

LABORATORIO 2

SEGMENTACIÓN DE CLIENTES PARA FINANZASALPES

GRUPO 10



PREPROCESAMIENTO

- Perfilamiento
 - Se graficaron las distribuciones de los datos de manera individual y se identificaron posibles grupos. Asimismo los rangos entre variables.
- Calidad de datos
 - Lo mas importante fue la imputación de los datos con KNNImputer para no desechar cerca de 300 datos. Además, se escalaron los datos para que estos quedaran entre 0 y 1 para que las unidades entre todos fueran comparables. Todo lo demás fue el procedimiento normal que se realizaría en calidad de datos.



ALGORITMOS UTILIZADOS

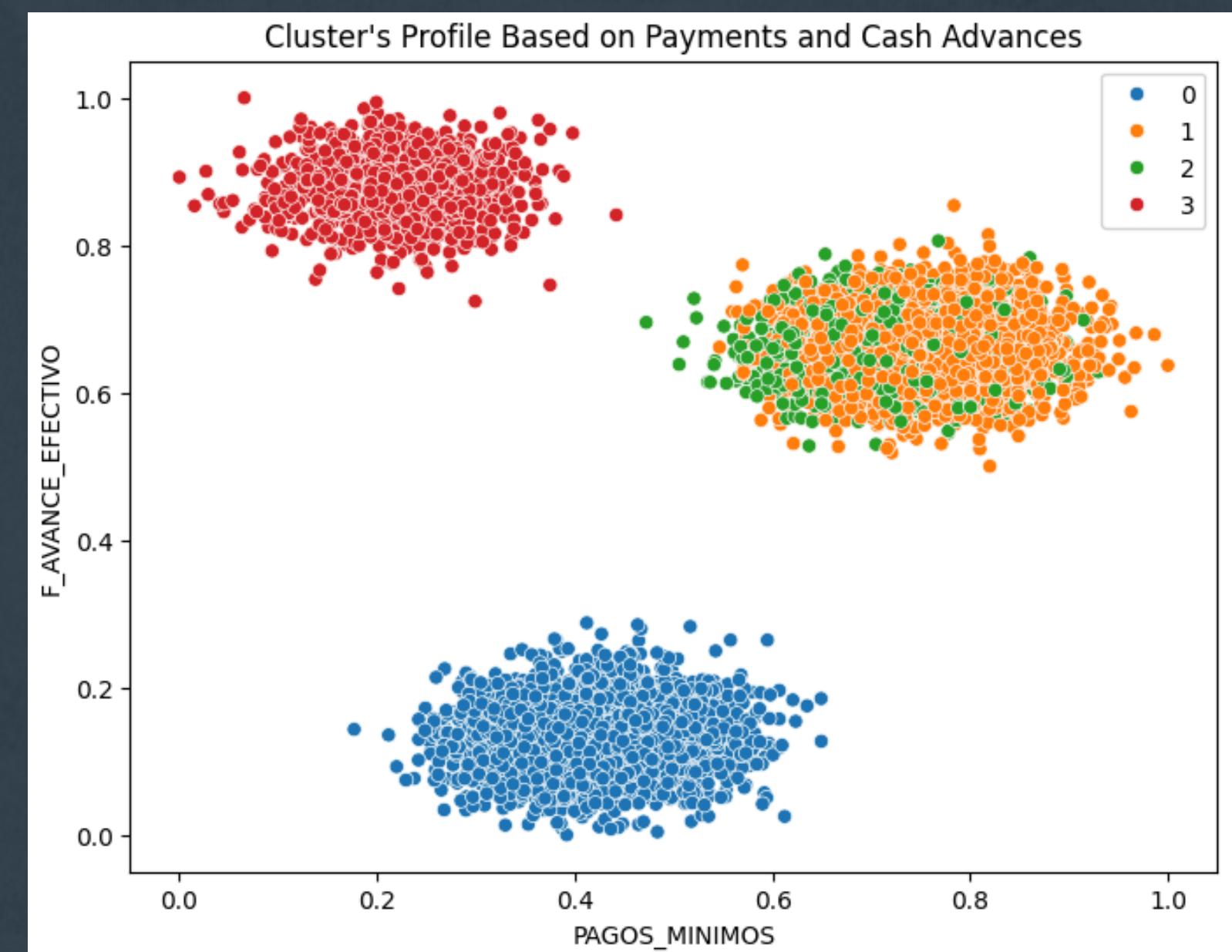
K-Means

DBScan

GMM

K-MEANS

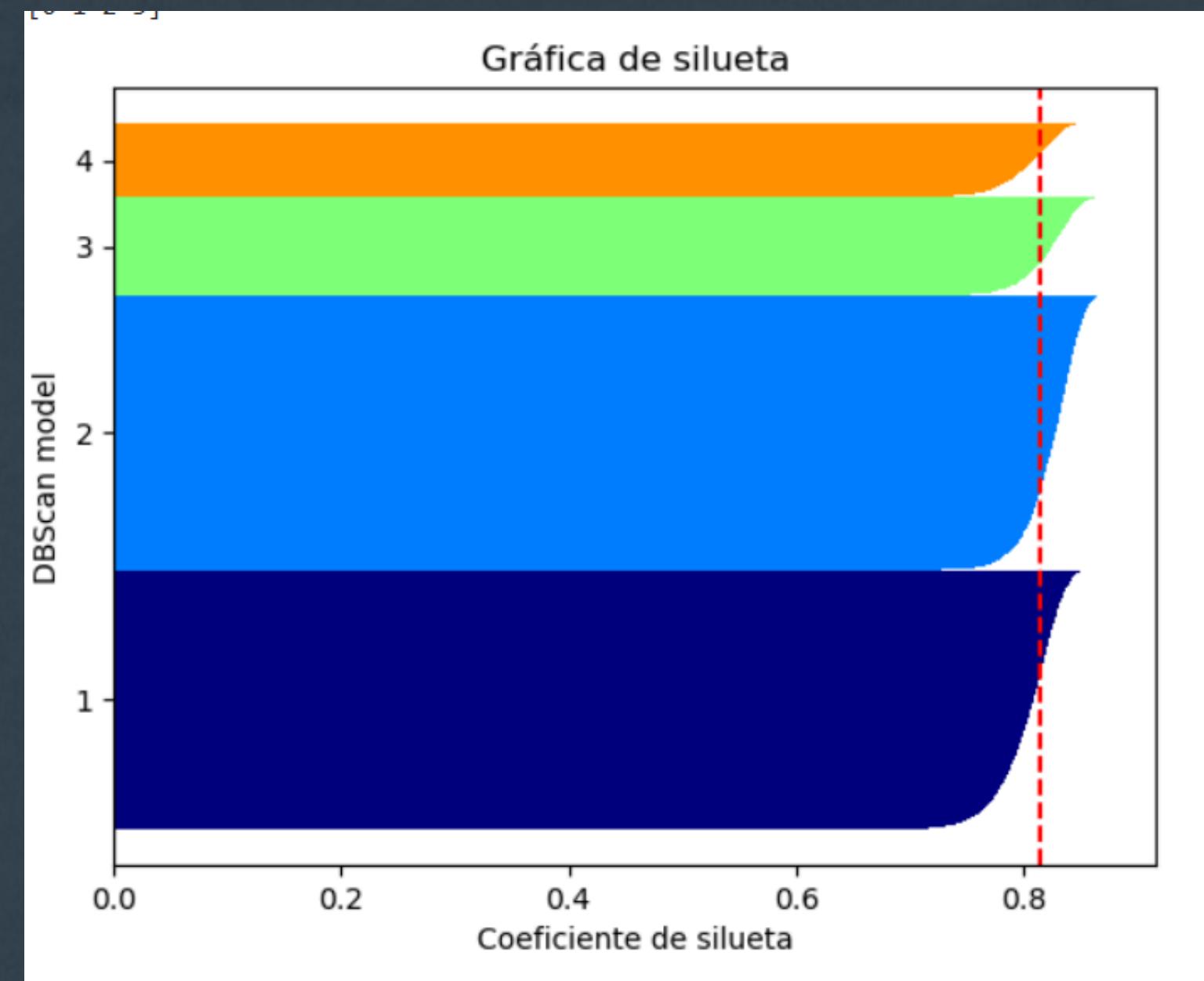
por: Juan Diego Osorio



DBSCAN

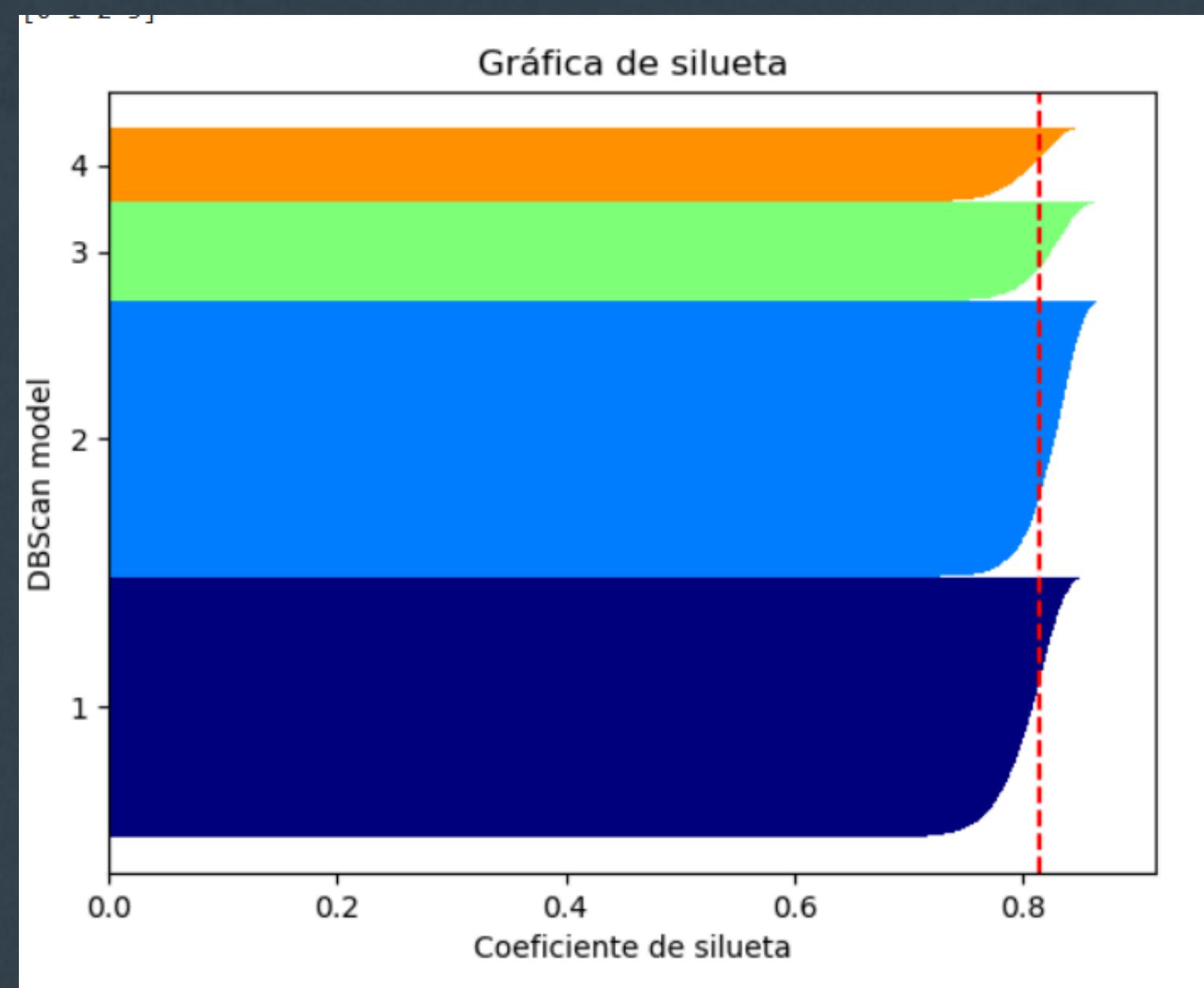
por: Javier Barrera

Luego de realizar el modelo se puede observar que los datos están bien ajustados. Esto indica que los puntos dentro de cada clúster están bien agrupados y que los clústeres están bien separados entre sí. Además, se incluyen los cuatro grupos encontrados por el algoritmo.



INSTANCIAS Y OUTLIERS

En el primer grupo se encuentran 3284 instancias, en el segundo hay 3481, en el tercero 1247 y en el último hay 938. Además, se sabe que no hay bastantes outliers debido a que se definió que la distancia mínima que se definió para que una instancia fuera de un grupo fue bastante considerable.

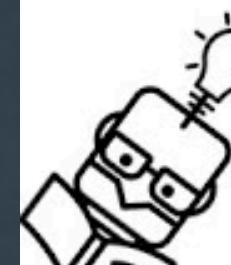


MODELO GAUSSIAN MIXTURE MODELS (GMM)

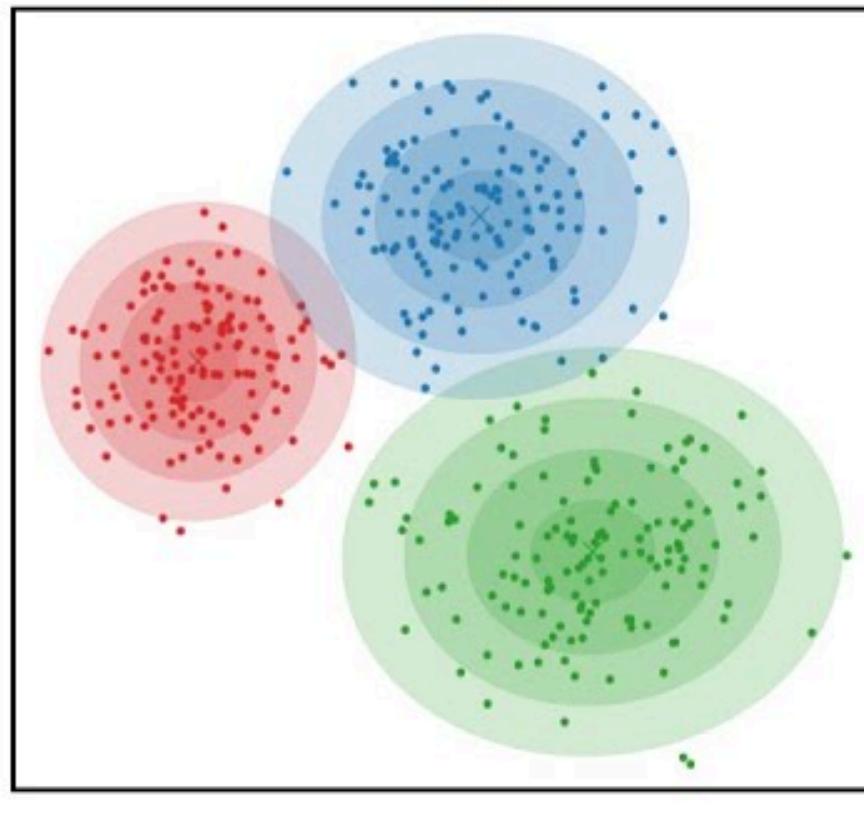
POR: DANIEL FELIPE ORTIZ

Gaussian Mixture Models

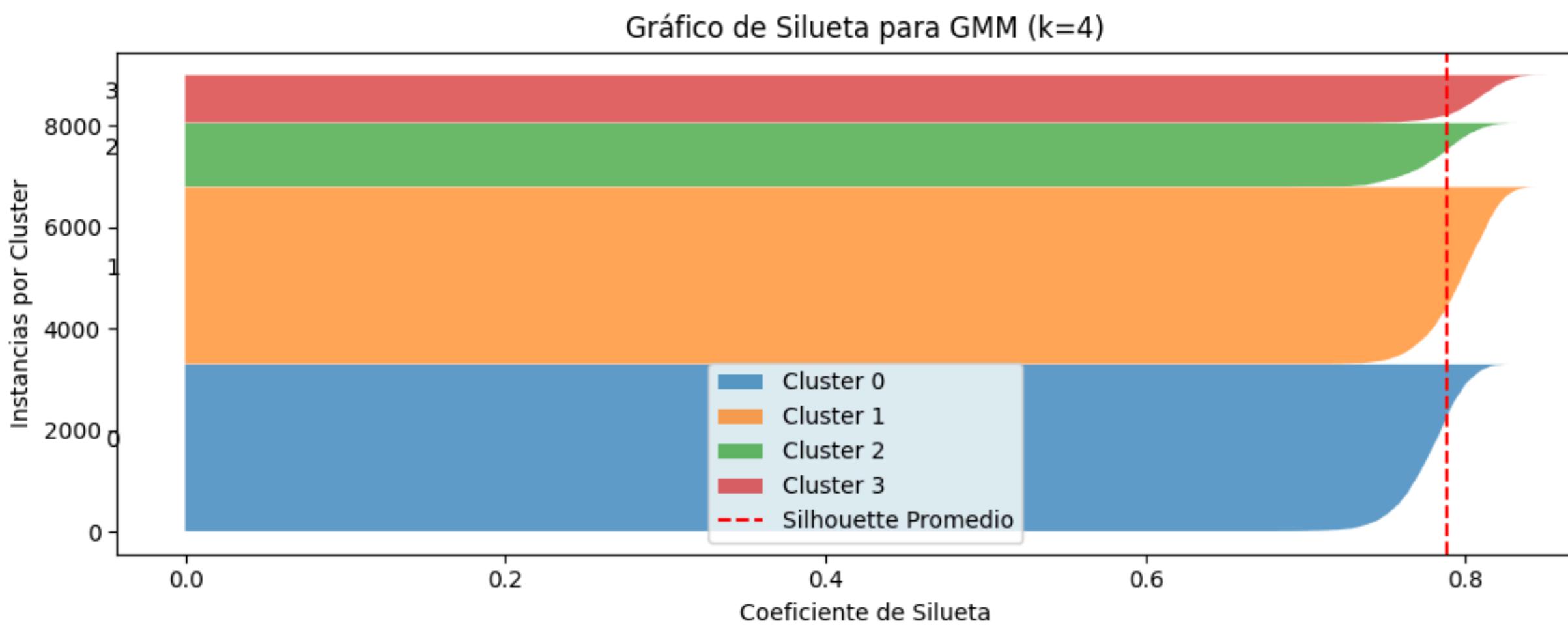
$$p(x) = \sum_{i=1}^K \phi_i \mathcal{N}(x|\mu_i, \sigma_i)$$



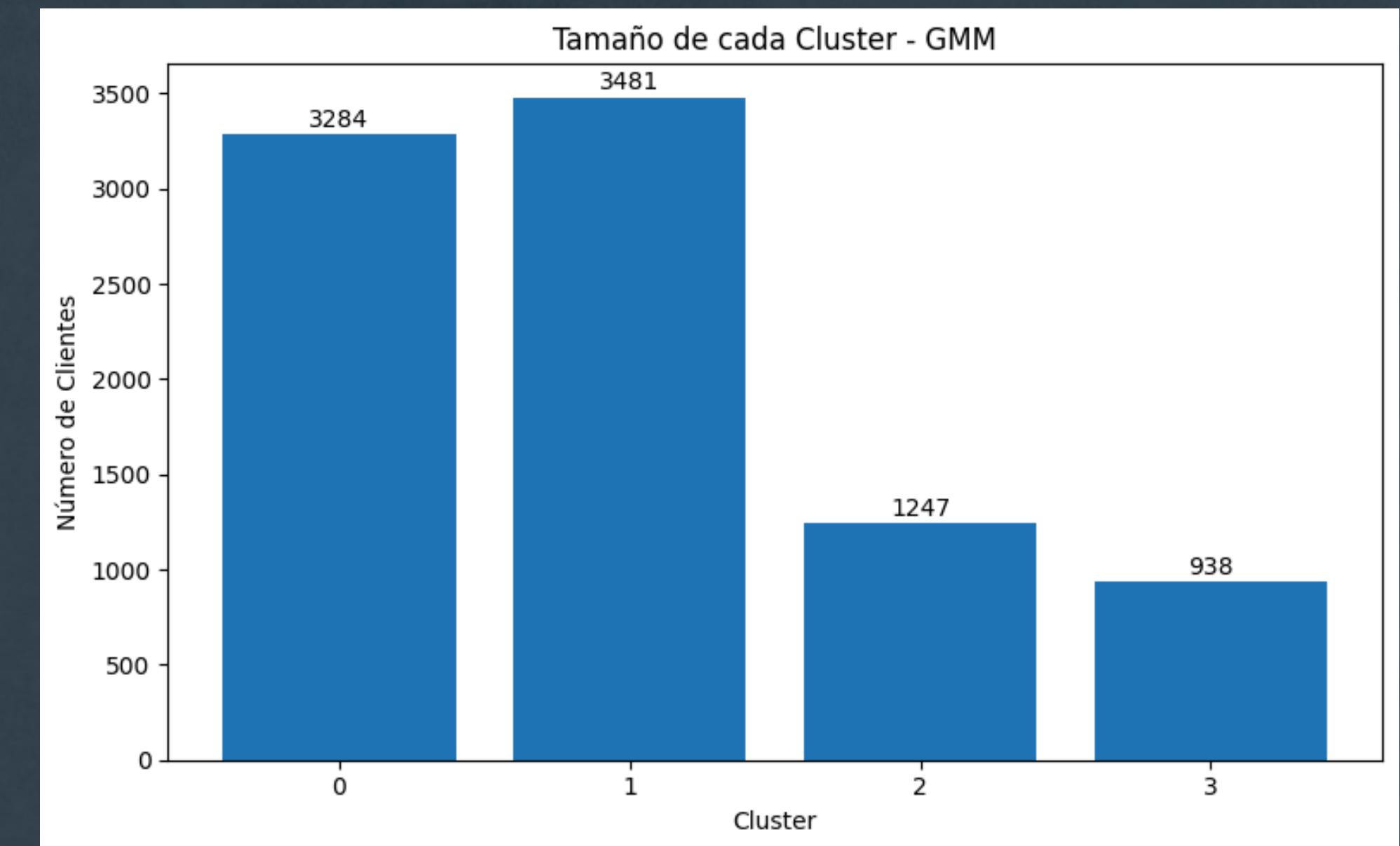
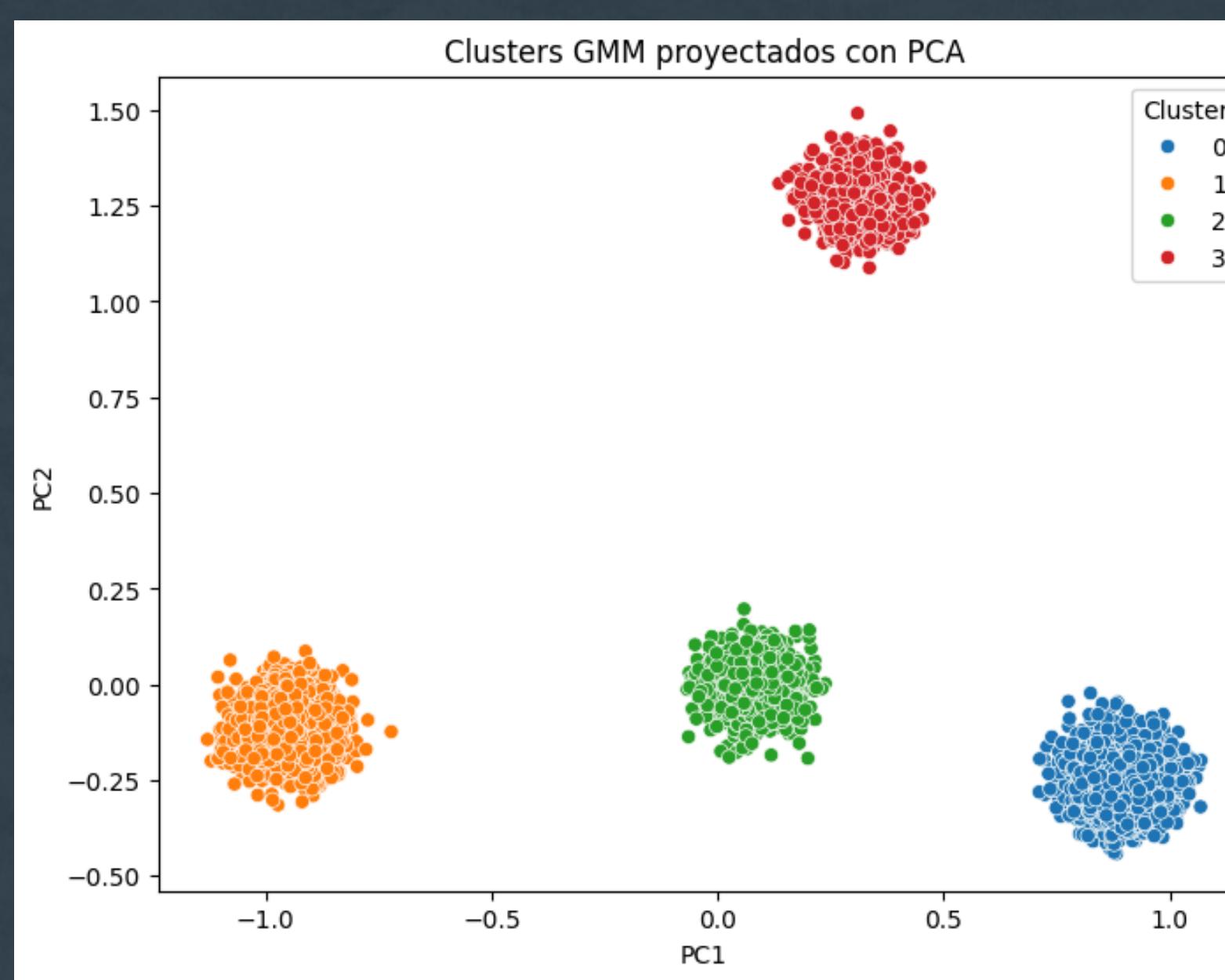
$$\sum_{i=1}^K \phi_i = 1$$



RESULTADOS NUMERO DE CLUSTERS



RESULTADO DE 4 CLUSTERS



RELACION CLUSTERS CON COLUMNAS DEL DATABASE



Aplicamos el modelo GMM con 4 clusters, logrando una segmentación clara y bien diferenciada, respaldada tanto por el gráfico PCA como por el análisis del coeficiente de silueta. Cada grupo presenta patrones únicos: clientes responsables con compras puntuales y bajo uso de crédito (Cluster 0), usuarios intensivos del crédito con compras a plazos (Cluster 1), clientes con comportamiento financiero equilibrado (Cluster 2) y perfiles con alto poder adquisitivo y mayor uso de avances (Cluster 3). Esta segmentación permite a FinanzasAlpes diseñar estrategias comerciales específicas y personalizadas según el perfil de cada cliente.