

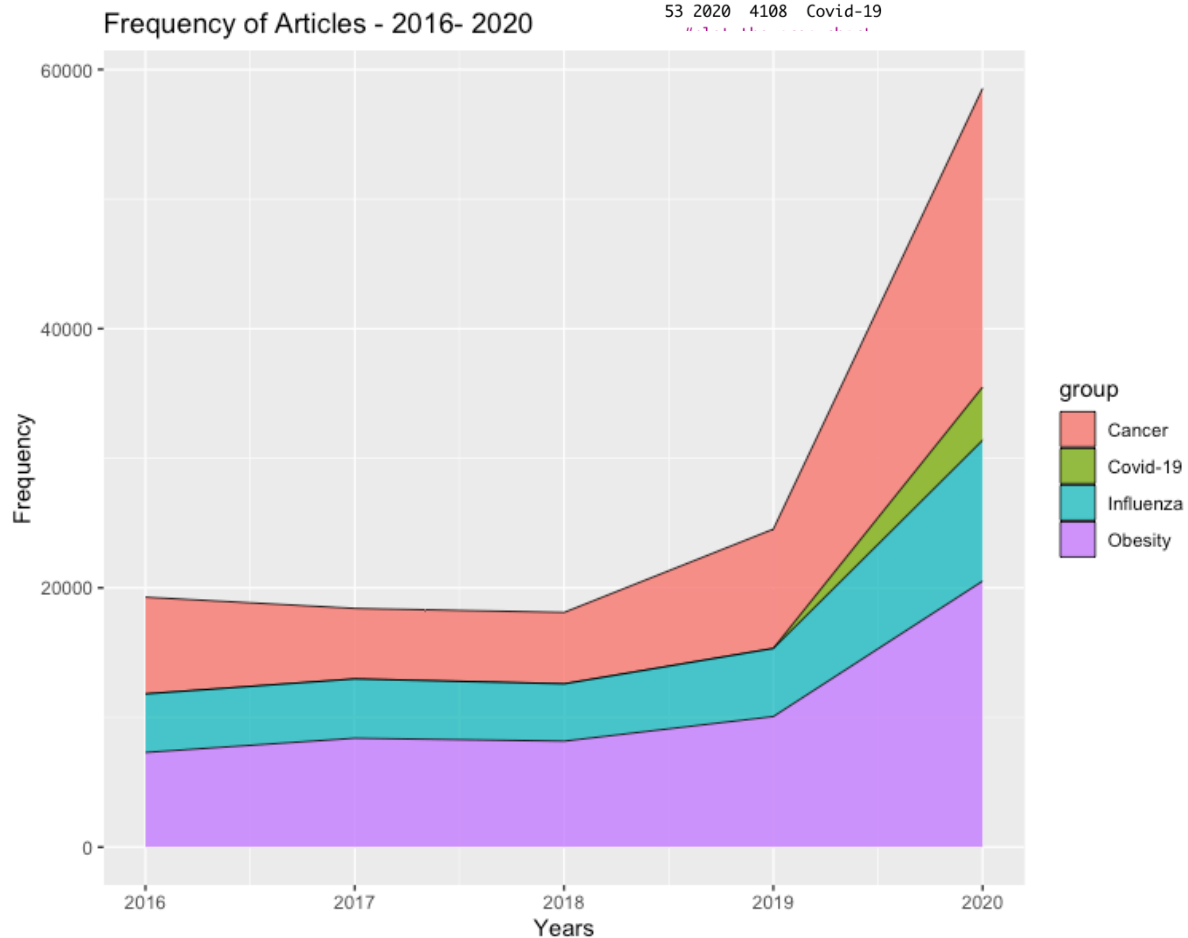
Homework 2 – Maryam Asghari

Downloading five years of scientific articles from PubMed dataset using EasyPubMed library.

1. For each of the following terms: Influenza, Obesity, Cancer, Covid-19 create an area chart and compare their frequencies for each of the following years: 2016, 2017, 2018, 2019, 2020.

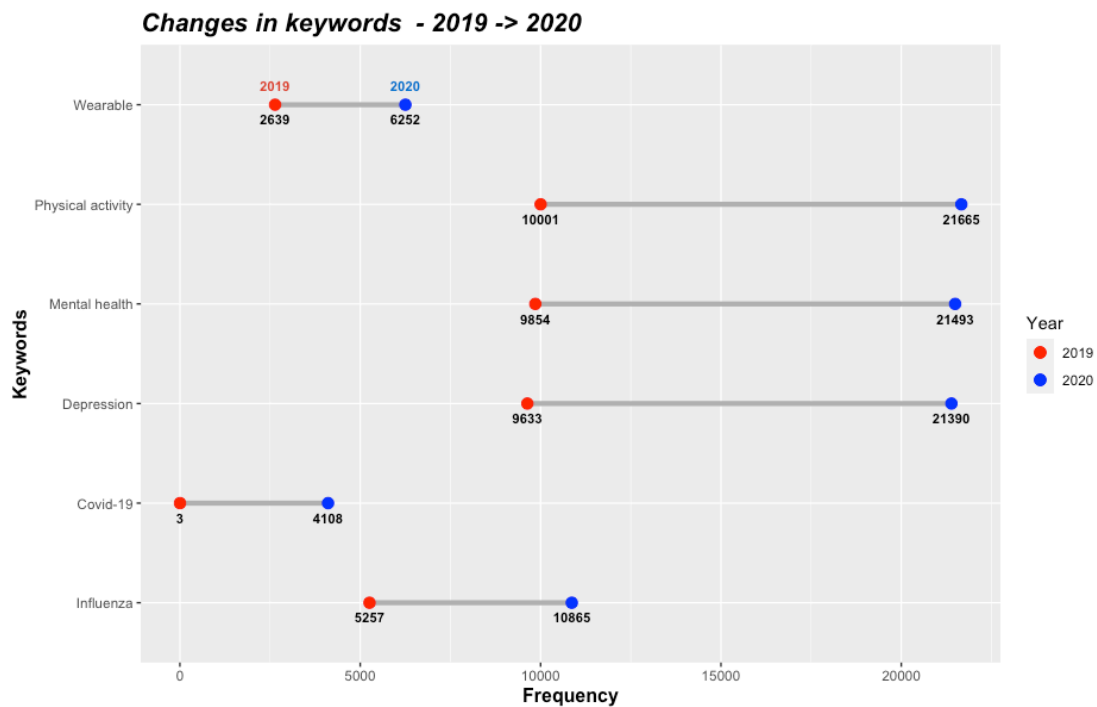
Total number of articles: 138,884

```
> df
  year  Freq  group
2  2016  4515 Influenza
3  2017  4578 Influenza
4  2018  4418 Influenza
5  2019  5257 Influenza
6  2020 10865 Influenza
21 2016  7300 Obesity
31 2017  8396 Obesity
41 2018  8172 Obesity
51 2019 10070 Obesity
61 2020 20522 Obesity
22 2016  7466 Cancer
32 2017  5441 Cancer
42 2018  5513 Cancer
52 2019  9204 Cancer
62 2020 23055 Cancer
1  2016     1 Covid-19
23 2017     0 Covid-19
33 2018     0 Covid-19
43 2019     3 Covid-19
53 2020  4108 Covid-19
```

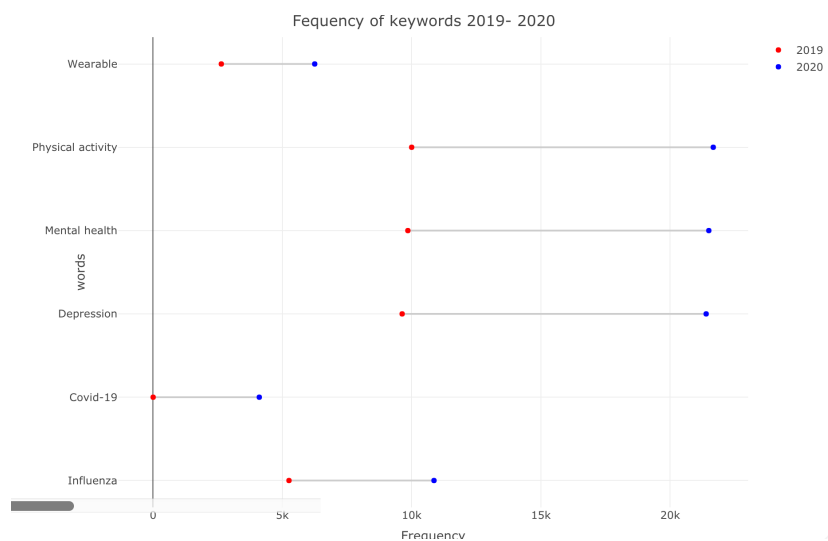


2. Please use a dumbbell-chart and report the changes in following keywords for 2020 and 2019: Influenza, Covid-19, Depression, Mental health, Physical activity, Wearable

```
> data <- data.frame(words, Y_2019, Y_2020,
+ data
+ )
  words Y_2019 Y_2020
1 Influenza  5257 10865
2 Covid-19    3   4108
3 Depression 9633 21390
4 Mental health 9854 21493
5 Physical activity 10001 21665
6 Wearable  2639  6252
```



Plotly : Please check out this in R



3. Get Covid-19 statistics from the city of Massachusetts: design a Choropleth map and visualize (number of confirmed cases) on the map of Boston or MA.

Confirmed Cases Reported =26,867

CATEGORY	NUMBER OF CONFIRMED CASES
County	
Barnstable	502
Berkshire	355
Bristol	1394
Dukes	12
Essex	3413
Franklin	148
Hampden	1694
Hampshire	224
Middlesex	5983
Nantucket	9
Norfolk	2838
Plymouth	2141
Suffolk	5579
Worcester	2128
Unknown	447

Massachusetts Covid Cases

