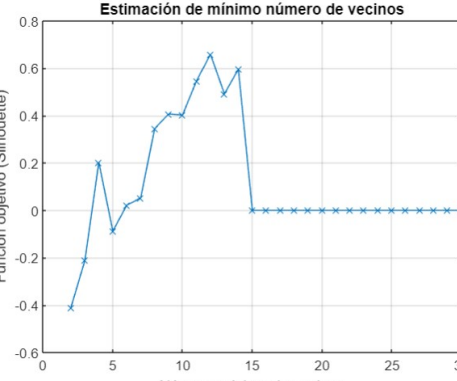| | |
|---|---|
| Clustering Jerárquico |  |
| Fuzzy c-means |  |
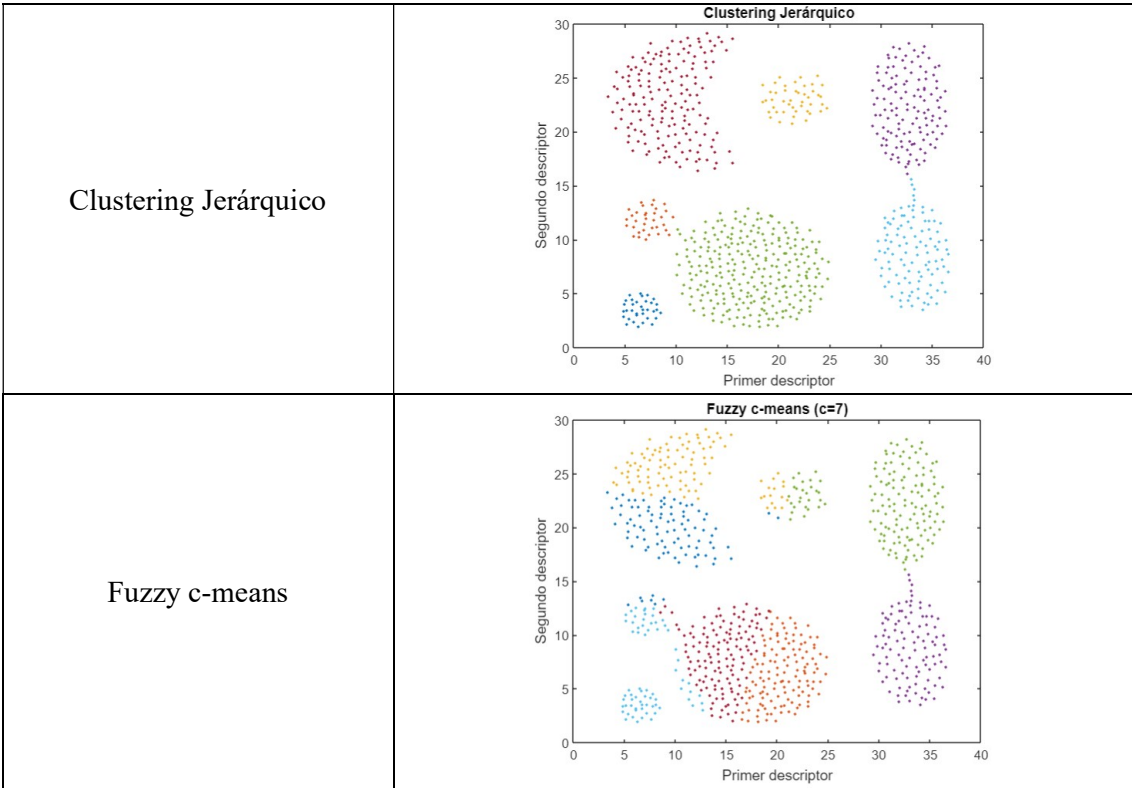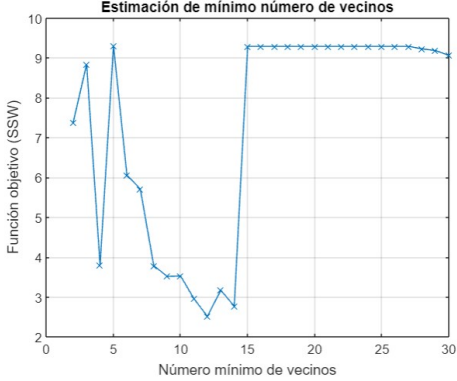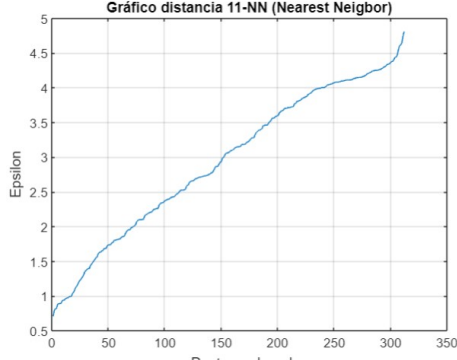
## 2.3 Métricas

| | k-means | k-medoids | DBSCAN | Clustering Jerárquico | Fuzzy c-means |
|---|---|---|---|---|---|
| **SSW** | 3,4442 | 3,4888 | 3,5124 | 3,6068 | 3,7040 |
| **SSB** | 11,8195 | 11,7849 | 11,6671 | 11,6411 | 11,7384 |
| **WB-index** | 2,0398 | 2,0723 | 2,1074 | 2,1688 | 2,2088 |
| **Silhouette** | 0,6613 | 0,6521 | 0,6658 | 0,6439 | 0,5632 |

# 3. Dataset Path-based2
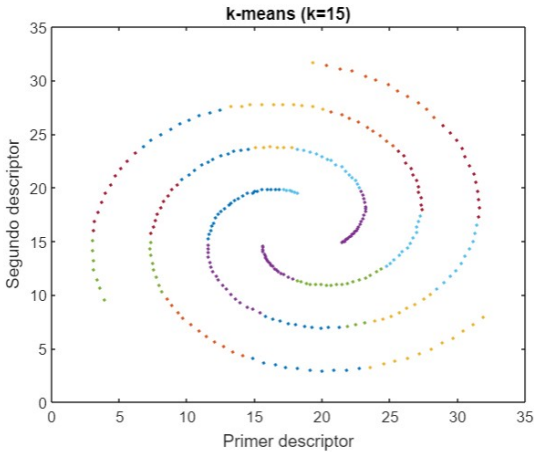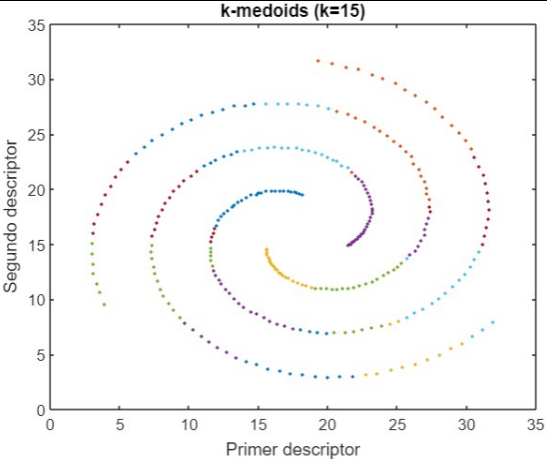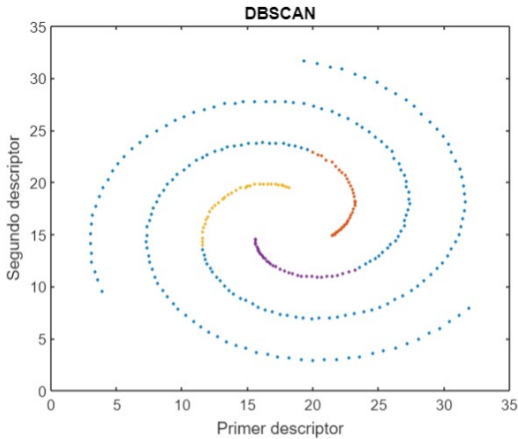
## 3.1 Estimación de parámetros

| Algoritmo | Estimación de parámetros |
|---|---|
| k-means | k = 15 <br>  |

| | |
|---|---|
| | k = 15 |
| k-medoids | **Diagrama de codo k-medoids**<br><br>Función objetivo vs Número de clusters |
| | Mínimo Nro. de vecinos =12<br>epsilon =2,2825 |
| DBSCAN | **Estimación de mínimo número de vecinos**<br>Función objetivo (SSW) vs Número mínimo de vecinos<br><br>**Estimación de mínimo número de vecinos**<br>Función objetivo (Silhouette) vs Número mínimo de vecinos<br><br>**Gráfico distancia 11-NN (Nearest Neigbor)**<br>Epsilon vs Puntos ordenados |

| | |
|---|---|
| Clustering Jerárquico | Nro. de clusters = 15<br>Average linkage<br><br><br><br><br><br> |
| Fuzzy c-means | c = 15<br><br> |

## 3.2 Visualización de clusters

| Algoritmo | Visualización de clusters |
|---|---|
| k-means |  |
| k-medoids |  |
| DBSCAN | <br>(3 clústeres en total más un cluster de anomalías)<br>(102 observaciones anómalas) |

| | |
|---|---|
| Clustering Jerárquico |  |
| Fuzzy c-means |  |
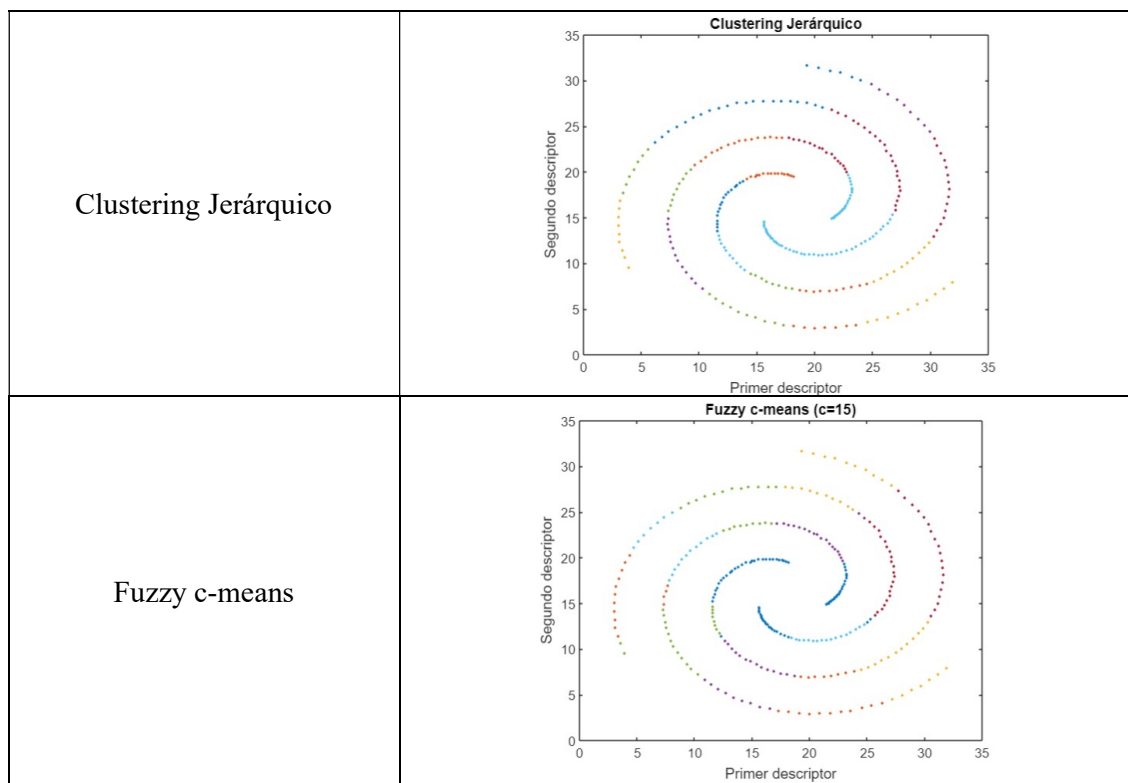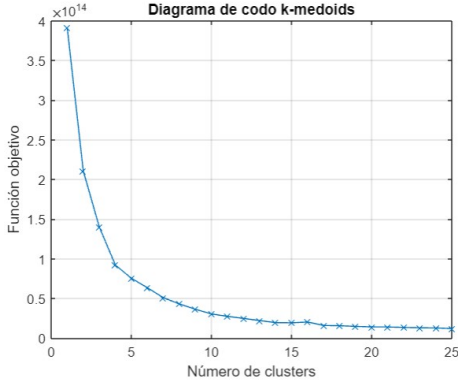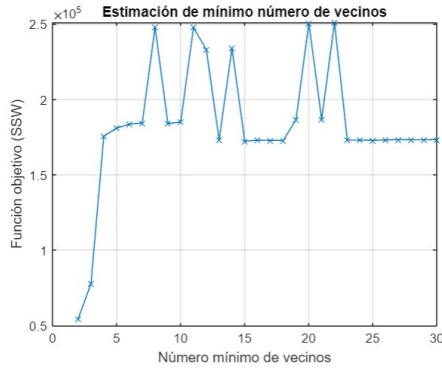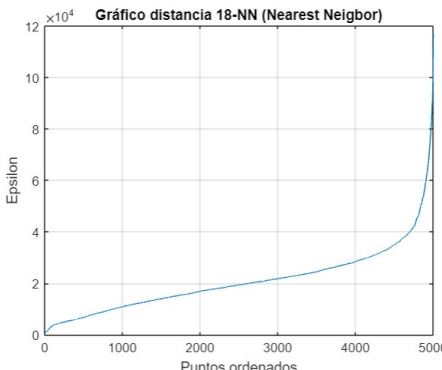
## 3.3 Métricas

| | k-means | k-medoids | DBSCAN | Clustering Jerárquico | Fuzzy c-means |
|---|---|---|---|---|---|
| **SSW** | 2,5454 | 2,6119 | 2,5133 | 2,6600 | 2,5221 |
| **SSB** | 9,0522 | 9,0185 | 4,2708 | 8,9698 | 9,0251 |
| **WB-index** | 4,2179 | 4,3442 | 1,7655 | 4,4482 | 4,1917 |
| **Silhouette** | 0,5150 | 0,4589 | 0,6579 | 0,4907 | 0,4904 |

# 4. Dataset S3

## 4.1 Estimación de parámetros
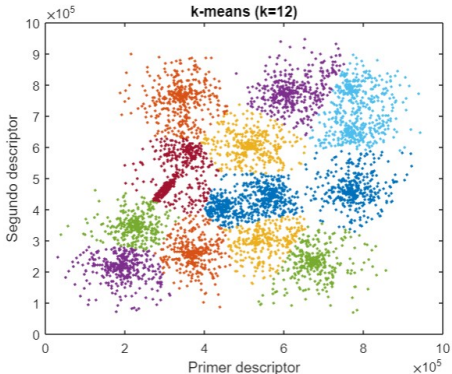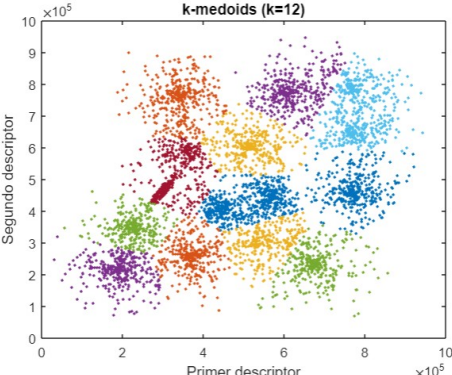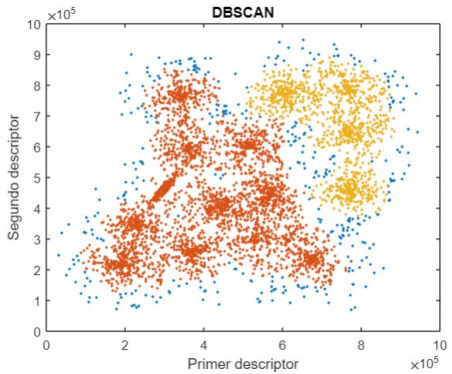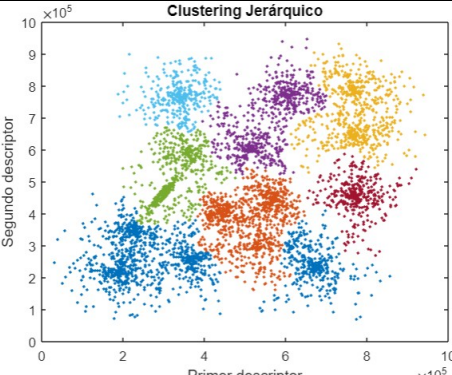
| Algoritmo | Estimación de parámetros |
|---|---|
| k-means | k = 12 <br>  |

| | |
|---|---|
| k-medoids | k = 12<br><br>![Diagrama de codo k-medoids]<br>Diagrama de codo k-medoids<br>(×10^14 Función objetivo vs Número de clusters, 0 a 25) |
| DBSCAN | Mínimo Nro. de vecinos = 19<br>epsilon = 3,0029 e4<br><br>Estimación de mínimo número de vecinos<br>(×10^5 Función objetivo (SSW) vs Número mínimo de vecinos)<br><br>Estimación de mínimo número de vecinos<br>(Función objetivo (Silhouette) vs Número mínimo de vecinos)<br><br>Gráfico distancia 18-NN (Nearest Neigbor)<br>(×10^4 Epsilon vs Puntos ordenados) |
| Clustering Jerárquico | Nro. de clusters = 8<br>Average linkage |

| | |
|---|---|
| |  |
| |  |
| |  |
| Fuzzy c-means | c = 12 <br>  |

## 4.2 Visualización de clusters

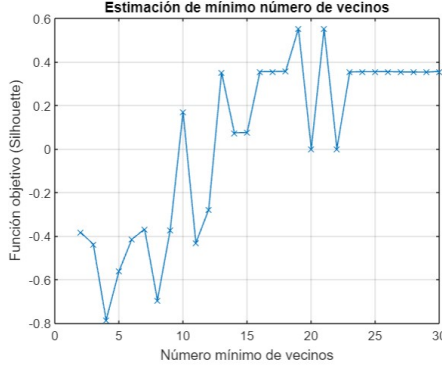| Algoritmo | Visualización de clusters |
|---|---|
| k-means |  |
| k-medoids |  |
| DBSCAN | <br>(2 clústeres en total más un cluster de anomalías)<br>(347 observaciones anómalas) |
| Clustering Jerárquico |  |

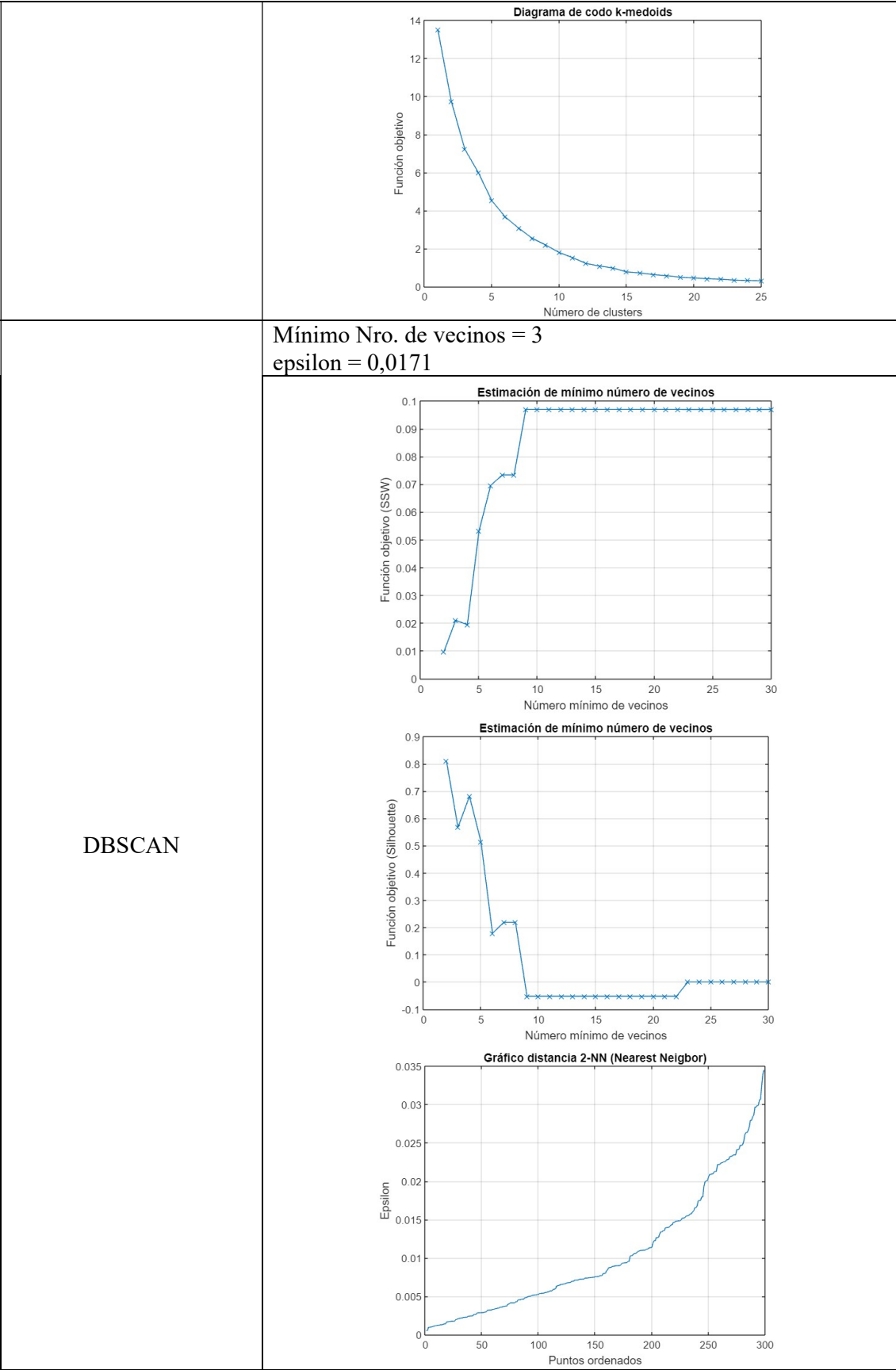| Fuzzy c-means |  |
|---|---|
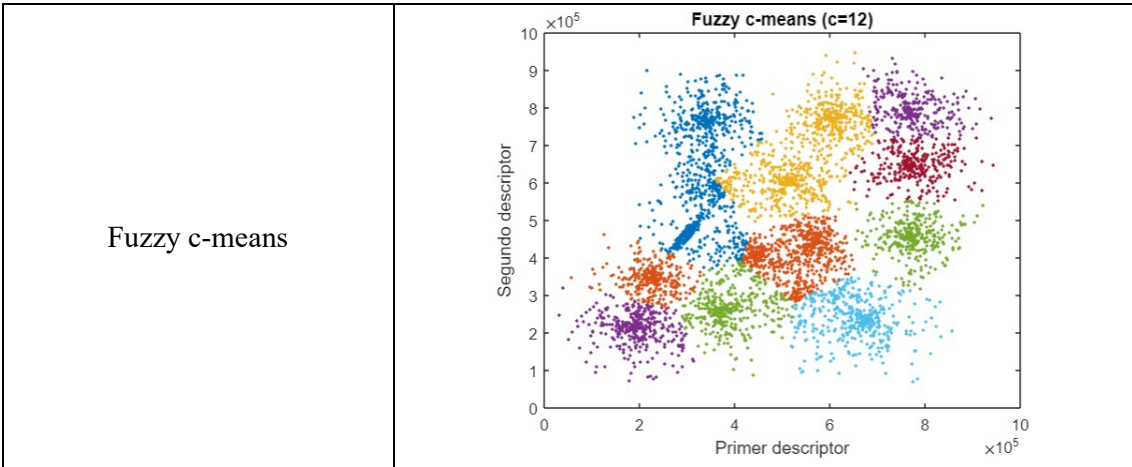
## 4.3 Métricas

|  | **k-means** | **k-medoids** | **DBSCAN** | **Clustering Jerárquico** | **Fuzzy c-means** |
|---|---|---|---|---|---|
| **SSW** | 6,1708 e4 | 6,1714 e4 | 1,8648 e5 | 8,5517 e4 | 6,0830 e4 |
| **SSB** | 2,4934 e5 | 2,4925 e5 | 1,5493 e5 | 2,4655 e5 | 2,4980 e5 |
| **WB-index** | 2,9698 | 2,9712 | 2,4072 | 2,7748 | 2,9222 |
| **Silhouette** | 0,6189 | 0,6179 | 0,5512 | 0,5743 | 0,6266 |

## 5. Dataset Zelnik1

### 5.1 Estimación de parámetros

| Algoritmo | Estimación de parámetros |
|---|---|
| k-means | k = 18  |
| k-medoids | k = 18 |

| | |
|---|---|
| | **Diagrama de codo k-medoids**<br> |
| | Mínimo Nro. de vecinos = 3<br>epsilon = 0,0171 |
| DBSCAN | **Estimación de mínimo número de vecinos**<br><br>**Estimación de mínimo número de vecinos**<br><br>**Gráfico distancia 2-NN (Nearest Neigbor)**<br> |

| | Nro. de clusters = 18<br>Average linkage |
|---|---|
| Clustering Jerárquico | <br><br> |

| | c = 18 |
|---|---|
| Fuzzy c-means |  |

## 5.2 Visualización de clusters

| Algoritmo | Visualización de clusters |
|---|---|
| k-means |  |
| k-medoids |  |

| | DBSCAN |
|---|---|
| DBSCAN | <br>(16 clústeres en total más un cluster de anomalías)<br>(91 observaciones anómalas) |
| Clustering Jerárquico |  |
| Fuzzy c-means |  |

## 5.3 Métricas

| | k-means | k-medoids | DBSCAN | Clustering Jerárquico | Fuzzy c-means |
|---|---|---|---|---|---|
| **SSW** | 0,0354 | 0,0356 | 0,0210 | 0,0392 | 0,0554 |
| **SSB** | 0,1686 | 0,1682 | 0,1034 | 0,1666 | 0,1695 |
| **WB-index** | 3,7855 | 3,8055 | 3,2480 | 4,2383 | 5,8885 |
| **Silhouette** | 0,7239 | 0,7216 | 0,5670 | 0,6621 | 0,3486 |