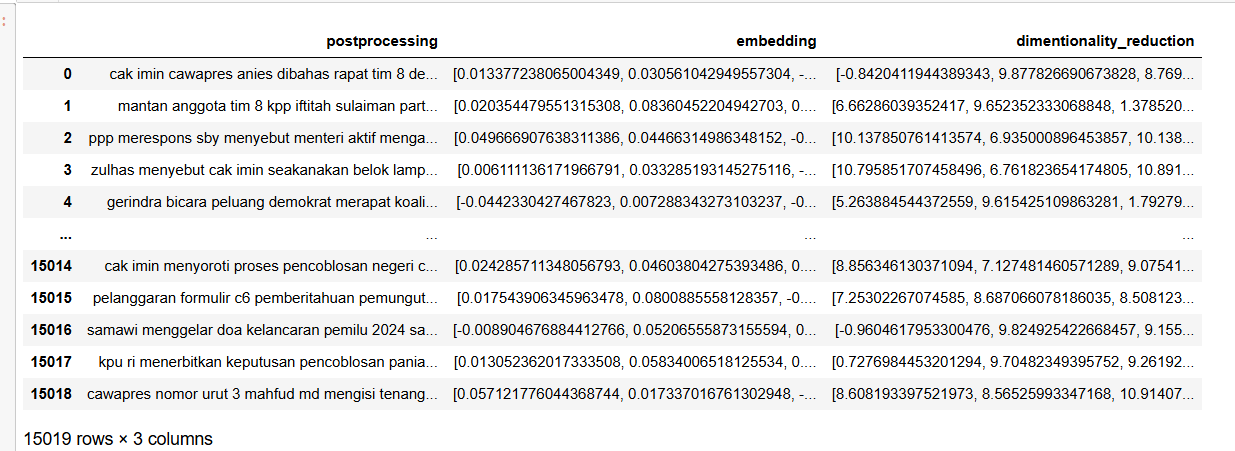
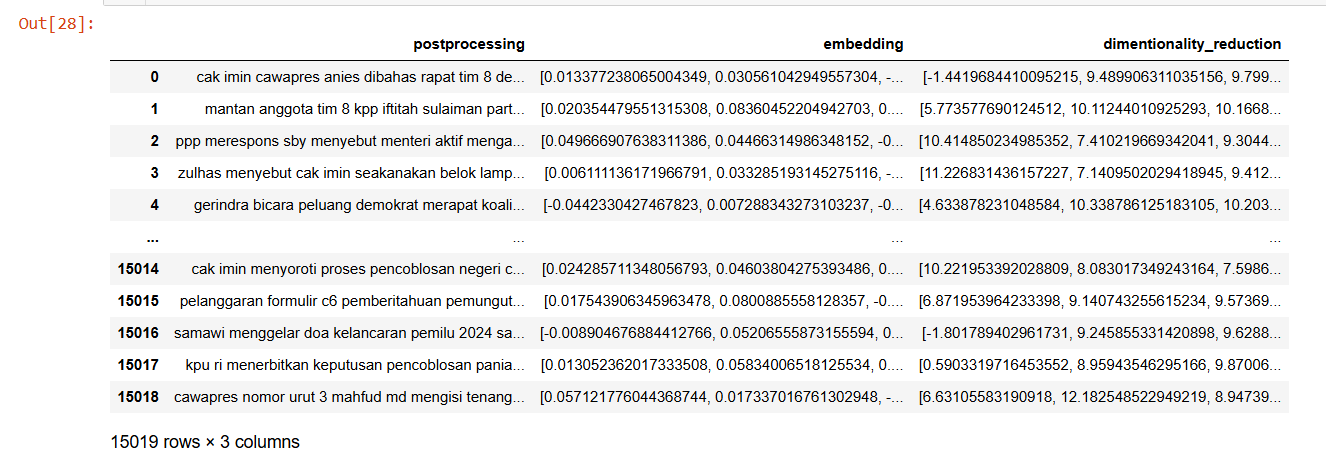
A screenshot of a computer

Description automatically generated  
  




Umap random\_state 42

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A close-up of numbers

Description automatically generated

A number and numbers on a white background

Description automatically generated

A graph with different colored lines

Description automatically generated  
  
A graph with colored spots and red x marks

Description automatically generated  
A screenshot of a computer

Description automatically generated

Evaluasi 4 cluster

A graph of a graph

Description automatically generated with medium confidence  
  
A graph with colored spots

Description automatically generated with medium confidence  
  
A close-up of numbers

Description automatically generated  
  
evaluasi 5 cluster

A white background with black numbers

Description automatically generated

Berdasarkan evaluasi metrik clustering yang Anda berikan:

\*\*Untuk 4 cluster:\*\*

- Silhouette Score: 0.5543

- SSE (Sum of Squared Errors): 95659.75

- Davies-Bouldin Index: 0.7086

- Calinski-Harabasz Index: 17630.22

- Dunn Index: 0.00085

\*\*Untuk 5 cluster:\*\*

- Silhouette Score: 0.5750

- SSE (Sum of Squared Errors): 74563.56

- Davies-Bouldin Index: 0.6460

- Calinski-Harabasz Index: 18024.58

- Dunn Index: 0.00251

\*\*Mana yang lebih baik?\*\*

Untuk menentukan cluster yang lebih baik, mari kita evaluasi masing-masing metrik secara singkat:

1. \*\*Silhouette Score\*\*: Lebih tinggi pada 5 cluster (0.5750) dibandingkan dengan 4 cluster (0.5543). Ini menunjukkan bahwa objek dalam cluster memiliki kesamaan yang lebih tinggi dan terpisah lebih baik dari cluster lainnya pada 5 cluster.

2. \*\*SSE (Sum of Squared Errors)\*\*: Lebih rendah pada 5 cluster (74563.56) dibandingkan dengan 4 cluster (95659.75). SSE yang lebih rendah menunjukkan bahwa cluster pada 5 cluster lebih padat dan lebih kompak secara keseluruhan.

3. \*\*Davies-Bouldin Index\*\*: Lebih rendah pada 5 cluster (0.6460) dibandingkan dengan 4 cluster (0.7086). Nilai yang lebih rendah menunjukkan bahwa cluster pada 5 cluster lebih baik terpisah satu sama lain.

4. \*\*Calinski-Harabasz Index\*\*: Lebih tinggi pada 5 cluster (18024.58) dibandingkan dengan 4 cluster (17630.22). Indeks yang lebih tinggi menunjukkan bahwa cluster pada 5 cluster lebih baik terpisah dan lebih padat.

5. \*\*Dunn Index\*\*: Lebih tinggi pada 5 cluster (0.00251) dibandingkan dengan 4 cluster (0.00085). Dunn Index yang lebih tinggi menunjukkan bahwa jarak antara cluster-clusternya lebih besar pada 5 cluster.

\*\*Kesimpulan\*\*: Dari evaluasi metrik-metrik di atas, 5 cluster memberikan hasil yang lebih baik atau setidaknya kompetitif dalam semua metrik yang ditinjau. Secara khusus, Silhouette Score yang lebih tinggi dan SSE yang lebih rendah pada 5 cluster menunjukkan bahwa 5 cluster mungkin lebih baik dalam membagi data menjadi kelompok-kelompok yang lebih homogen dan terpisah dengan baik.

Jadi, berdasarkan evaluasi ini, 5 cluster kemungkinan lebih baik daripada 4 cluster untuk mewakili struktur data dengan lebih baik. Namun, keputusan akhir juga dapat dipengaruhi oleh tujuan analisis dan interpretasi domain khusus dari data Anda.  
  
A screenshot of a computer

Description automatically generated