Homogeneity of Clusters

After adding punctuation and stop words:
Average accuracy among the two clusters with 100 items each (calculated using
`compute_homogeneity`): 0.771856657398826