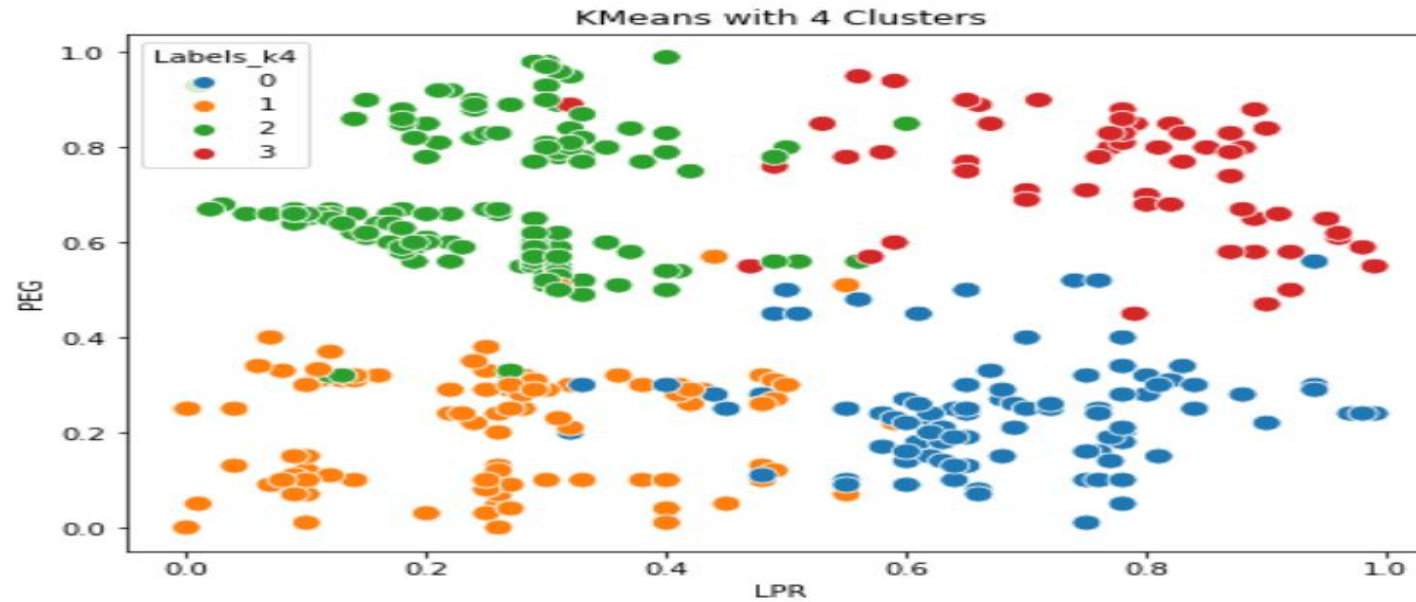


K MEAN CON 4 CLUSTERS

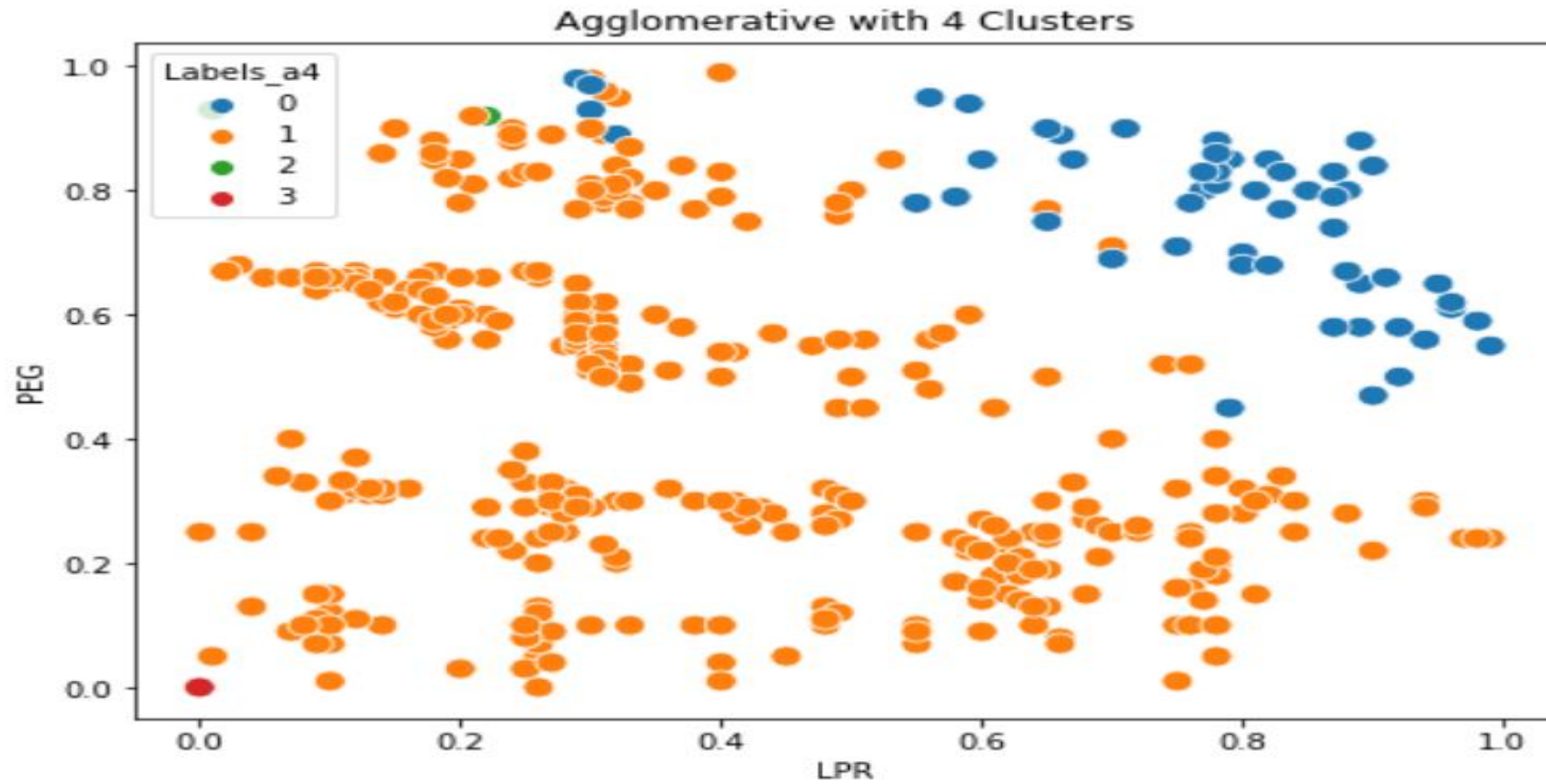


Se observa que con el k-mean se distinguen claramente las 4 etiquetas

parece una buena opcion tomar 4 cluster:

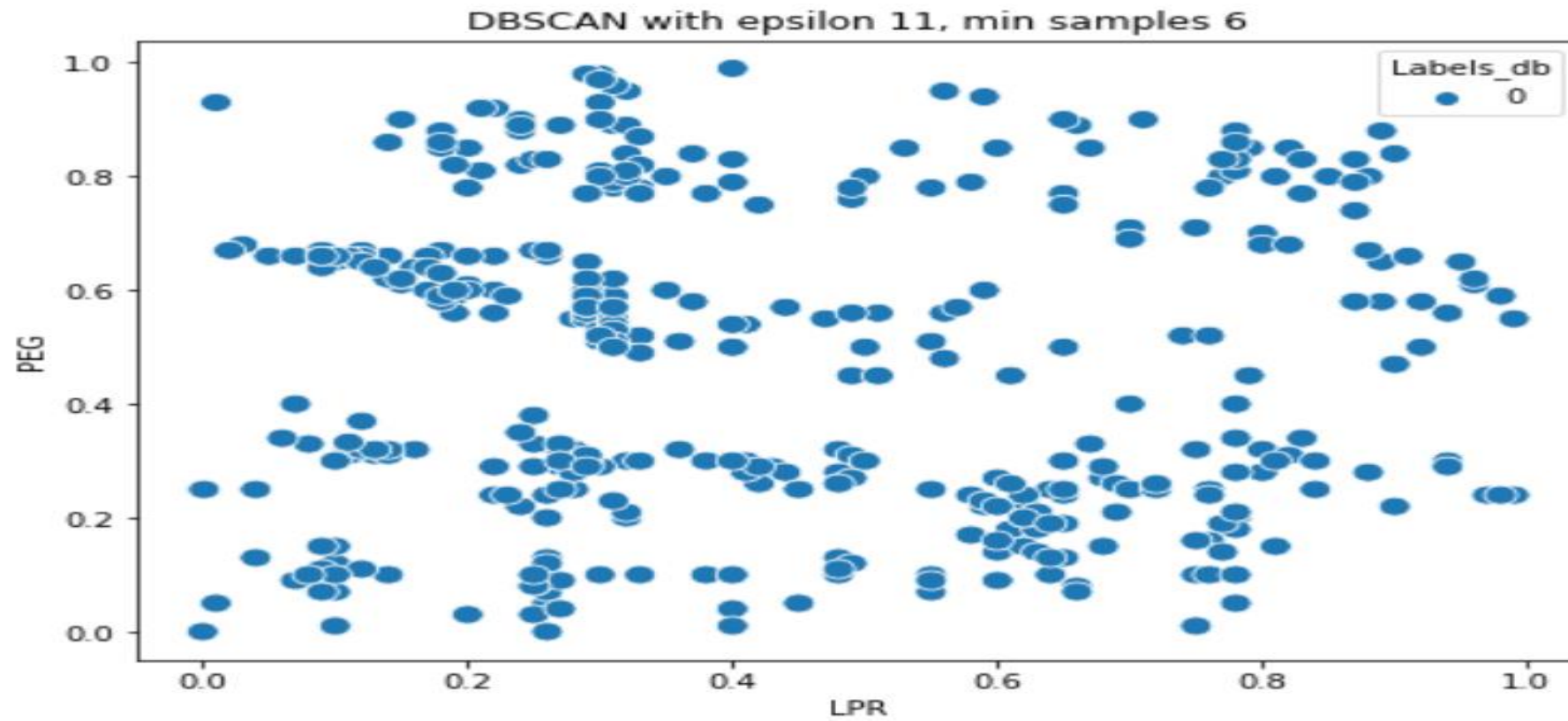
- . cluster 0: ALTO LPR Y BAJO PEG
- . cluster 1: BAJO LPR Y ALTO PEG
- . cluster 2: BAJO LPR Y BAJO PEG
- . cluster 3: ALTO LPR Y ALTO PEG

CLUSTERING JERARQUICO



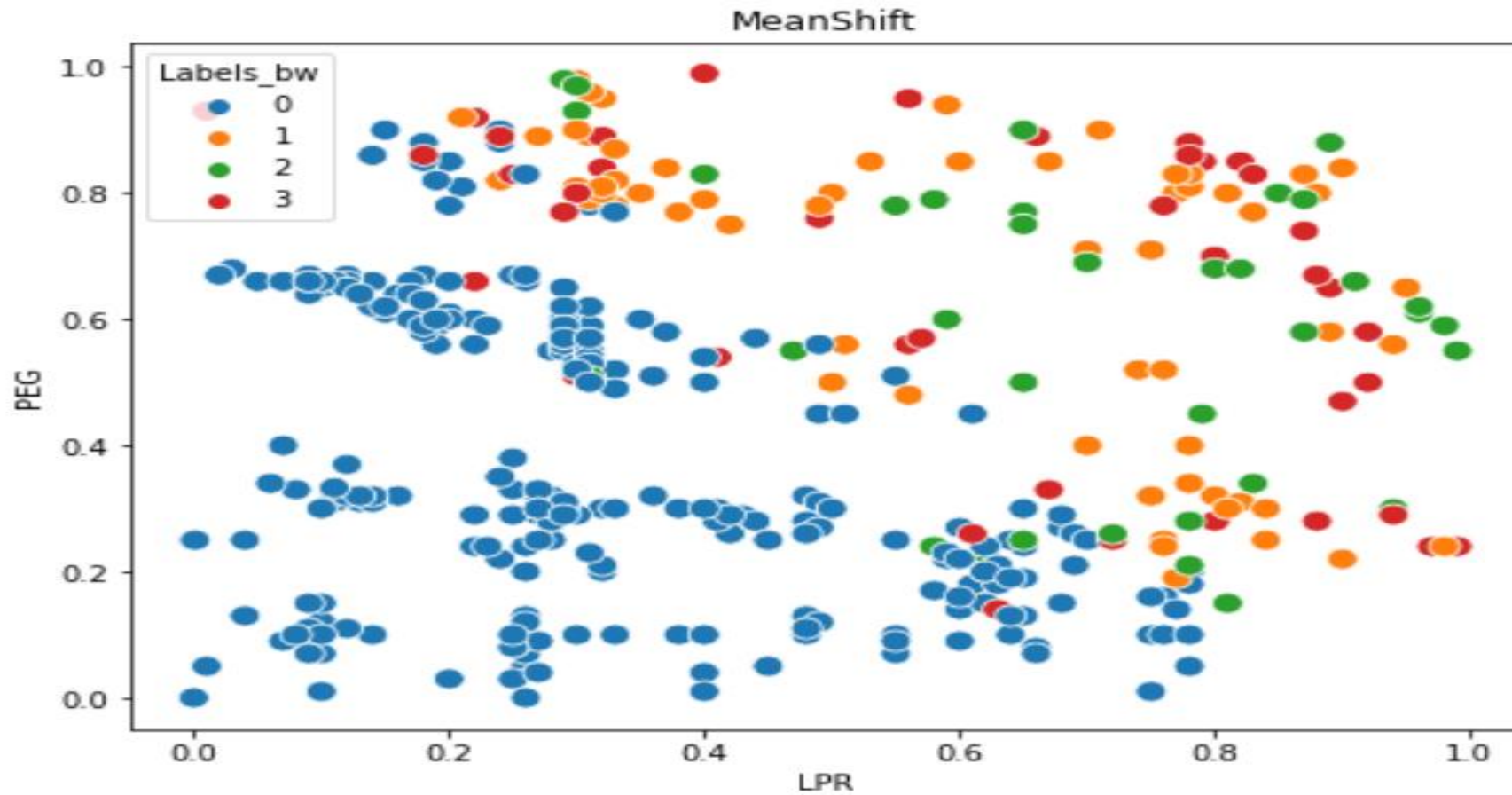
Se observa que con el clustering jerárquico no se visualizan las 4 labels. solo se aprecian la 0 y la 1. El K-mean seria una mejor opción.

DBSCAN



El DBSCAN no sería el mejor algoritmo puesto que solo se visualiza 1 label.

MEAN SHIFT



El mean shift no sería el mejor algoritmo puesto que solo se visualizan 3 labels y no están mejor clasificados que el k-mean

CONCLUSIONES

De los 4 algoritmos se escoge k-mean con 4 labels puesto que es el que mejor divide a los datos. Los otros algoritmos de agrupamiento como lo son : clustering jerárquico, DBSCAN y mean shift no realizan el clustering de una forma adecuada por lo tanto no son la mejor opcion.

PORCENTAJE DE ERROR

```
from sklearn.metrics import accuracy_score, mean_absolute_percentage_error  
  
(1 - accuracy_score(data1["UNS_Kmeans"], data1["UNS"]))*100
```

97.02233250620348