
Clusters by Descriptions

1 Using Word2Vec

- We got 10 clusters using Word2Vec method. Clusters according to their themes are following–

Cluster 0: ['hanging', 'holder', 'light']

Cluster 1: ['christmas', 'paper', 'vintage']

Cluster 2: ['bag', 'jumbo', 'lunch']

Cluster 3: ['clock', 'drawer', 'red']

Cluster 4: ['of', 'pack', 'set']

Cluster 5: ['blue', 'pink', 'red']

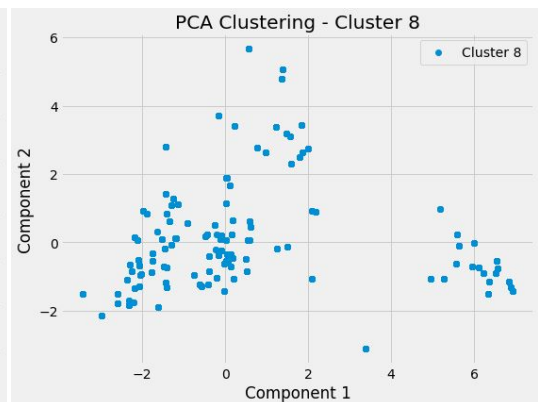
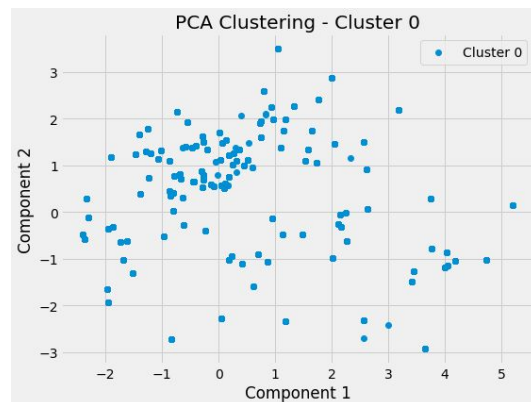
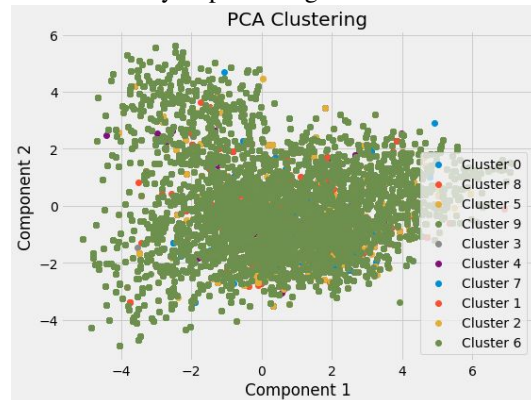
Cluster 6: ['decoration', 'heart', 'white']

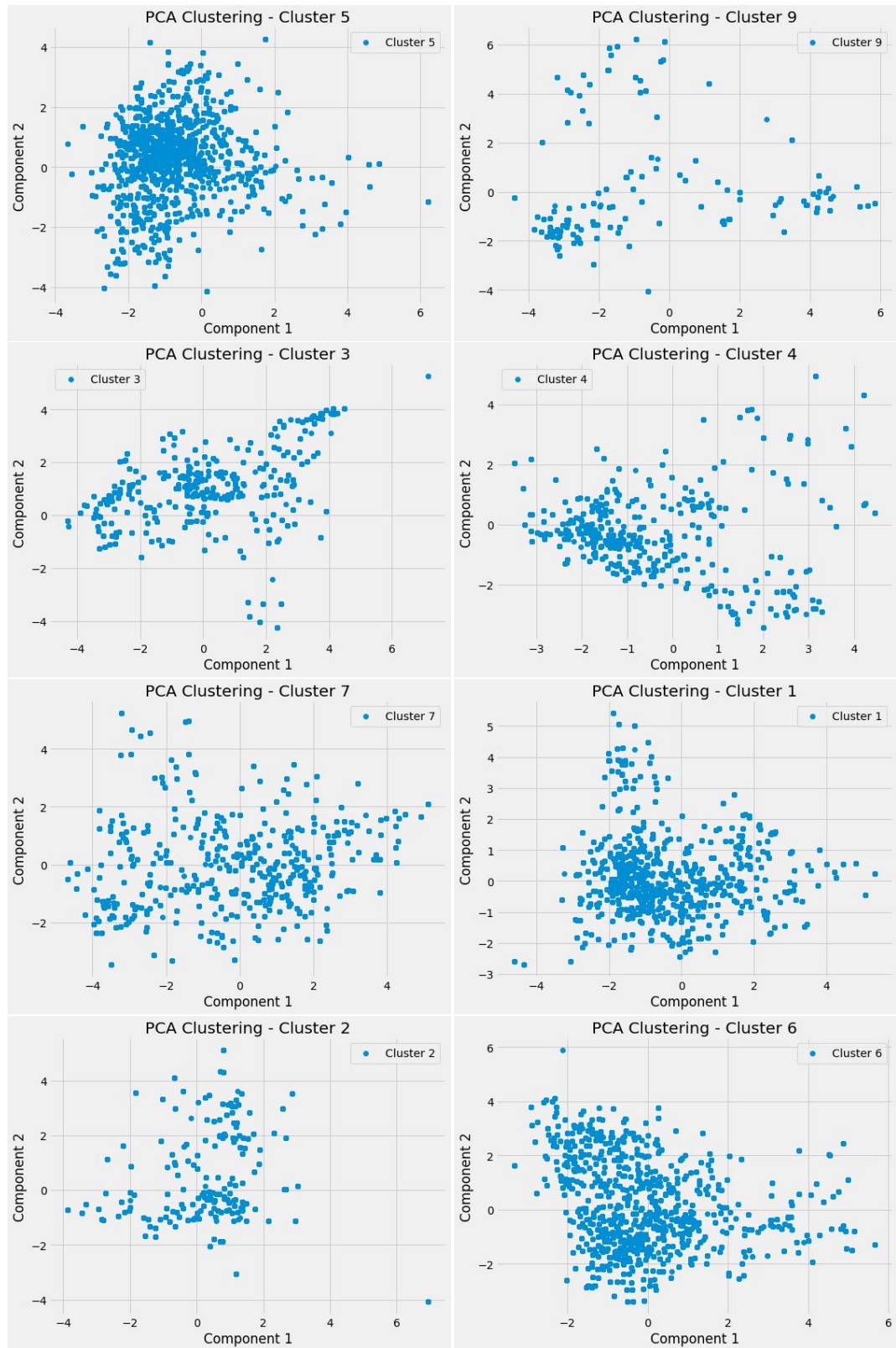
Cluster 7: ['box', 'design', 'retrospot']

Cluster 8: ['home', 'metal', 'sign']

Cluster 9: ['bottle', 'card', 'water']

- The PCA analysis plots we got –





2 Using BERT

- Using BERT we also created 10 clusters. Clusters according to their themes are following–
Cluster 0: of, red, set, box, cake
Cluster 1: bag, set, of, vintage, red
Cluster 2: set, bag, of, blue, lunch
Cluster 3: set, bag, red, of, vintage
Cluster 4: set, of, heart, design, bag
Cluster 5: of, set, design, vintage, bag
Cluster 6: bag, retrospot, jumbo, heart, red
Cluster 7: set, red, retrospot, of, colour
Cluster 8: set, heart, christmas, vintage, of
Cluster 9: set, of, bag, vintage, heart
- Occurances of every cluster is –
Cluster 1: 47329 occurrences
Cluster 9: 43537 occurrences
Cluster 3: 41791 occurrences
Cluster 6: 41636 occurrences
Cluster 5: 41447 occurrences
Cluster 0: 41423 occurrences
Cluster 7: 38196 occurrences
Cluster 8: 37784 occurrences
Cluster 4: 37350 occurrences
Cluster 2: 36336 occurrences