

Aimee Haas, Sean Fletcher, SJ Franklin

Professor Mansouri

COS 470 - Natural Language Processing - Lab 2

February 3, 2023

Analysis of Word Cloud Results

While the Word Cloud package creates a slightly different result each time the program is run, we can observe that the same set of words seems to take priority in dominating the screen, as can be seen in the image below. Namely, the words “Covid”, “case”, “day”, and “least” are quite prominent. From this information, paired with the secondary words that appear in the Word Cloud such as “countries” and “year”, we can infer that our dataset is most likely working with articles that report on the number of daily COVID-19 cases that have happened over various periods of time in multiple areas. This is further corroborated by the formatting of the files we are working with.

