## **Breakdown of Key Hyperparameters**

## 1 KMeans Hyperparameters

| Hyperparameter | Current Setting                            | Role & Impact  |
|----------------|--|--|
| n_clusters     | <b>6</b> (Optimized via Silhouette Method) | Defines the number of clusters. Directly impacts results.            |
| init           | Defaults to 'k-means++'                    | Helps with better initial centroid placement to improve convergence. |
| max_iter       | Defaults to 300                            | Limits how long the algorithm runs before stopping.                  |
| tol            | Defaults to 1e-4                           | Determines when the model converges (stops updating centroids).      |
| n_init         | Defaults to 10                             | Runs KMeans multiple times to avoid poor local minima.               |

```
Training KMeans with different K values...

K=2: Silhouette Score=0.5673, Inertia=83212.63, Training Time=0.0194 sec K=3: Silhouette Score=0.5858, Inertia=47215.83, Training Time=0.0190 sec K=4: Silhouette Score=0.6813, Inertia=19465.56, Training Time=0.0228 sec K=5: Silhouette Score=0.6976, Inertia=14089.74, Training Time=0.0253 sec K=6: Silhouette Score=0.7251, Inertia=9835.57, Training Time=0.0333 sec K=7: Silhouette Score=0.7026, Inertia=6635.59, Training Time=0.0661 sec K=8: Silhouette Score=0.7160, Inertia=4525.62, Training Time=0.0443 sec K=9: Silhouette Score=0.7161, Inertia=3701.13, Training Time=0.0447 sec K=10: Silhouette Score=0.7126, Inertia=3250.03, Training Time=0.0444 sec
```

Highest Silhouette Score: 0.7251

## **2** DBSCAN Hyperparameters

| Hyperparameter | Current Setting            | Role & Impact   |
|----------------|----------------------------|---|
| eps            | 1.50 (Optimized)           | Defines how far two points must be to be in the same cluster. A crucial tuning parameter. |
| min_samples    | 10 (Optimized)             | The minimum number of points required to form a dense cluster.                            |
| metric         | Defaults to<br>'euclidean' | Defines how distance between points is measured.  |
| algorithm      | Defaults to 'auto'         | Automatically chooses the best implementation of DBSCAN.                                  |

```
Training DBSCAN with different parameter combinations...
eps=0.50, min_samples=3: Silhouette Score=0.3488, Clusters Found=45, Training Time=0.012365 sec
eps=0.50, min_samples=5: Silhouette Score=0.2887, Clusters Found=23, Training Time=0.016127 sec
eps=0.50, min_samples=7: Silhouette Score=0.3358, Clusters Found=8, Training Time=0.031650 sec
eps=0.50, min_samples=10: Silhouette Score=0.4491, Clusters Found=4, Training Time=0.037517 sec
eps=1.00, min_samples=3: Silhouette Score=0.7135, Clusters Found=6, Training Time=0.029120 sec
eps=1.00, min_samples=5: Silhouette Score=0.7070, Clusters Found=6, Training Time=0.024105 sec
eps=1.00, min_samples=7: Silhouette Score=0.6807, Clusters Found=7, Training Time=0.014233 sec
eps=1.00, min_samples=10: Silhouette Score=0.6190, Clusters Found=9, Training Time=0.022724 sec
eps=1.50, min_samples=3: Silhouette Score=0.6894, Clusters Found=5, Training Time=0.030598 sec
eps=1.50, min_samples=5: Silhouette Score=0.6894, Clusters Found=5, Training Time=0.034070 sec
eps=1.50, min_samples=7: Silhouette Score=0.6888, Clusters Found=5, Training Time=0.032876 sec
eps=1.50, min_samples=10: Silhouette Score=0.7223, Clusters Found=6, Training Time=0.033617 sec
eps=2.00, min_samples=3: Silhouette Score=0.6133, Clusters Found=5, Training Time=0.043511 sec
eps=2.00, min_samples=5: Silhouette Score=0.6133, Clusters Found=5, Training Time=0.037941 sec
eps=2.00, min_samples=7: Silhouette Score=0.6133, Clusters Found=5, Training Time=0.044316 sec
eps=2.00, min samples=10: Silhouette Score=0.6133, Clusters Found=5, Training Time=0.036891 sec
eps=2.50, min_samples=3: Silhouette Score=0.6923, Clusters Found=5, Training Time=0.034854 sec
eps=2.50, min_samples=5: Silhouette Score=0.6923, Clusters Found=5, Training Time=0.040720 sec
eps=2.50, min_samples=7: Silhouette Score=0.6923, Clusters Found=5, Training Time=0.041262 sec
eps=2.50, min_samples=10: Silhouette Score=0.6923, Clusters Found=5, Training Time=0.035165 sec
eps=3.00, min_samples=3: Silhouette Score=0.6923, Clusters Found=5, Training Time=0.043361 sec
eps=3.00, min_samples=5: Silhouette Score=0.6923, Clusters Found=5, Training Time=0.028957 sec
eps=3.00, min_samples=7: Silhouette Score=0.6923, Clusters Found=5, Training Time=0.036390 sec
eps=3.00, min_samples=10: Silhouette Score=0.6923, Clusters Found=5, Training Time=0.019533 sec
eps=3.50, min_samples=3: Silhouette Score=0.6923, Clusters Found=5, Training Time=0.018263 sec
eps=3.50, min_samples=5: Silhouette Score=0.6923, Clusters Found=5, Training Time=0.018759 sec
eps=3.50, min_samples=7: Silhouette Score=0.6923, Clusters Found=5, Training Time=0.018493 sec
eps=3.50, min_samples=10: Silhouette Score=0.6923, Clusters Found=5, Training Time=0.023596 sec
eps=4.00, min_samples=3: Silhouette Score=0.6923, Clusters Found=5, Training Time=0.018119 sec
eps=4.00, min_samples=5: Silhouette Score=0.6923, Clusters Found=5, Training Time=0.033741 sec
eps=4.00, min_samples=7: Silhouette Score=0.6923, Clusters Found=5, Training Time=0.017404 sec
eps=4.00, min_samples=10: Silhouette Score=0.6923, Clusters Found=5, Training Time=0.026521 sec
eps=4.50, min_samples=3: Silhouette Score=0.5180, Clusters Found=4, Training Time=0.019079 sec
eps=4.50, min_samples=5: Silhouette Score=0.5180, Clusters Found=4, Training Time=0.018052 sec
eps=4.50, min_samples=7: Silhouette Score=0.5180, Clusters Found=4, Training Time=0.017792 sec
eps=4.50, min_samples=10: Silhouette Score=0.5180, Clusters Found=4, Training Time=0.017603 sec
eps=5.00, min_samples=3: Silhouette Score=0.5866, Clusters Found=3, Training Time=0.017943 sec
eps=5.00, min_samples=5: Silhouette Score=0.5866, Clusters Found=3, Training Time=0.018959 sec
eps=5.00, min_samples=7: Silhouette Score=0.5866, Clusters Found=3, Training Time=0.017376 sec
eps=5.00, min_samples=10: Silhouette Score=0.5866, Clusters Found=3, Training Time=0.017498 sec
```

Highest Silhouette Score: 0.7223

## 3 Agglomerative Hyperparameters (WINNER)

| Hyperparameter | Current Setting                     | Role & Impact                                 |
|----------------|-------------------------------------|---|
| n_clusters     | 6 (Optimized via Silhouette Method) | Defines the number of clusters.               |
| linkage        | 'ward' (Fixed)                      | Uses variance minimization to merge clusters. |
| affinity       | Defaults to 'euclidean'             | Determines distance measure for clustering.   |

```
Training Agglomerative Clustering with different K values...

K=2: Silhouette Score=0.5698, Inertia=86589.58, Training Time=0.0339 sec K=3: Silhouette Score=0.5806, Inertia=47923.96, Training Time=0.0562 sec K=4: Silhouette Score=0.6813, Inertia=19483.46, Training Time=0.0664 sec K=5: Silhouette Score=0.6978, Inertia=14107.64, Training Time=0.0536 sec K=6: Silhouette Score=0.7257, Inertia=9854.77, Training Time=0.0593 sec K=7: Silhouette Score=0.7015, Inertia=6705.23, Training Time=0.0593 sec K=8: Silhouette Score=0.7086, Inertia=4819.96, Training Time=0.0641 sec K=9: Silhouette Score=0.7088, Inertia=3821.04, Training Time=0.0557 sec K=10: Silhouette Score=0.7088, Inertia=3354.75, Training Time=0.0575 sec
```

Highest Silhouette Score: 0.7257