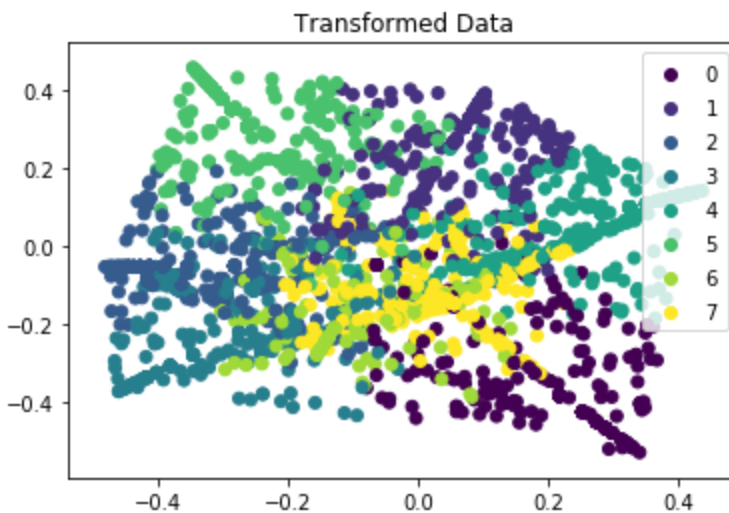


Rex Asabor

Task 1:

['the', 'and', 'this', 'I', 'we', 'me', 'us', 'that']

Task 2:



The transformed data is more unstructured, it shows data points with no easily identifiable bounds between colors and that do not have neat separable clusters.

Task 4

month	February	March	April
tweet_labels			
0	149	122	104
1	174	62	107
2	143	86	83
3	163	125	92
4	77	146	79

Task 5

February Majority: ['ly', 'solidarity', 'possible', '03', 'kind', 'society', 'pakistan'] → "Solidarity could lead to a society kind of like Pakistan in 2003"

March Majority: ['gov', 'possible', 'helping'] → "The Government Possibly Helping"

April Majority: ['trump', 'possible', 'population', 'cure'] → "Trump promotes a possible cure for the population"

The April Majority topic seems to be present in all months.

Most of these make sense, but the february majority topic does not. Testing new stop words or finetuning other parameters may be able to help.