

Evan Coulson
MATH189Z
Worked With: Lauren West

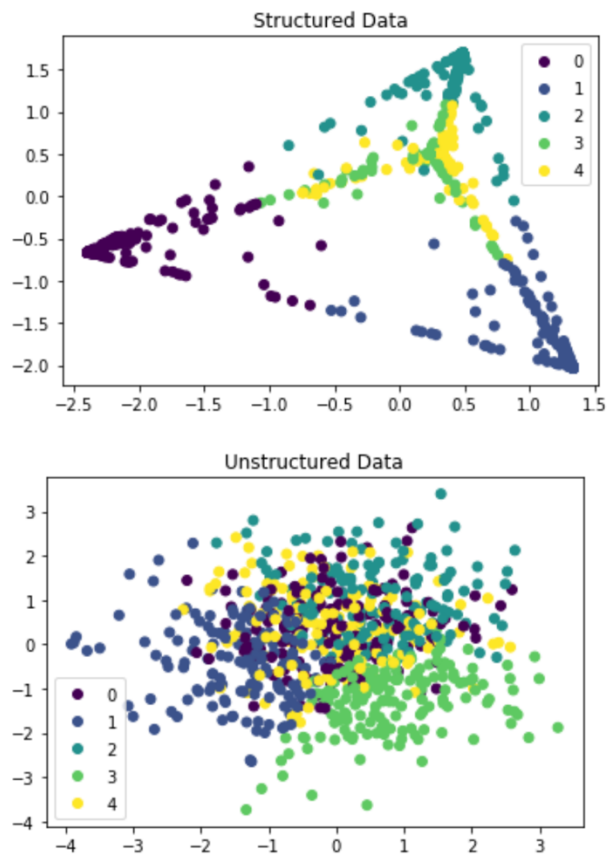
1)

[

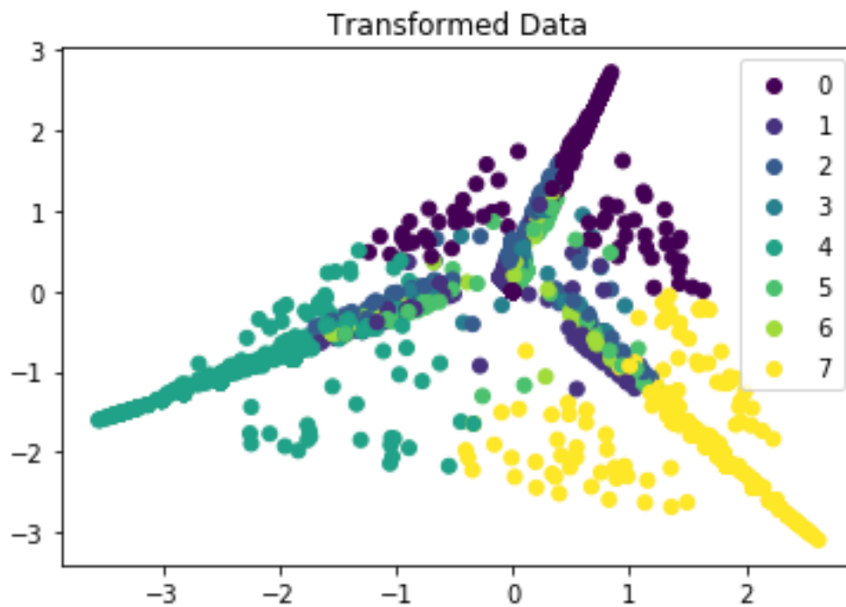
```
'http', 'com', 'net', 'COVID-19', 'corona',  
'said', 'https', 'let', 'covid', '19',  
'coronavirus', '000', 'twitter', 'amp',  
'2020', '04', 'www', 'like', 'just', 've',  
'thank', 'use', 'know', 'says', 'really',  
'covid_19', 'going', 'doesn', 'don', 'covid19'  
'did', 'got', 'll', '10'
```

]

2)

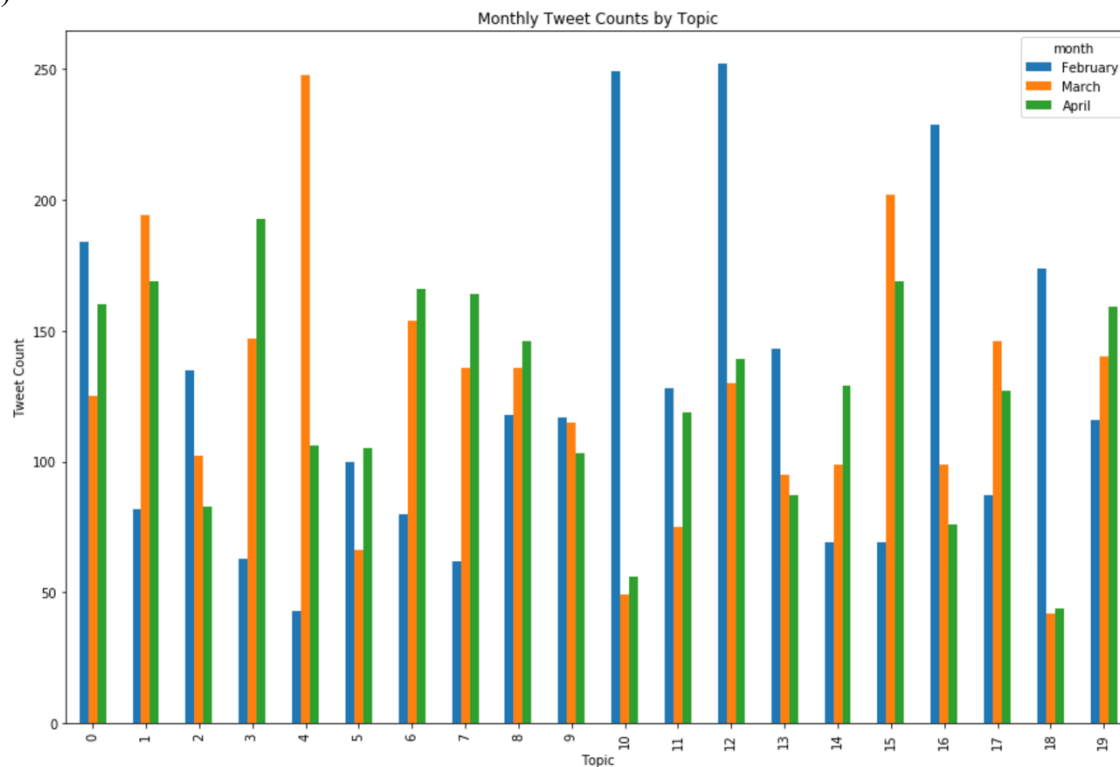


3)



My transformed data has a clear structure with all tweet topics being linearly related in the topic space. The topics also compose an interrelated structure between all the topics asides from topic 0, 4, and 7. Each topic seems to share some commonality between them. Other topics which are at the extreme ends of the structure do not share many similarities.

4)



5)

Present: Topic 0

February: Topic 10, 12

March: Topic 4

April: Topic 3

The most consistent topic throughout the three months was topic 0. This topic consisted of the keywords

['china', 'fight', 'chinese', 'started', 'live', 'outbreak', 'news', 'days', 'amid', 'couple', 'tom', 'message']

This topic seems to refer to the origins of the virus's outbreak and the live news cycle about the virus

The two most present topics, topic 10 and topic 12, in February were very similar, as they had the respective key words of

['hospital', 'died', 'wuhan', 'patients', 'china', 'infected', 'patient', 'chinese', 'novel', 'outbreak', 'city', 'director']

['china', 'covid19', 'people', 'new', 'number', 'deaths', 'reports', 'cases', 'confirmed', 'total', 'death', 'iran']

These keywords refer to Wuhan being the epicenter of the outbreak, the virus spreading to Iran, and the large number of confirmed deaths and cases. They also refer to the hospitals being overrun.

The most common topic in March was topic 4, which had the keywords

['white', 'news', 'positive', 'trump', 'negative', 'breaking', 'realdonaldtrump', 'house', 'tested', 'tests', 'test', 'president']

This topic can be summarized as the white house's response to testing in the United states, particularly directed toward president Trump and his twitter account.

In April the focus shifted quite a bit as Topic 3 was the most common and had the keywords

['time', 'care', 'health', 'staff', 'healthcare', 'doctors', 'hospitals', 'medical', 'workers', 'support', 'minister', 'pay']

Topic 3 focuses more on giving attention to medical staff that are being overrun by the number of cases they are encountering. This topic also encompasses a bit about the current layoffs in America as it captures keywords such as pay and time off.