

ESCUELA POLITÉCNICA NACIONAL

REDES NEURONALES

TAREA 9

(PARTE 1: Datasets 1 a 5)



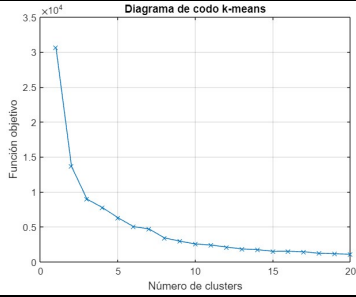
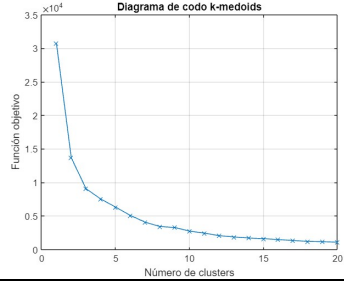
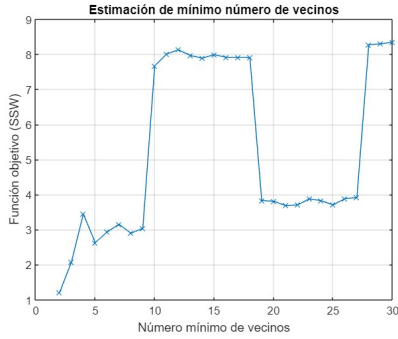
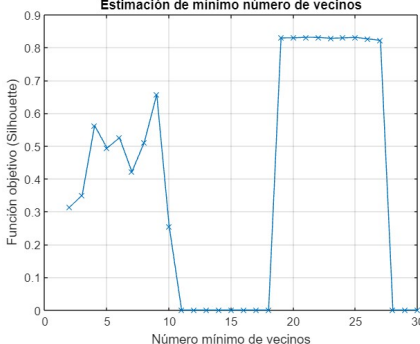
David Fabián Cevallos Salas

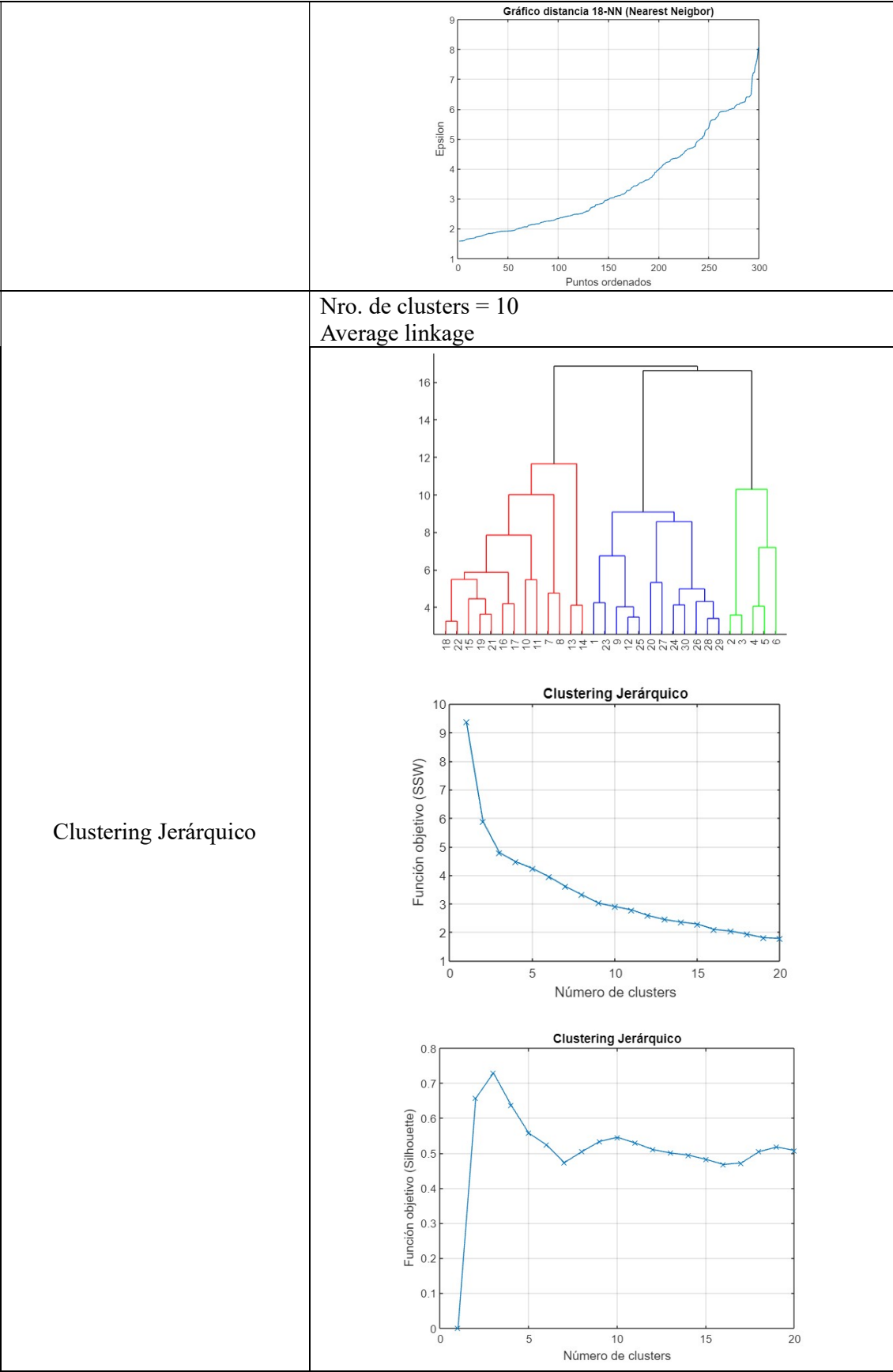
2023-08-29

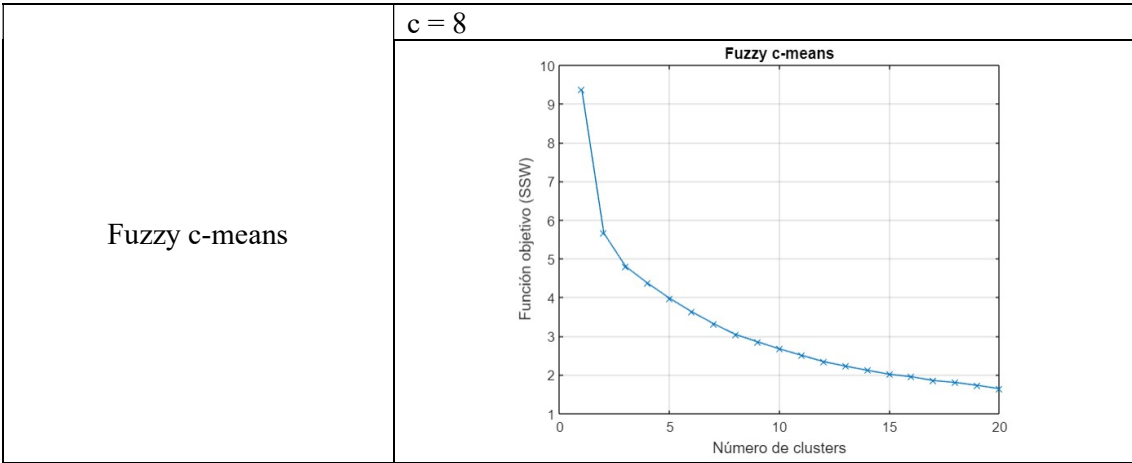
APRENDIZAJE NO SUPERVISADO

1. Dataset Path-Based1

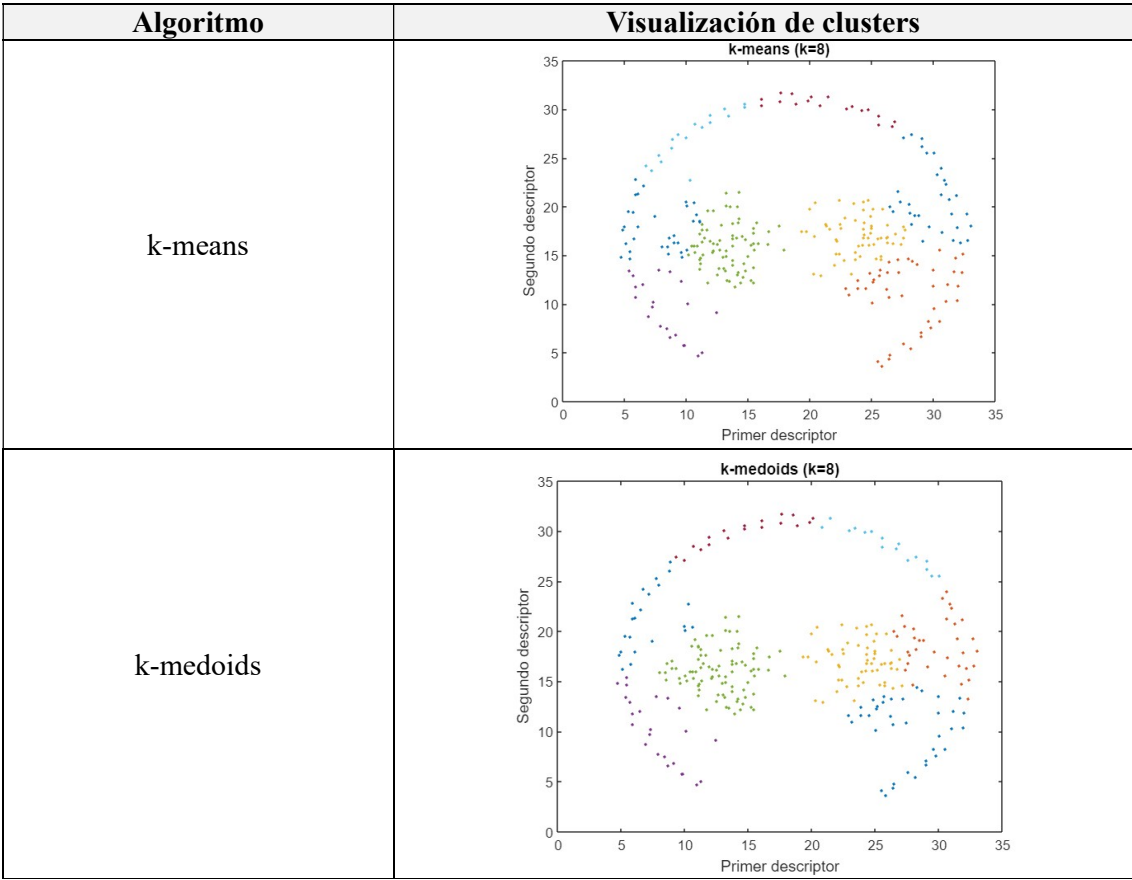
1.1 Estimación de parámetros

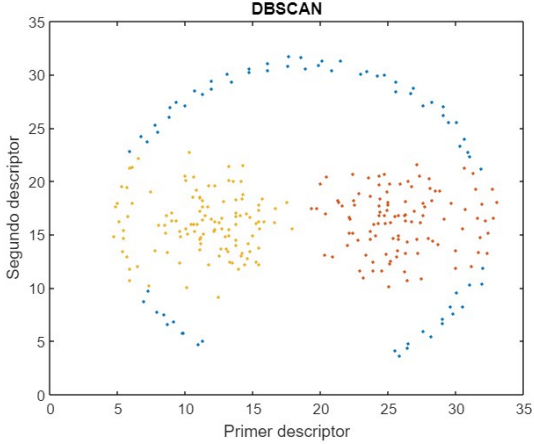
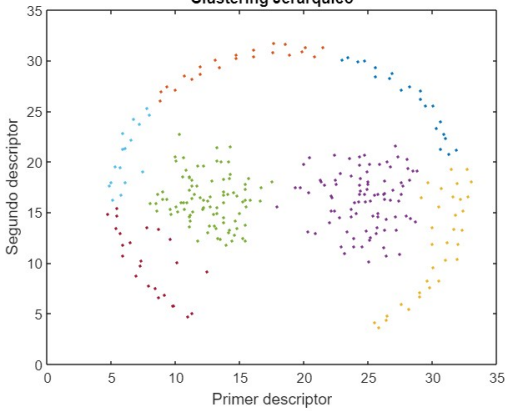
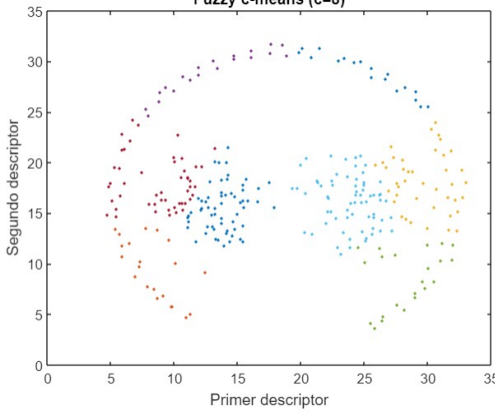
Algoritmo	Estimación de parámetros
k-means	$k = 8$
	 <p>Diagrama de codo k-means</p>
k-medoids	$k = 8$
	 <p>Diagrama de codo k-medoids</p>
DBSCAN	Mínimo Nro. de vecinos = 19 epsilon = 3,4482
	 <p>Estimación de mínimo número de vecinos</p>
	 <p>Estimación de mínimo número de vecinos</p>





1.2 Visualización de clusters



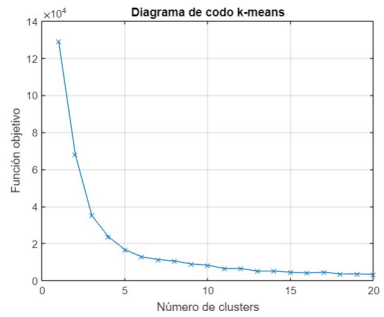
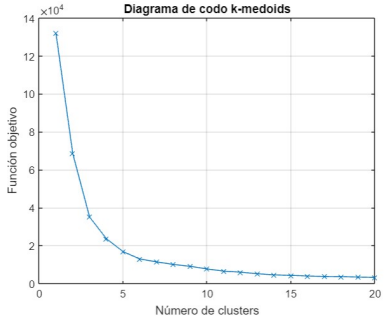
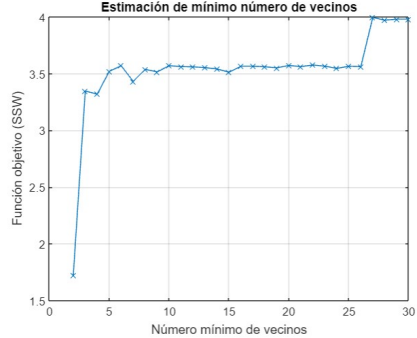
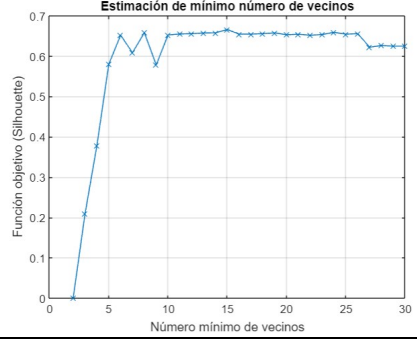
DBSCAN	 <p>(2 clústeres en total más un cluster de anomalías) (71 observaciones anómalas)</p>
Clustering Jerárquico	
Fuzzy c-means	

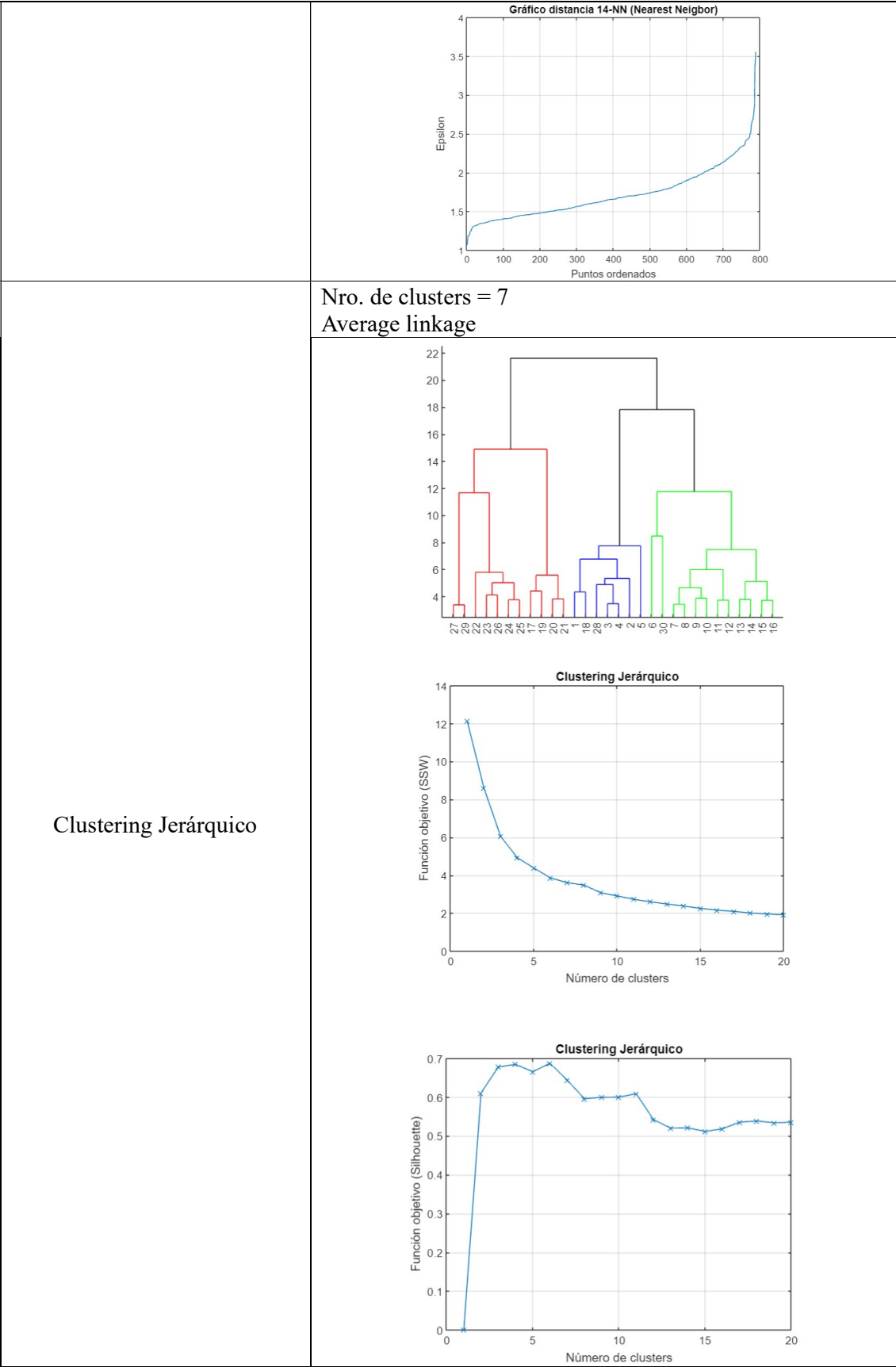
1.3 Métricas

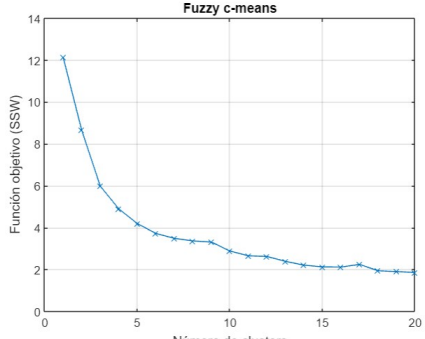
	k-means	k-medoids	DBSCAN	Clustering Jerárquico	Fuzzy c-means
SSW	3,0668	3,0575	3,8258	2,8998	3,0472
SSB	8,9472	8,9552	7,3168	8,9096	8,9276
WB-index	2,7422	2,7314	1,0458	3,2547	2,7306
Silhouette	0,5409	0,5684	0,8296	0,5401	0,5356

2. Dataset Aggregation

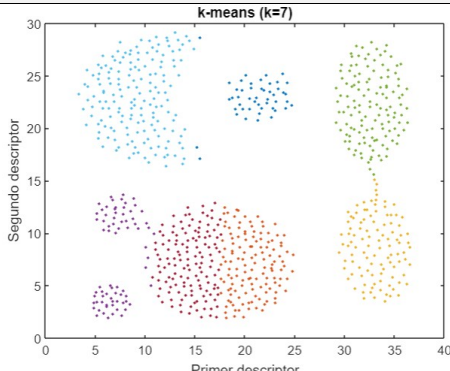
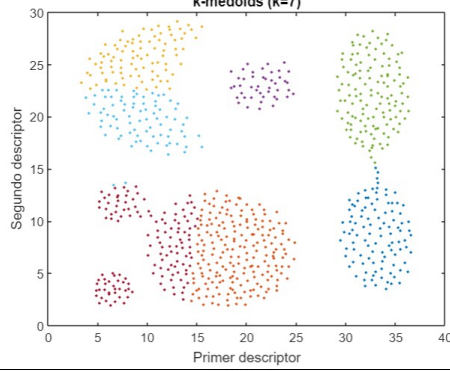
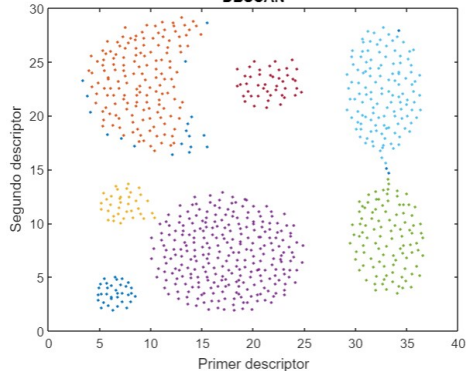
2.1 Estimación de parámetros

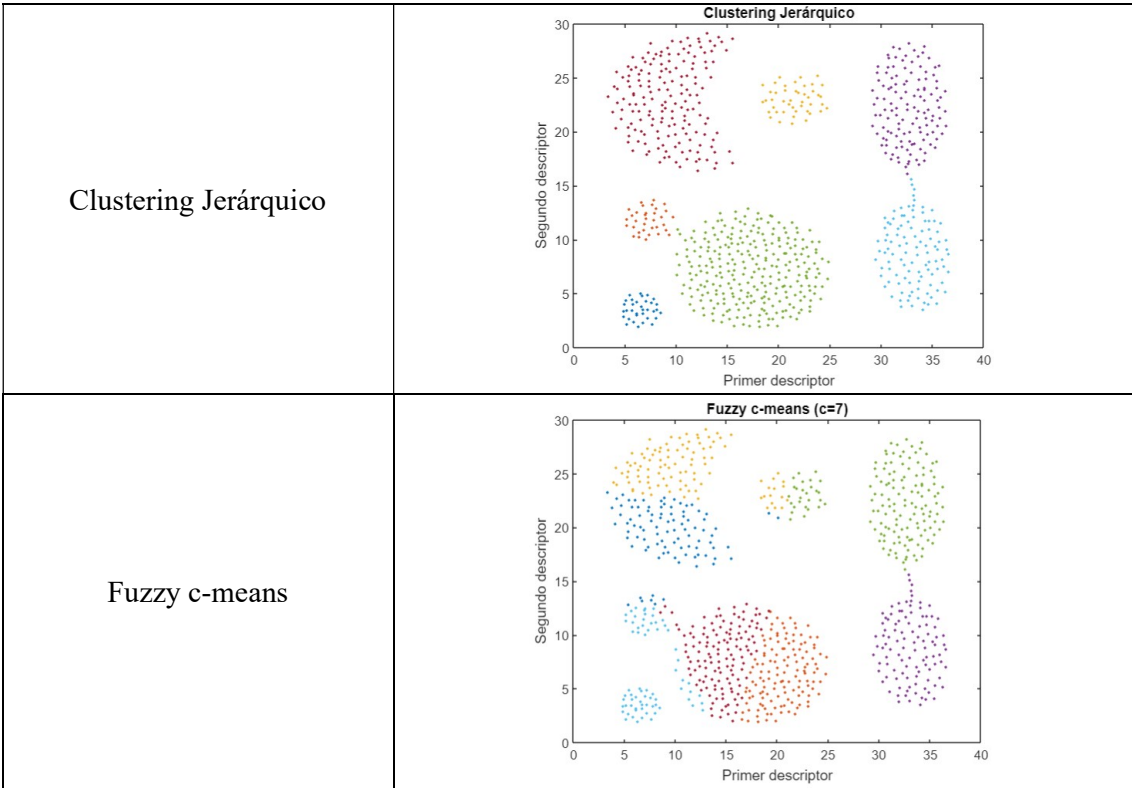
Algoritmo	Estimación de parámetros
k-means	<p>$k = 7$</p> 
k-medoids	<p>$k = 7$</p> 
DBSCAN	<p>Mínimo Nro. de vecinos = 15 epsilon = 1,8200</p>  



Fuzzy c-means	$c = 7$
	

2.2 Visualización de clusters

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

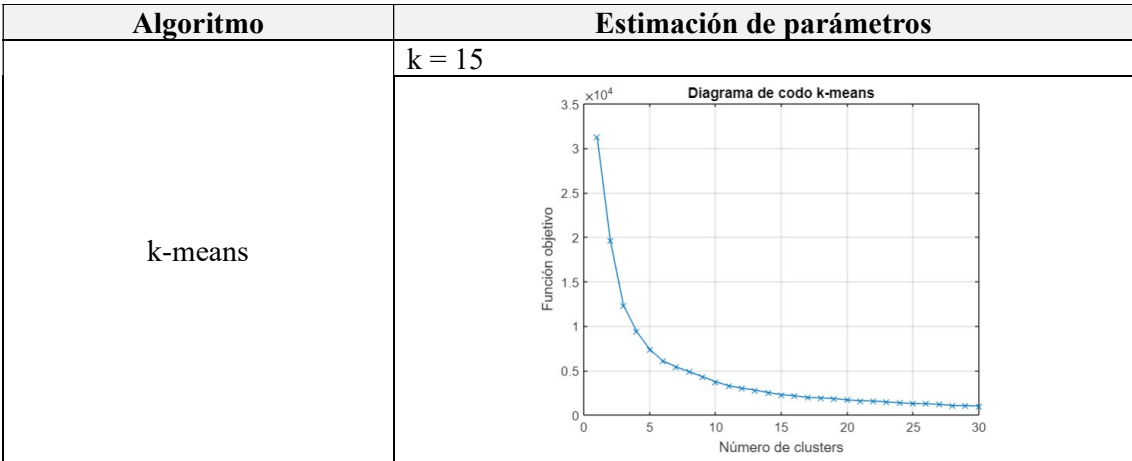


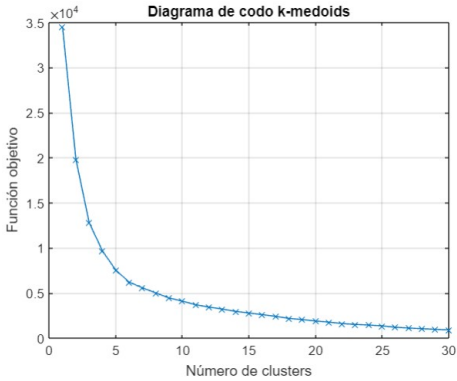
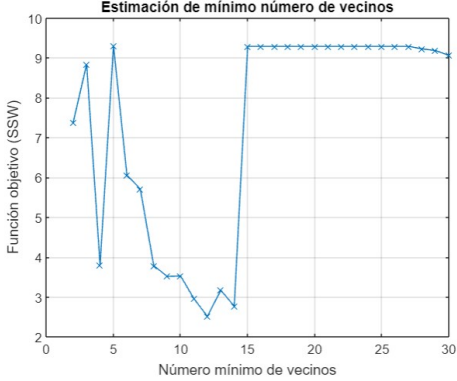
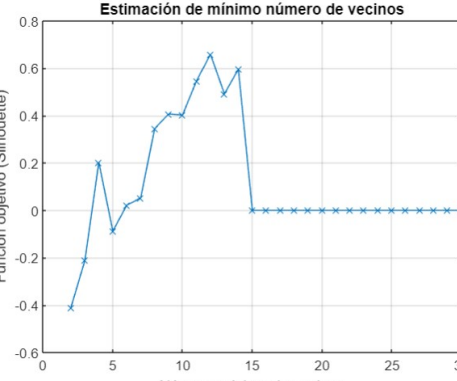
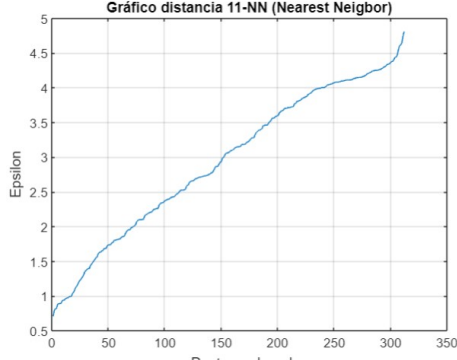
2.3 Métricas

	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

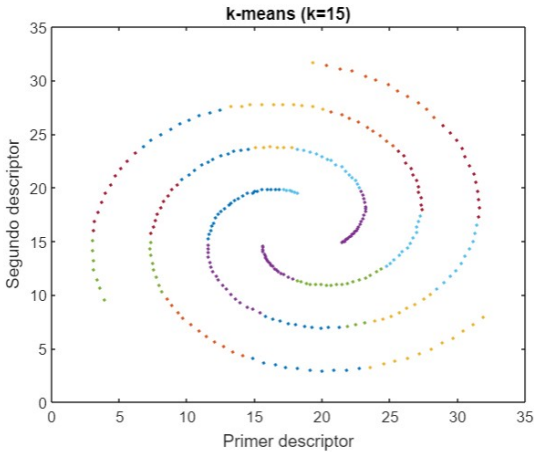
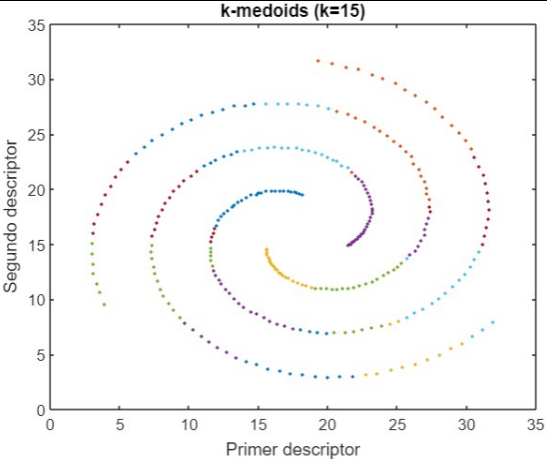
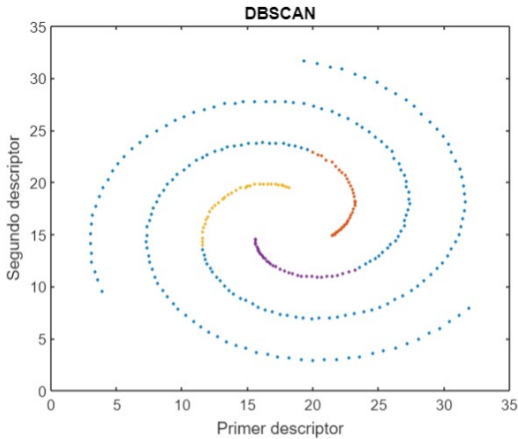
3.1 Estimación de parámetros

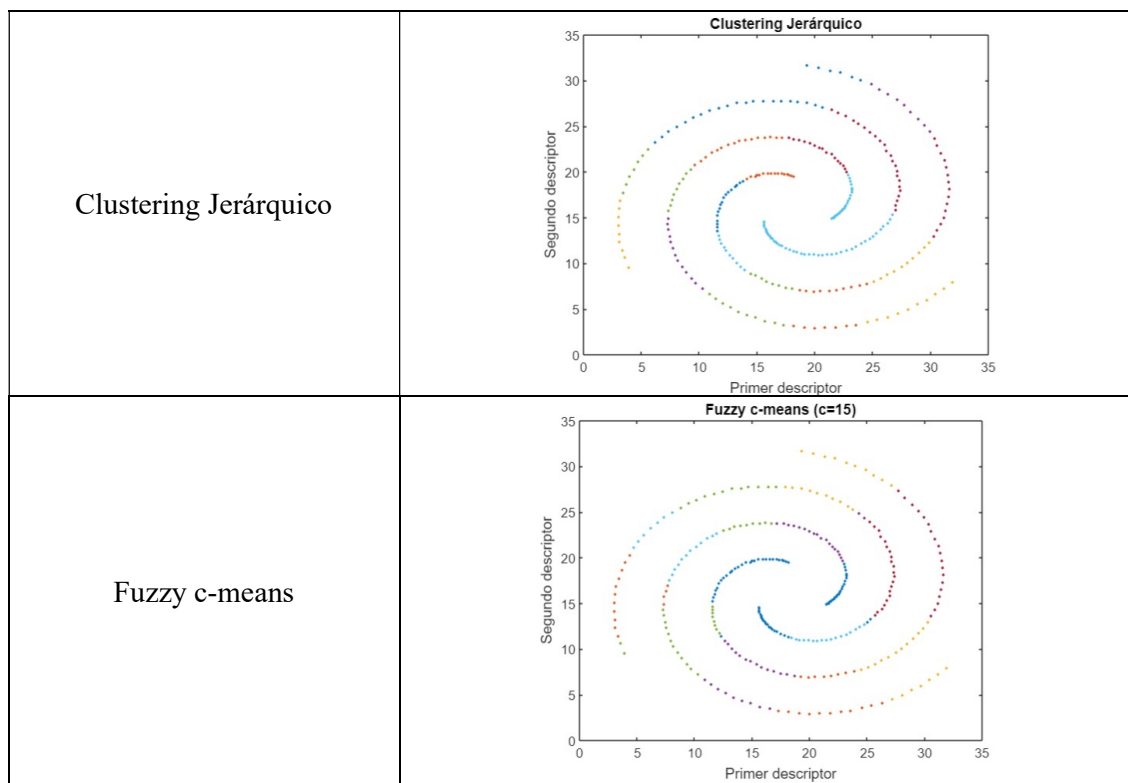


<p>k-medoids</p>	<p>k = 15</p>  <p>Diagrama de codo k-medoids</p> <p>Este gráfico muestra la función objetivo (Y-axis, multiplicada por 10^4) en función del número de clusters (X-axis, de 0 a 30). La curva comienza en un valor alto (aproximadamente 3.5 x 10^4) y disminuye rápidamente, mostrando un cambio de pendiente significativo alrededor de 15 clusters, lo que sugiere que este número podría ser el óptimo.</p>
<p>DBSCAN</p>	<p>Mínimo Nro. de vecinos =12 epsilon =2,2825</p>  <p>Estimación de mínimo número de vecinos</p> <p>Este gráfico muestra la función objetivo (SSW) (Y-axis, de 2 a 10) en función del número mínimo de vecinos (X-axis, de 0 a 30). La curva muestra una fluctuación inicial, pero a partir de 15 vecinos, el valor se estabiliza en un nivel alto (alrededor de 9.5), lo que indica que este número podría ser el óptimo.</p>  <p>Estimación de mínimo número de vecinos</p> <p>Este gráfico muestra la función objetivo (Silhouette) (Y-axis, de -0.6 a 0.8) en función del número mínimo de vecinos (X-axis, de 0 a 30). La curva muestra una fluctuación inicial, pero a partir de 15 vecinos, el valor se estabiliza en un nivel alto (alrededor de 0.6), lo que indica que este número podría ser el óptimo.</p>  <p>Gráfico distancia 11-NN (Nearest Neighbor)</p> <p>Este gráfico muestra la distancia (Y-axis, de 0.5 a 5) en función del número de puntos ordenados (X-axis, de 0 a 350). La curva muestra una tendencia general ascendente, lo que sugiere que el número de puntos ordenados podría ser el óptimo.</p>

Clustering Jerárquico	<p>Nro. de clusters = 15 Average linkage</p>
	<div data-bbox="787 262 1218 598"></div> <div data-bbox="787 640 1218 1018"><p>Clustering Jerárquico</p></div> <div data-bbox="787 1050 1218 1417"><p>Clustering Jerárquico</p></div>
Fuzzy c-means	<p>c = 15</p> <div data-bbox="787 1459 1218 1829"><p>Fuzzy c-means</p></div>

3.2 Visualización de clusters

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

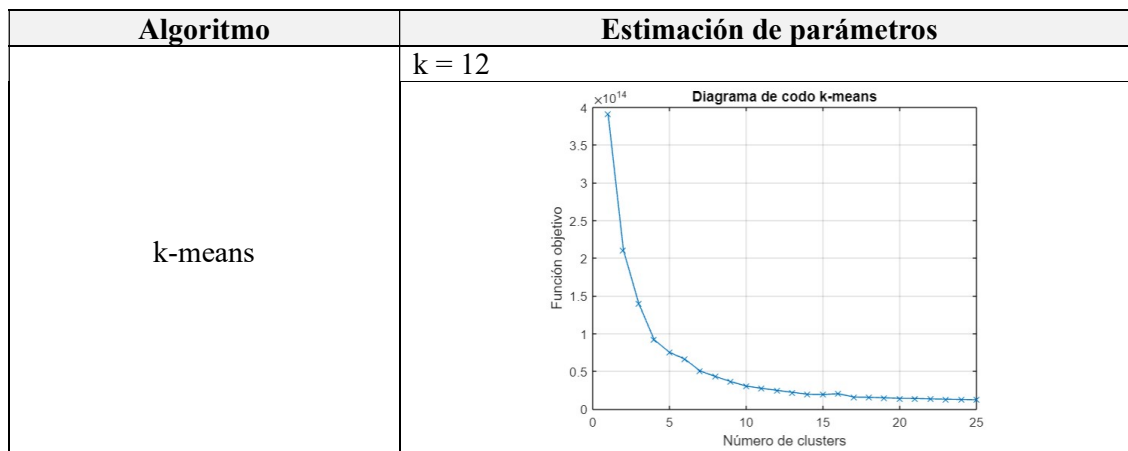


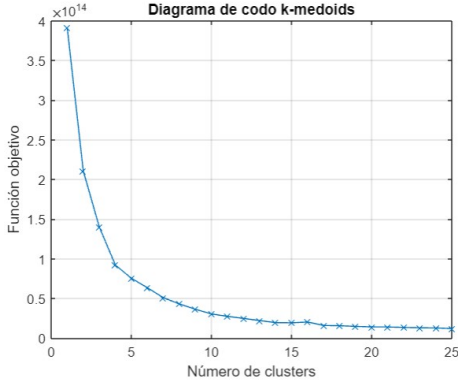
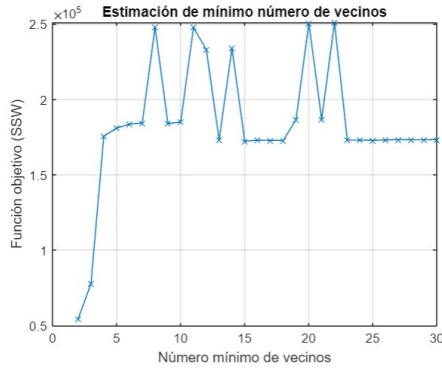
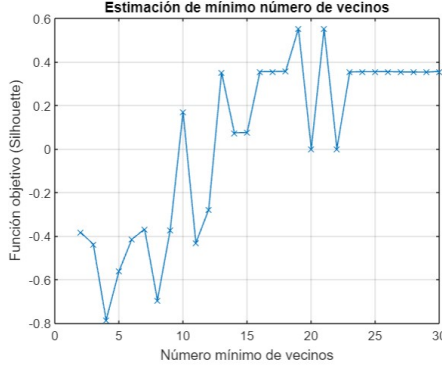
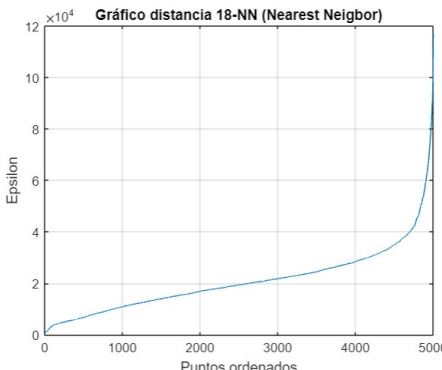
3.3 Métricas

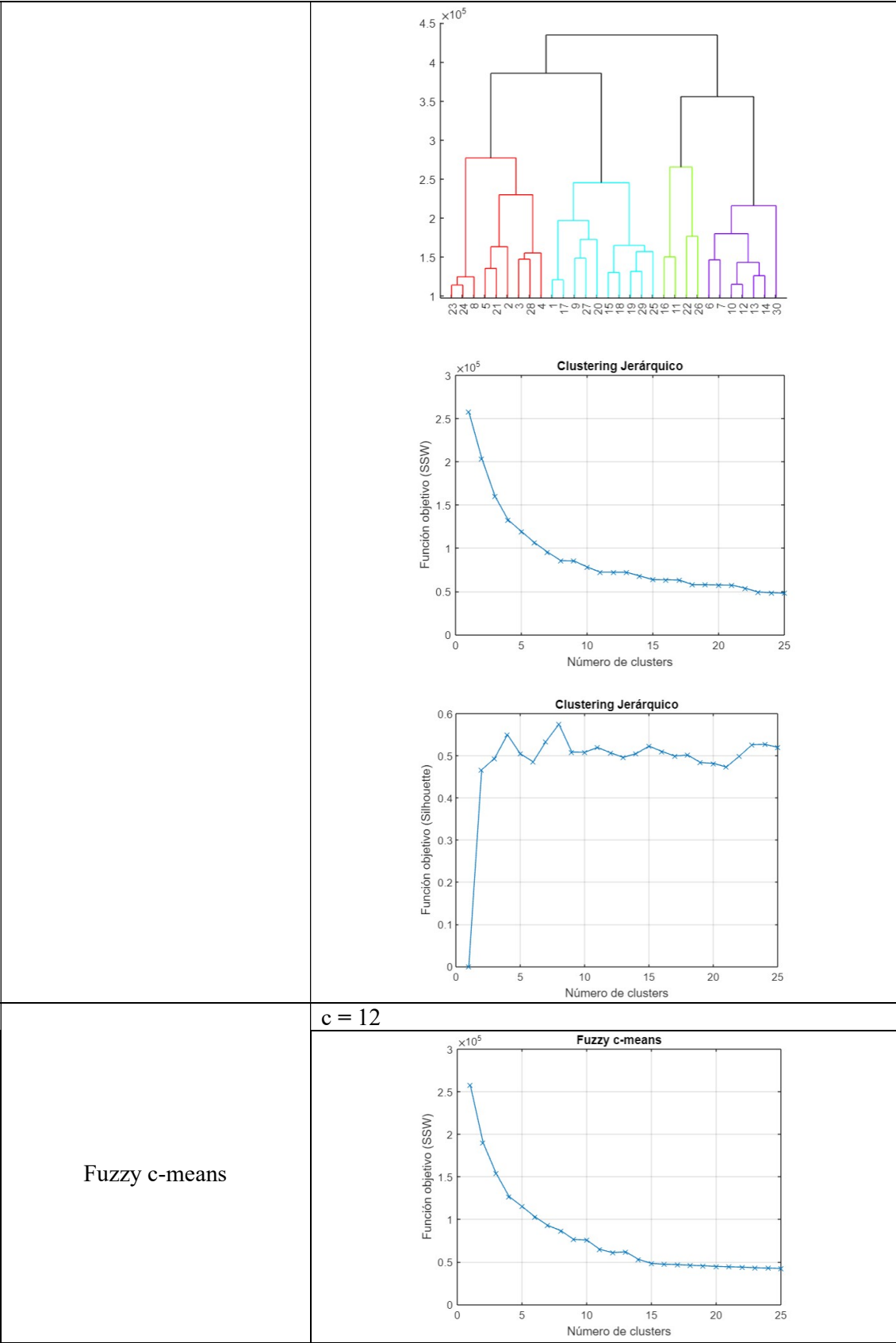
	k-means	k-medoids	DBSCAN	Clustering Jerárquico	Fuzzy c-means
SSW	2,5454	2,6119	2,5133	2,6600	2,5221
SSB	9,0522	9,0185	4,2708	8,9698	9,0251
WB-index	4,2179	4,3442	1,7655	4,4482	4,1917
Silhouette	0,5150	0,4589	0,6579	0,4907	0,4904

4. Dataset S3

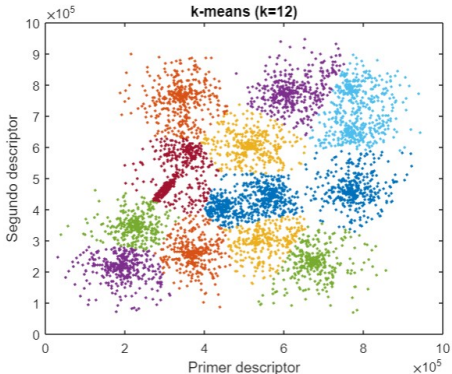
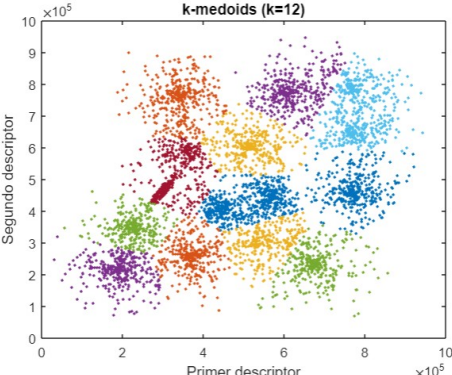
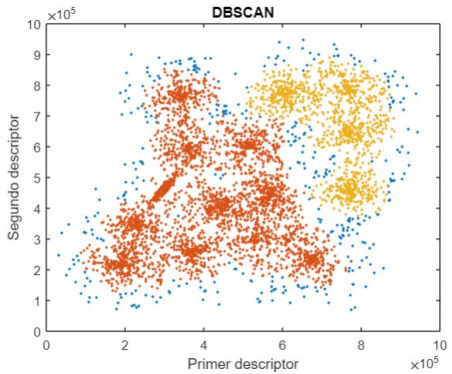
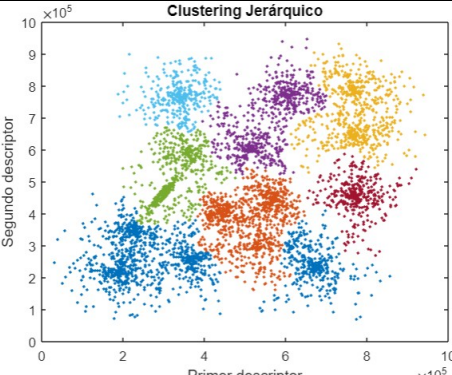
4.1 Estimación de parámetros

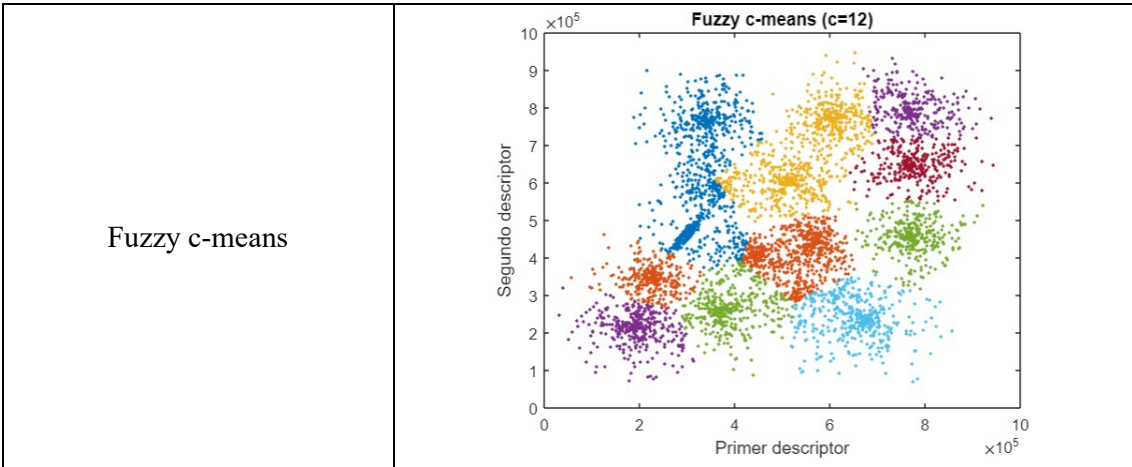


<p>k-medoids</p>	<p>k = 12</p> 
<p>DBSCAN</p>	<p>Mínimo Nro. de vecinos = 19 epsilon = 3,0029 e4</p>   
<p>Clustering Jerárquico</p>	<p>Nro. de clusters = 8 Average linkage</p>



4.2 Visualización de clusters

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



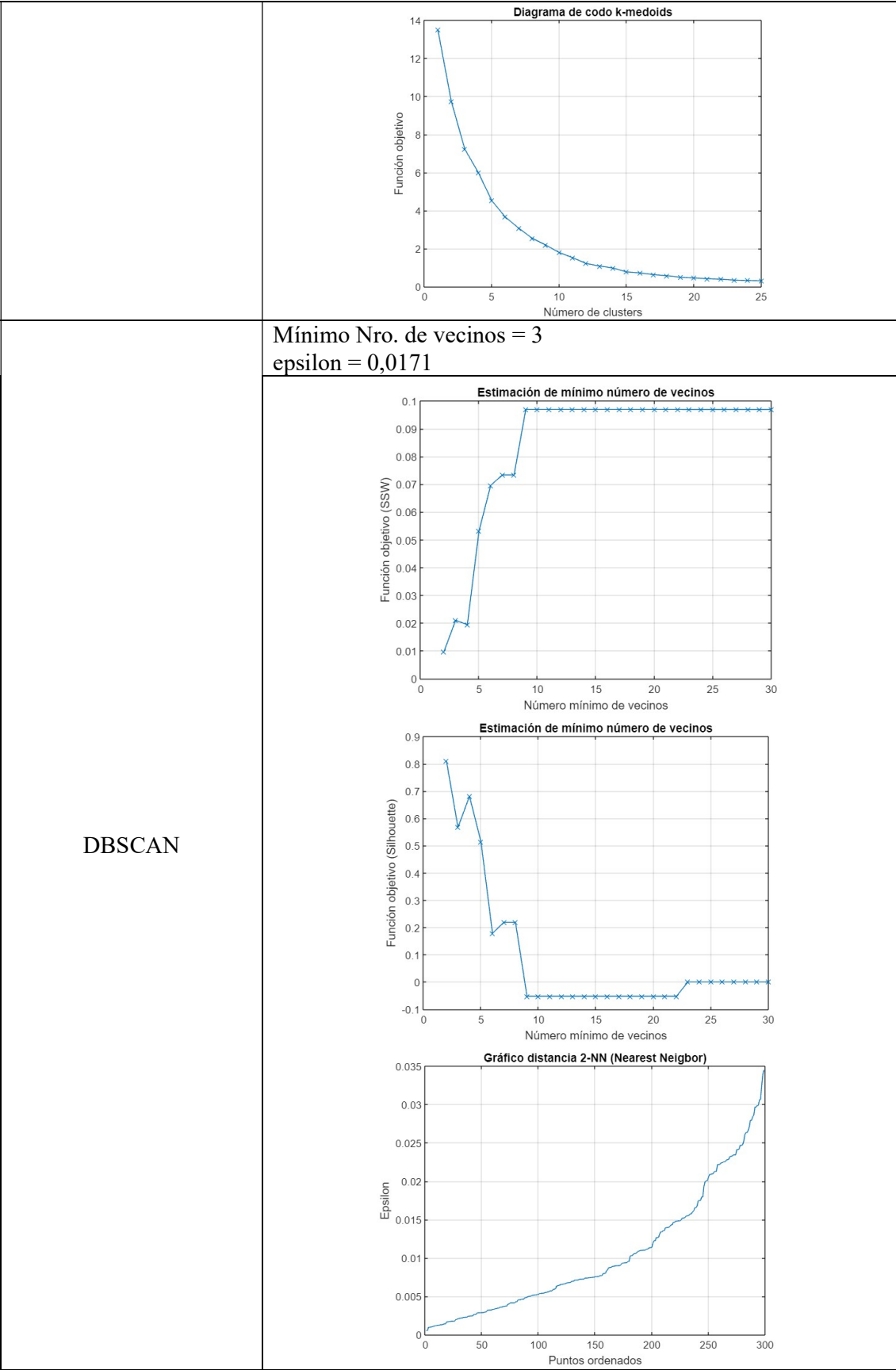
4.3 Métricas

	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

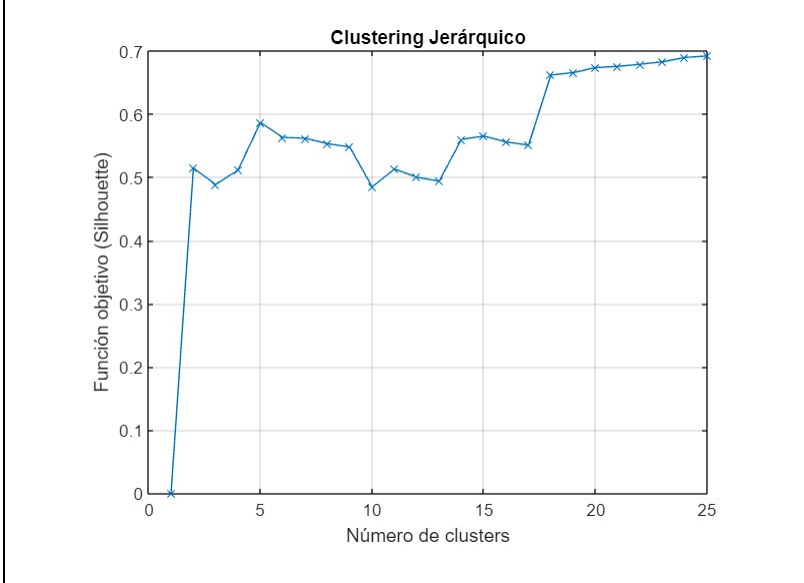
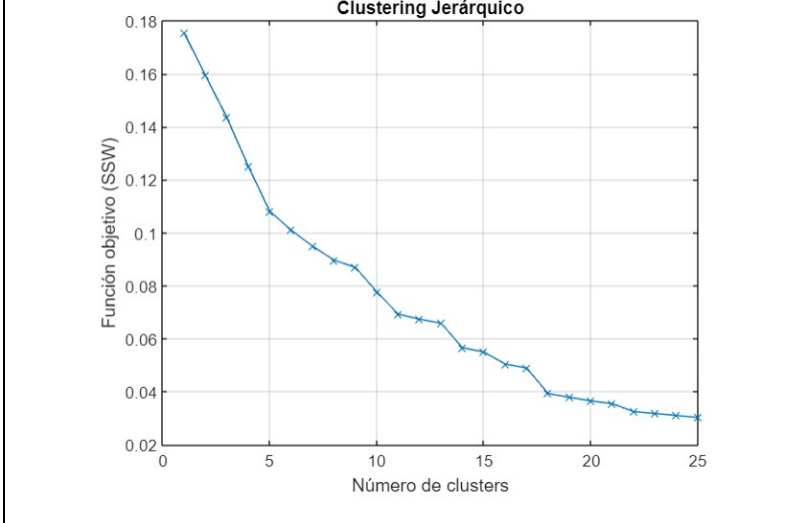
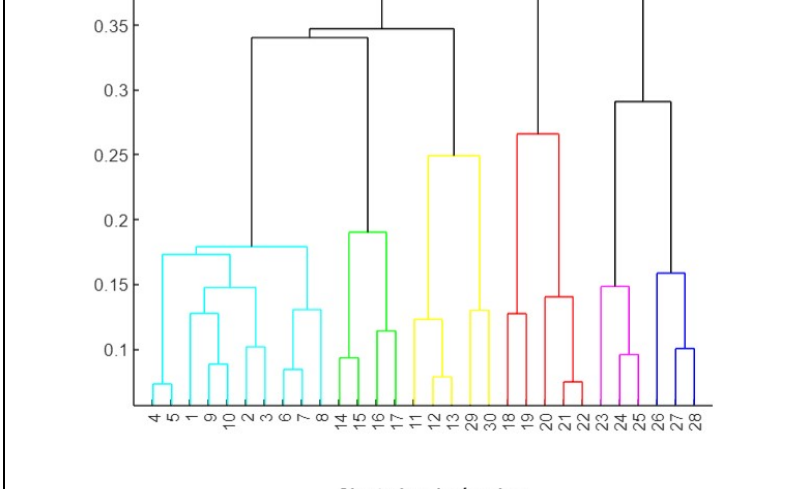
Algoritmo	Estimación de parámetros
k-means	k = 18
k-medoids	k = 18

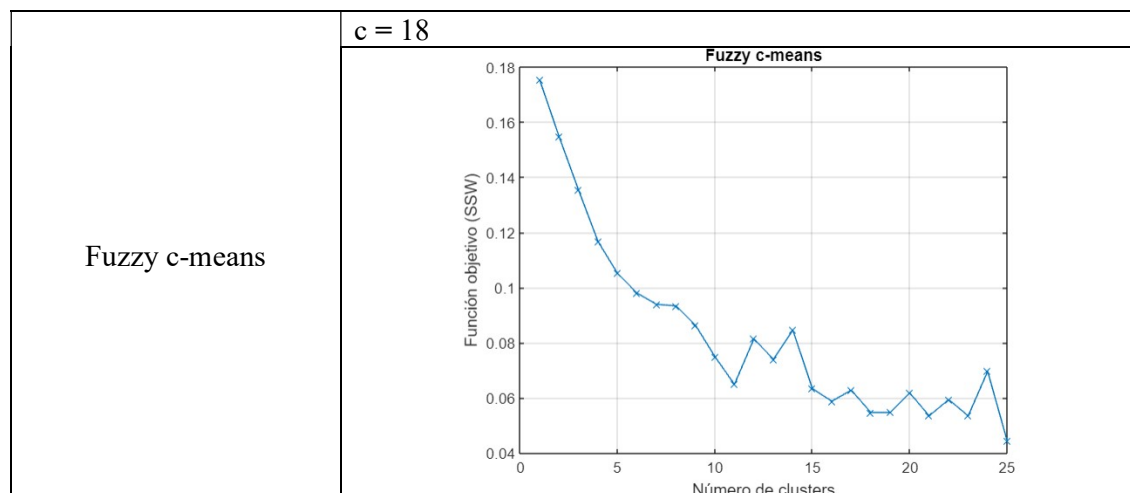


Clustering Jerárquico

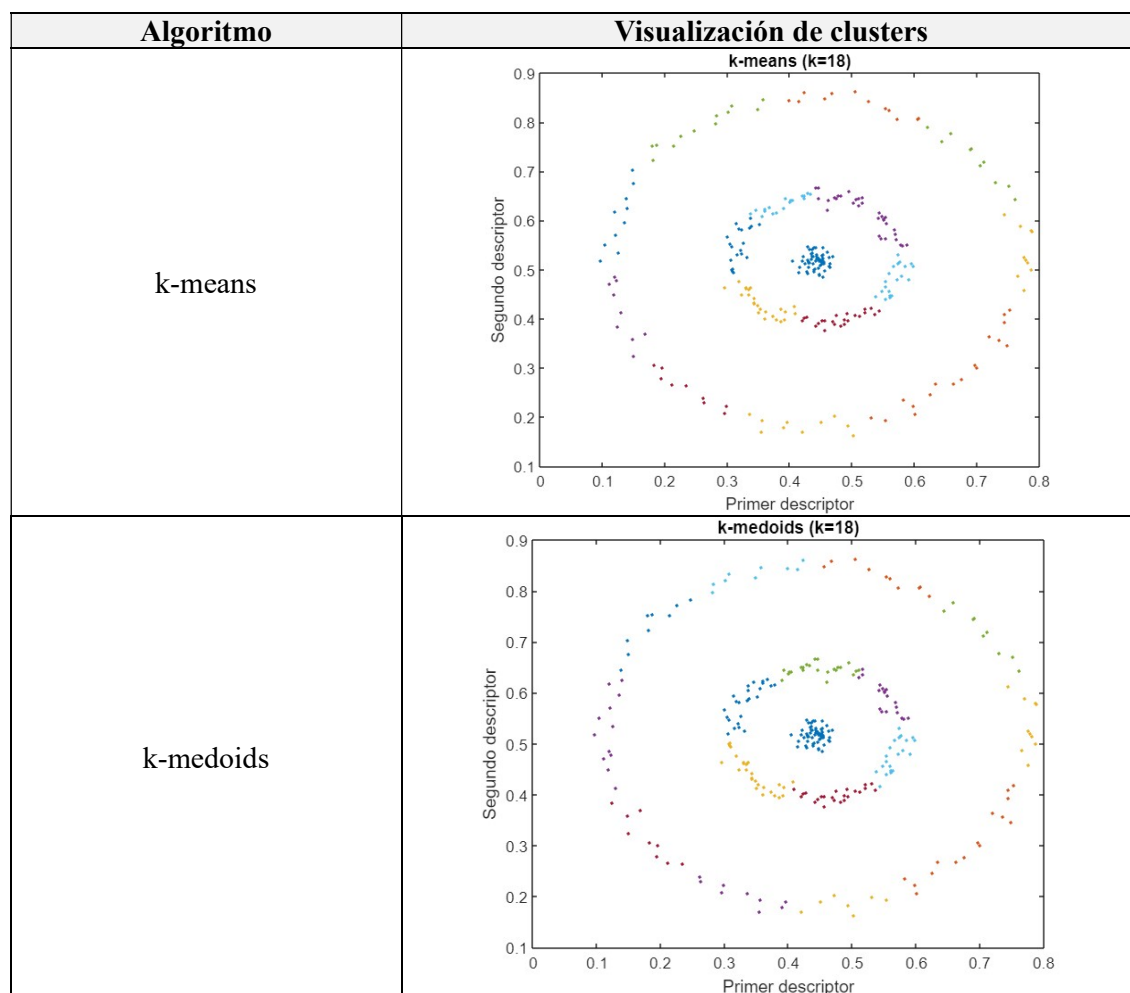
A bar chart showing the proportion of respondents who have been in a romantic relationship in the last 12 months, broken down by gender and age group. The y-axis represents the proportion, with a tick mark at 0.4. The x-axis shows age groups: 18-24, 25-34, 35-44, 45-54, 55-64, and 65+. For each age group, there are two bars: a blue bar for 'Male' and a red bar for 'Female'. The proportions generally increase with age, with the highest proportions in the 65+ age group.

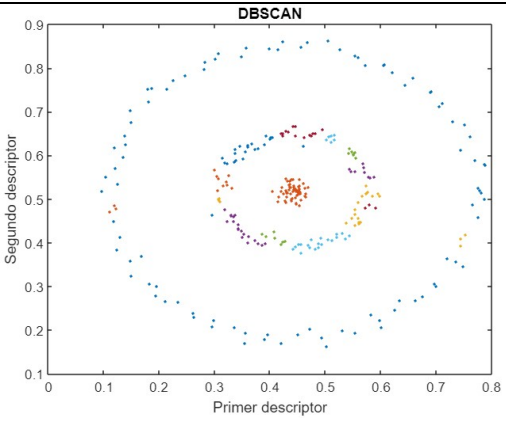
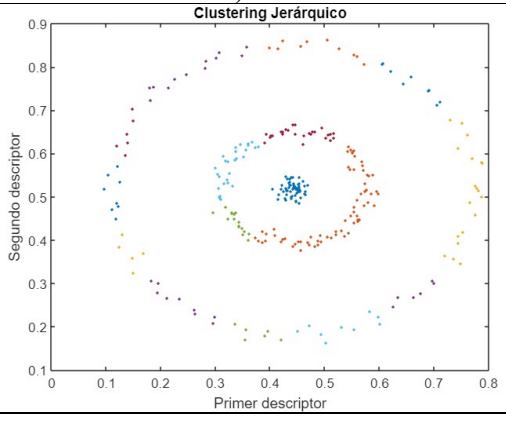
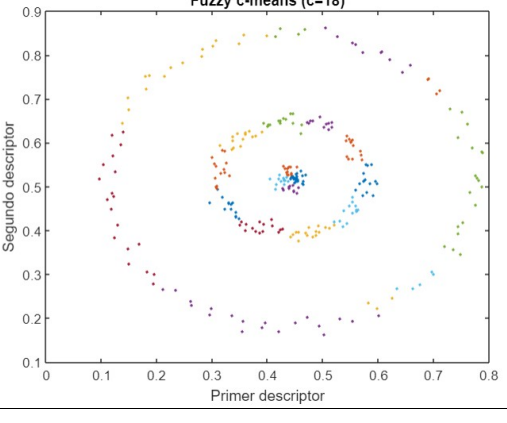
Age Group	Male	Female
18-24	~0.15	~0.18
25-34	~0.25	~0.28
35-44	~0.35	~0.38
45-54	~0.45	~0.48
55-64	~0.55	~0.58
65+	~0.65	~0.68





5.2 Visualización de clusters



DBSCAN	 <p>(16 clústeres en total más un cluster de anomalías) (91 observaciones anómalas)</p>
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