

Silhouette Score: 0.0794 (-1  $\leq$  Silhouette Score  $\leq$  1) The best value is 1 and the worst value is -1. Values near 0 indicate overlapping clusters. Negative values generally indicate that a sample has been assigned to the wrong cluster, as a different cluster is more similar.

Dendrogram Options - Distance Metric: euclidean, Linkage Method: average, Data Values Format: tfidf (Norm: I2)

Dendrogram Distances - 0.00716, 0.00962, 0.01053, 0.01062, 0.01067, 0.01141, 0.01356, 0.01487, 0.01532, 0.01844, 0.02162, 0.02296, 0.02374, 0.02814, 0.03529

BaJin\_Autumn\_1940\_1: Scrubbing Options - Punctuation: kept, Lowercase: off, Digits: kept, Tags: kept,

Cutting Options - Cut by [number]: 2, Percentage Overlap: 0,

BaJin\_Autumn\_1940\_2: Scrubbing Options - Punctuation: kept, Lowercase: off, Digits: kept, Tags: kept,

Cutting Options - Cut by [number]: 2, Percentage Overlap: 0,

BaJin\_Fog\_1931\_1: Scrubbing Options - Punctuation: kept, Lowercase: off, Digits: kept, Tags: kept, Cutting Options - Cut by [number]: 2, Percentage Overlap: 0,

BaJin\_Fog\_1931\_2: Scrubbing Options - Punctuation: kept, Lowercase: off, Digits: kept, Tags: kept, Cutting Options - Cut by [number]: 2, Percentage Overlap: 0,

BaJin\_Lightning\_1935\_1: Scrubbing Options - Punctuation: kept, Lowercase: off, Digits: kept, Tags: kept,

Cutting Options - Cut by [number]: 2, Percentage Overlap: 0,

BaJin\_Lightning\_1935\_2: Scrubbing Options - Punctuation: kept, Lowercase: off, Digits: kept, Tags: kept.

Cutting Options - Cut by [number]: 2, Percentage Overlap: 0,

BaJin\_Rain\_1933\_1: Scrubbing Options - Punctuation: kept, Lowercase: off, Digits: kept, Tags: kept,

Cutting Options - Cut by [number]: 2, Percentage Overlap: 0,

BaJin\_Rain\_1933\_2: Scrubbing Options - Punctuation: kept, Lowercase: off, Digits: kept, Tags: kept,

Cutting Options - Cut by [number]: 2, Percentage Overlap: 0,

BaJin\_Spring\_1938\_1: Scrubbing Options - Punctuation: kept, Lowercase: off, Digits: kept, Tags: kept.

Cutting Options - Cut by [number]: 2, Percentage Overlap: 0,

BaJin\_Spring\_1938\_2: Scrubbing Options - Punctuation: kept, Lowercase: off, Digits: kept, Tags: kept,

Cutting Options - Cut by [number]: 2, Percentage Overlap: 0,

BaJin\_TheFamily\_1933\_1: Scrubbing Options - Punctuation: kept, Lowercase: off, Digits: kept, Tags: kept,

Cutting Options - Cut by [number]: 2, Percentage Overlap: 0,

BaJin\_TheFamily\_1933\_2: Scrubbing Options - Punctuation: kept, Lowercase: off, Digits: kept, Tags: kept,

Cutting Options - Cut by [number]: 2, Percentage Overlap: 0,

JiaPingwa\_MissingWolves\_2000\_1: Scrubbing Options - Punctuation: kept, Lowercase: off, Digits: kept, Tags: kept,

Cutting Options - Cut by [number]: 2, Percentage Overlap: 0,

JiaPingwa\_MissingWolves\_2000\_2: Scrubbing Options - Punctuation: kept, Lowercase: off, Digits: kept, Tags: kept,

Cutting Options - Cut by [number]: 2, Percentage Overlap: 0,

JiaPingwa\_ShaanxiOpera\_2005\_1: Scrubbing Options - Punctuation: kept, Lowercase: off, Digits: kept, Tags: kept,

Cutting Options - Cut by [number]: 2, Percentage Overlap: 0,

JiaPingwa\_ShaanxiOpera\_2005\_2: Scrubbing Options - Punctuation: kept, Lowercase: off, Digits: kept, Tags: kept,

Cutting Options - Cut by [number]: 2, Percentage Overlap: 0,