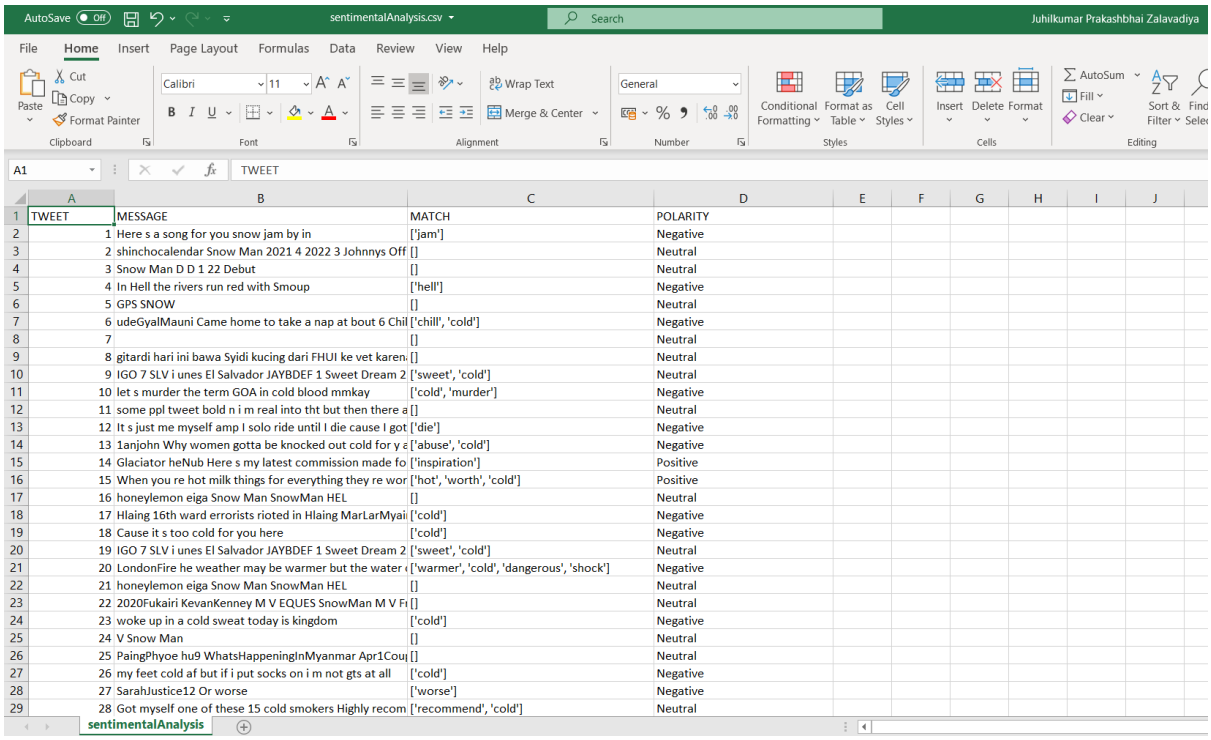
Juhilkumar P. Zalavadiya
B00872280

Problem #1: Sentimental Analysis

I have fetched live tweet texts from Twitter stream API and cleaned URLs and other noisy data to evaluate it easily. I have written a script to find positive and negative words from a tweet based on the predefined positive and negative word txt files [1]. After finding the words, I have checked the polarity of the tweet based on the number of positive and negative words available in the text.

- If positive words are greater than negative ones, I have set polarity to “Positive”.
- If negative words are greater than positive ones, I have set polarity to “Negative”.
- If a tweet has the same number of positive and negative words, then I have set polarity to “Neutral”.

ScreenShots of the CSV file:



	A	B	C	D
1	TWEET	MESSAGE	MATCH	POLARITY
2	1	Here s a song for you snow jam by in	['jam']	Negative
3	2	shinchocalendar Snow Man 2021 4 2022 3 Johnnys Off	[]	Neutral
4	3	Snow Man D D 1 22 Debut	[]	Neutral
5	4	In Hell the rivers run red with Smoup	['hell']	Negative
6	5	GPS SNOW	[]	Neutral
7	6	udeGyalMauni Came home to take a nap at bout 6 Chil	['chill', 'cold']	Negative
8	7		[]	Neutral
9	8	gitardi hari ini bawa Syidi kucing dari FHUI ke vet karen	[]	Neutral
10	9	IGO 7 SLV i unes El Salvador JAYBDEF 1 Sweet Dream 2	['sweet', 'cold']	Neutral
11	10	let s murder the term GOA in cold blood mmkay	['cold', 'murder']	Negative
12	11	some ppl tweet bold n i m real into tht but then there a	[]	Neutral
13	12	It s just me myself amp I solo ride until I die cause I got	['die']	Negative
14	13	1anjohn Why women gotta be knocked out cold for y a	['abuse', 'cold']	Negative
15	14	Glaciator heNub Here s my latest commission made fo	['inspiration']	Positive
16	15	When you re hot milk things for everything they re wor	['hot', 'worth', 'cold']	Positive
17	16	honeylemon eiga Snow Man SnowMan HEL	[]	Neutral
18	17	Hlaing 16th ward errorists rioted in Hlaing MarLarMyai	['cold']	Negative
19	18	Cause it s too cold for you here	['cold']	Negative
20	19	IGO 7 SLV i unes El Salvador JAYBDEF 1 Sweet Dream 2	['sweet', 'cold']	Neutral
21	20	LondonFire he weather may be warmer but the water	['warmer', 'cold', 'dangerous', 'shock']	Negative
22	21	honeylemon eiga Snow Man SnowMan HEL	[]	Neutral
23	22	2020Fukairi KevanKenney M V EQUES SnowMan M V Fi	[]	Neutral
24	23	woke up in a cold sweat today is kingdom	['cold']	Negative
25	24	V Snow Man	[]	Neutral
26	25	PaingPhyoe hu9 WhatsHappeningInMyanmar Apr1Cou	[]	Neutral
27	26	my feet cold af but if i put socks on i m not gts at all	['cold']	Negative
28	27	SarahJustice12 Or worse	['worse']	Negative
29	28	Got myself one of these 15 cold smokers Highly recom	['recommend', 'cold']	Neutral

Figure 1 ScreenShot of sentimental analysis

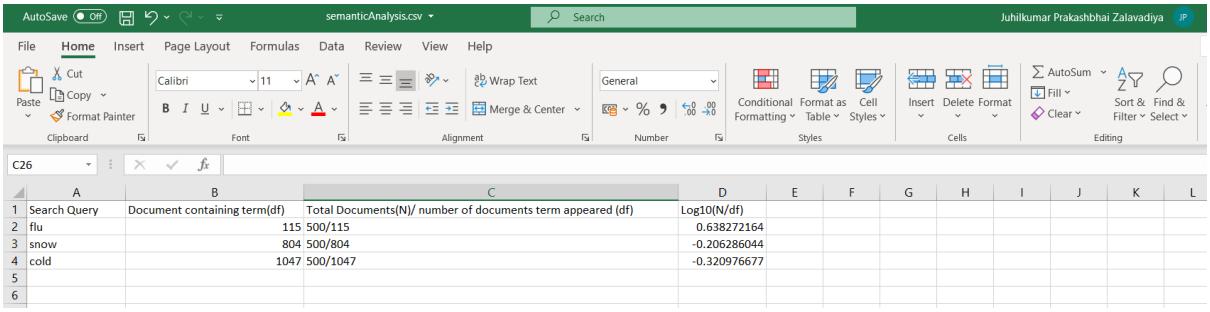
Juhilkumar P. Zalavadiya
B00872280

Problem #2: Semantic Analysis

In the semantic analysis, I have used previously stored cleaned tweets from the MongoDB database. I have also attached the screenshots of the outputs in this document and uploaded the scripts in GitLab.

ScreenShots of the output:

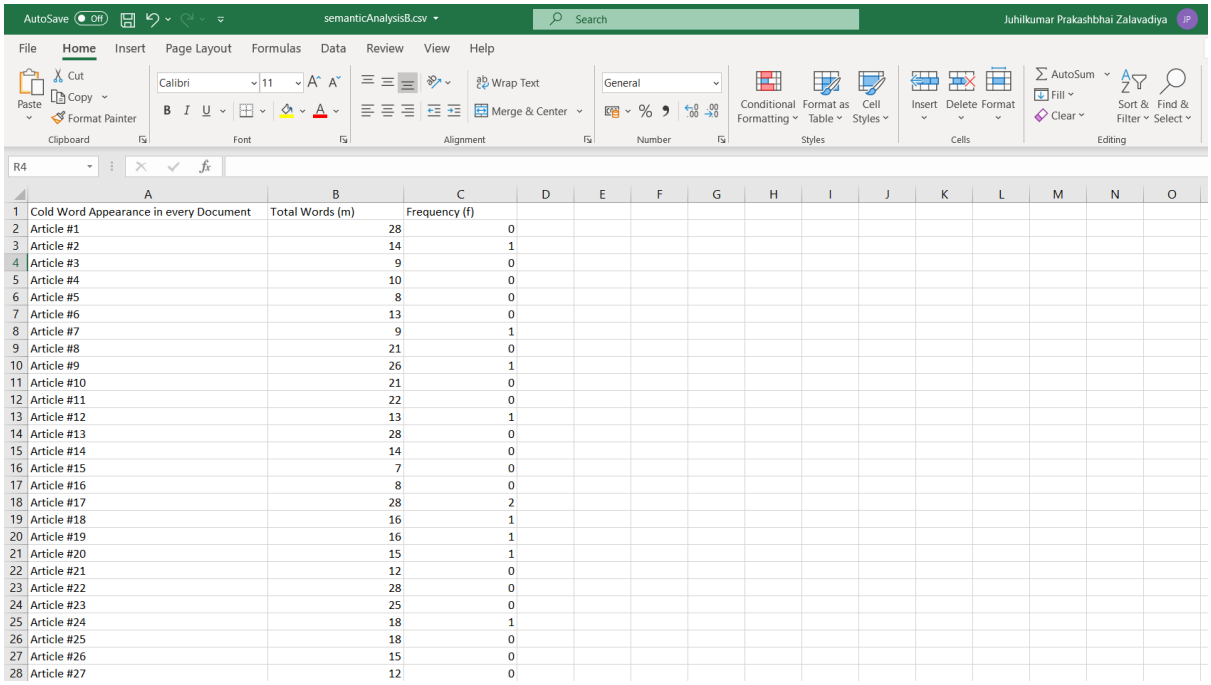
a) Term frequency-inverse document frequency



	A	B	C	D
1	Search Query	Document containing term(df)	Total Documents(N)/ number of documents term appeared (df)	Log10(N/df)
2	flu	115	500/115	0.638272164
3	snow	804	500/804	-0.206286044
4	cold	1047	500/1047	-0.320976677
5				
6				

Figure 2 term frequency of the required words

b) frequency count of the word “cold” per document



	A	B	C
1	Cold Word Appearance in every Document	Total Words (m)	Frequency (f)
2	Article #1	28	0
3	Article #2	14	1
4	Article #3	9	0
5	Article #4	10	0
6	Article #5	8	0
7	Article #6	13	0
8	Article #7	9	1
9	Article #8	21	0
10	Article #9	26	1
11	Article #10	21	0
12	Article #11	22	0
13	Article #12	13	1
14	Article #13	28	0
15	Article #14	14	0
16	Article #15	7	0
17	Article #16	8	0
18	Article #17	28	2
19	Article #18	16	1
20	Article #19	16	1
21	Article #20	15	1
22	Article #21	12	0
23	Article #22	28	0
24	Article #23	25	0
25	Article #24	18	1
26	Article #25	18	0
27	Article #26	15	0
28	Article #27	12	0

Figure 3 ScreenShot of frequency count of the total words with the frequency of word “cold”

Juhilkumar P. Zalavadiya
B00872280

3) The highest relative frequency of the word “cold”

```
C:\Users\lenovo\PycharmProjects\SementicAnalysisPart-c\venv\Scripts\python.exe C:/Users/lenovo/PycharmProjects/SementicAnalysisPart-c/main.py
Artical with highest relative frequency : Article #51
Total Words (m): 30
Frequency (f): 3
Relative frequency : 0.1

Process finished with exit code 0
```

Figure 4 ScreenShot of the highest relative frequency of word “cold”

Juhilkumar P. Zalavadiya
B00872280

References:

[1] “Positive words and Negative words”, cs.uic.edu, 2021. [Online]. Available: https://www.cs.uic.edu/~liub/FBS/sentiment-analysis.html [Accessed: 05-April-2021]