SENTIMENT ANALYSIS IN R
ASSIGNMENT SUBMISSION REPORT

NAME: VIDYA SAGAR. R

SOLUTION APPROACH:

1. First step - Use preprocessing python file, in that using Youtube API, comments from most watched videos on climate change are extracted and result_df is the input dataset

2. Second Step - Use updated Sentiment Analysis R script, to do the sentiment analysis of the comments. This sentiment analysis is based on the word cloud technique and every preprocessing step of text is implemented, each step explained in the code. At the end of the code,plots of output summary for every videos is generated

NOTE:

- Using result_df as input in the Sentiment Analysis Rscript, an output.csv file is generated. output.csv file contains the "text" and "sentiment" denotes the sentiment of the comment

- The maximum input data size considered is 25000 data records to avoid RAM exhausted error. So for any video , the maximum comment limit is considered as 25000 to process the analysis, even if the actual data have more than 35000 records

Video 1 : https://www.youtube.com/watch?v=VbiRNT_gWUQ&ab_channel=InsiderScience

Dataset:

df1

Out[2]:

	author	updated_at	like_count	text	public
0	@JustACarnotaurus	2023-12-27T22:03:58Z	0	Florida be like:😊🥺😊	True
1	@QwartzBeat	2023-12-27T19:21:34Z	0	While the question asks about how Earth would ...	True
2	@QwartzBeat	2023-12-27T18:15:02Z	0	The Intergovernmental Panel on Climate Change ...	True
3	@QwartzBeat	2023-12-27T18:14:26Z	0	If all the ice on Earth were to melt, sea leve...	True
4	@draghiceanu2006	2023-12-27T15:28:40Z	0	Do melted ice cubes over run the cup?	True
...	...	...	...	...	...
17667	@daviddaugherty2144	2015-03-04T03:45:32Z	0	We should build a dam cutting off the Gulf of ...	True
17668	@sensorieskin7367	2015-03-01T20:21:57Z	0	And tought it would be a big deal. Plenty land...	True
17669	@BrandonSLockey	2015-02-26T18:03:44Z	0	Scary! Music helped but still.	True
17670	@rarier	2015-02-26T01:51:39Z	0	Yay!! Valdosta sinks and my farm becomes expen...	True
17671	@bonurb1	2015-02-18T19:35:18Z	0	I have been vacationing in the Keys in Florida...	True

17672 rows × 5 columns

Sentiment output:

Dataset preparation.R ×	Sentiment_Analysis(Calculus Carbo... * ×	df1 ×	

Filter

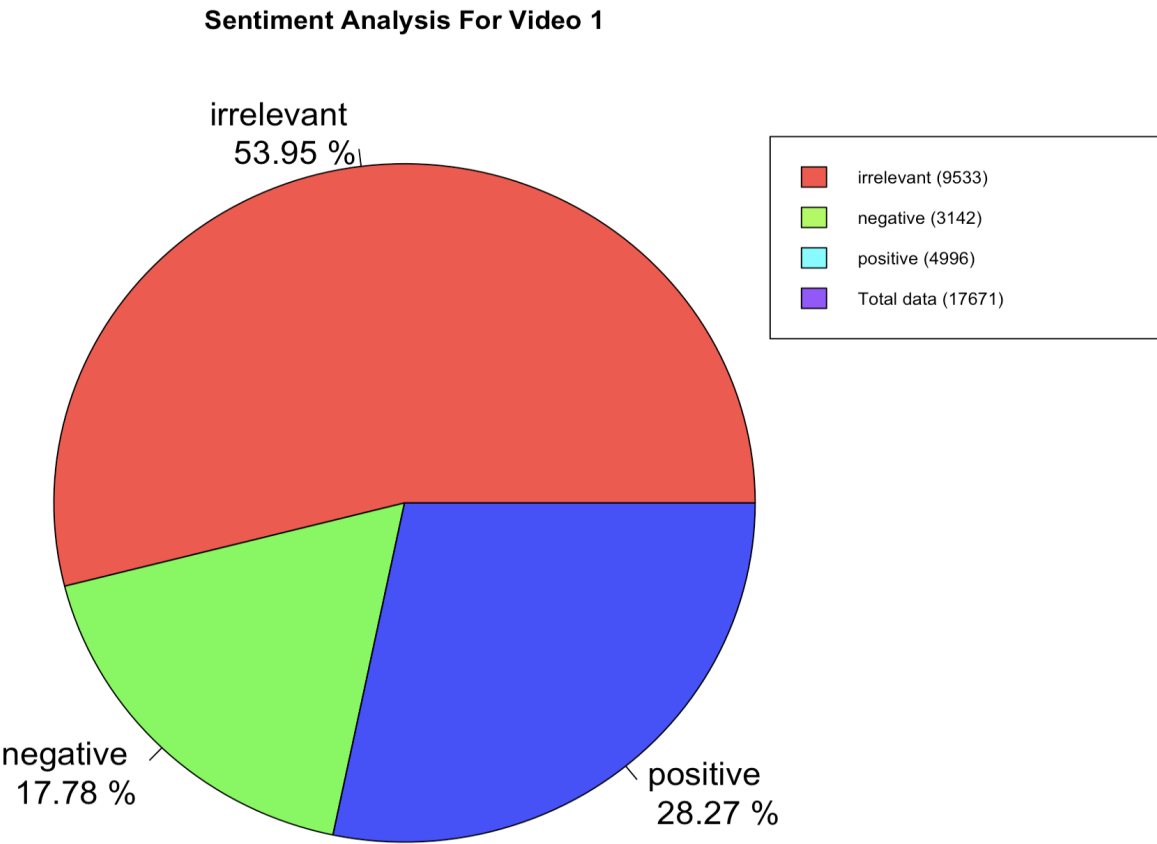|	NA.	text	sentiment
1	0	Florida be like:😊🥺😊	irrelevant
2	1	While the question asks about how Earth would look i...	positive
3	2	The Intergovernmental Panel on Climate Change (IPC...	positive
4	3	If all the ice on Earth were to melt, sea levels would ri...	positive
5	4	Do melted ice cubes over run the cup?	irrelevant
6	5	eh that's fine	irrelevant
7	6	CLIMATE CHANGE... Will only be a problem if and whe...	negative
8	7	It's not so bad, most ice seems to be beneath the sea ...	positive
9	8	Who was beating those drums?	negative
10	9	Excellent! Canada won't be affected by sea level rise! ...	positive
11	10	All ice has melted on earth many times in ancient pas...	irrelevant
12	11	I have a hard time believing there would be zero impa...	positive
13	12	Dovoljno je da ne gledate ovaj video jer bi u ovom sce...	irrelevant
14	13	But won't global warming also evaporate a lot of that ...	positive
15	14	Ma perche pensate solo a voi stessi??? Al vostro paese...	positive

Showing 1 to 16 of 17,671 entries, 3 total columns

Output Summary for Video 1:

Sentiment Analysis For Video 1

irrelevant
53.95 %

🟥	irrelevant (9533)
🟩	negative (3142)
🟦	positive (4996)
🟪	Total data (17671)

negative
17.78 %

positive
28.27 %

Video2

https://www.youtube.com/watch?v=yiw6_JakZFc&ab_channel=Kurzgesagt%E2%80%93InaNutshell

Video 2 Dataset:

df2

Out[3]:

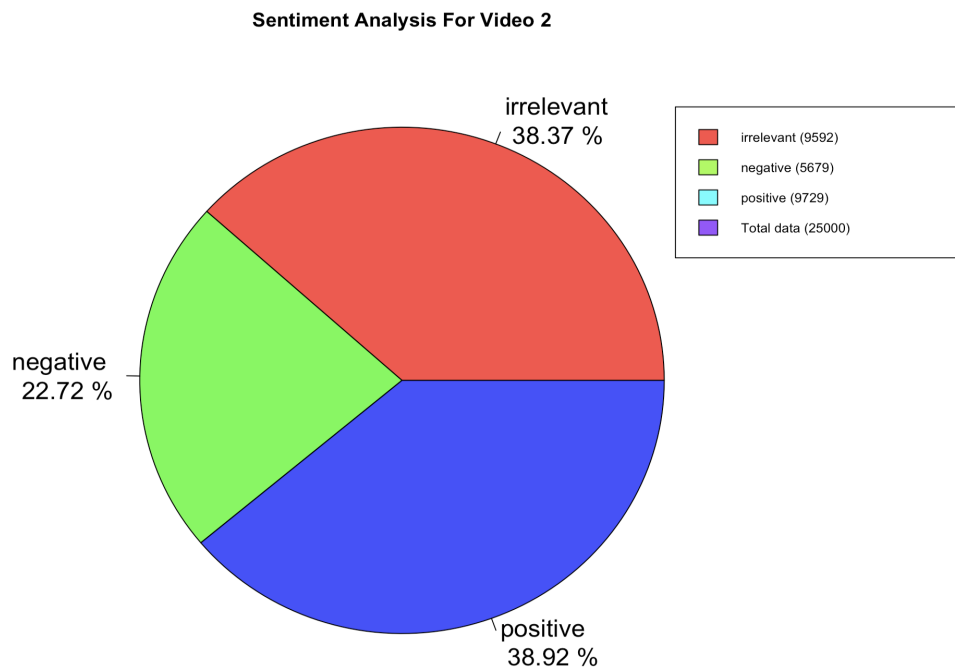|	author	updated_at	like_count	text	public
0	@kurzgesagt	2021-09-22T12:27:33Z	12586	You want to learn more about science? Check ou...	True
1	@maple5998	2023-12-26T08:25:23Z	1	Conclusion : Human Extinction is only way to f...	True
2	@samlair3342	2023-12-23T17:33:53Z	0	The good news about global warming is that it'...	True
3	@InnocentBread-iw3pk	2023-12-22T16:12:53Z	0	In other words mathew 6:24.holds water ❤️\nOnly...	True
4	@Dj-Parralax-Aphelion	2023-12-21T12:08:10Z	0	Companies could use Hydropower functional tech...	True
...	...	...	...	...	...
35073	@j1o970	2021-09-22T13:59:39Z	0	What	True
35074	@InfinityGaming420	2021-09-22T13:59:39Z	0	V	True
35075	@marcorogialli7827	2021-09-22T13:59:38Z	3	Hi	True
35076	@Yoel202	2021-09-22T13:59:36Z	72	E	True
35077	@kurzgesagt	2021-09-22T12:27:33Z	12586	You want to learn more about science? Check ou...	True

35078 rows × 5 columns

NOTE: Due to RAM constraint, input data is reduced to top 25000 entries of df2

Sentiment output:

NA.	text	sentiment
17673	17672	You want to learn more about science? Check out our ...	positive
17674	17673	Conclusion : Human Extinction is only way to fix Clim...	irrelevant
17675	17674	The good news about global warming is that it's not ...	positive
17676	17675	In other words mathew 6:24.holds water ❤ Only one ...	negative
17677	17676	Companies could use Hydropower functional technol...	irrelevant
17678	17677	Video is sponsored by the personal blog of Bill Gates....	irrelevant
17679	17678	Climate hoax.	negative
17680	17679	Another great reason to go vegan!	positive
17681	17680	FIX IT??? I didn't know it was broken.	negative
17682	17681	Z	irrelevant
17683	17682	Y	irrelevant
17684	17683	X	irrelevant
17685	17684	W	irrelevant
17686	17685	V	irrelevant
17687	17686	U	irrelevant

Showing 1 to 16 of 25,000 entries, 3 total columns

<u>Output Summary for Video 2:</u>

Sentiment Analysis For Video 2

irrelevant
38.37 %

negative
22.72 %

positive
38.92 %

- irrelevant (9592)
- negative (5679)
- positive (9729)
- Total data (25000)

Video 3 : https://www.youtube.com/watch?v=0pJ1oD-OXZo&ab_channel=BLACKPINK

Video 3 Dataset:

df3

Out[4]:

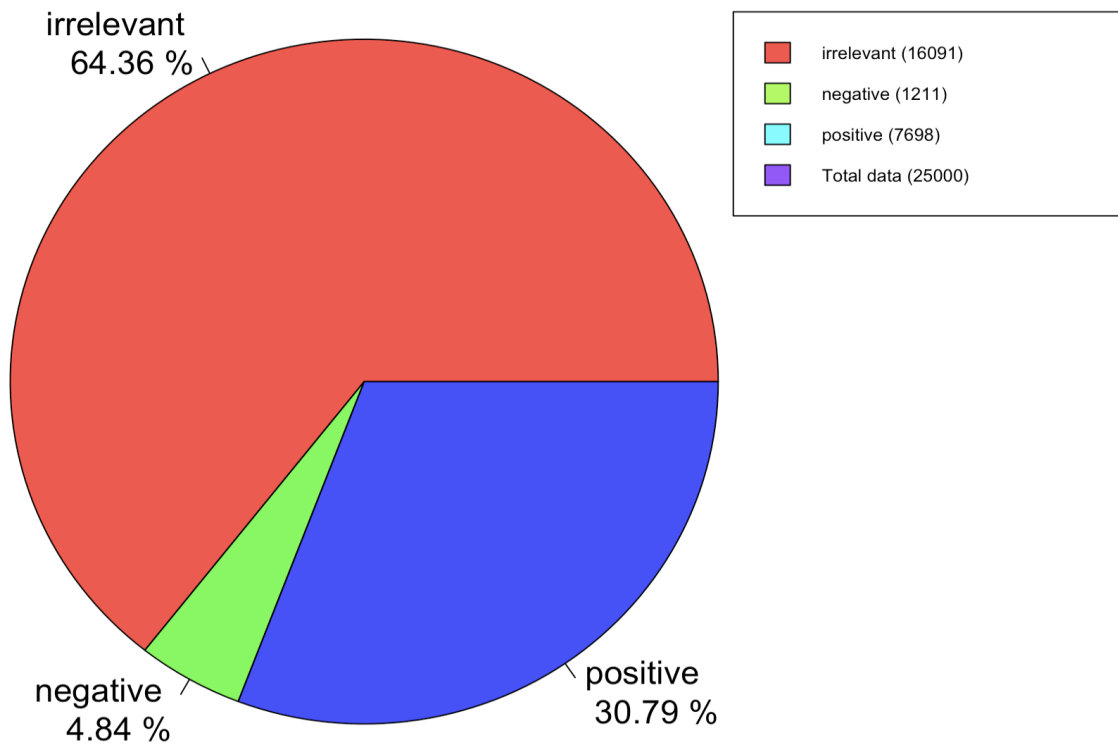|	author	updated_at	like_count	text	public
0	@SitiAisyah-kb2ne	2023-12-22T04:14:52Z	0	omgg jisoo??😭❤️	True
1	@Iam-li4eu	2023-12-07T15:34:33Z	0	Congratulations for being awarded the honorary...	True
2	@user-gr8tg8fr2x	2023-12-04T12:45:12Z	0	Please don't leave us 🥺🥺🥺	True
3	@iamjosh7290	2023-11-23T19:41:23Z	1	You 4 aren't exactly helping in any real way	True
4	@dearines	2023-11-23T09:40:40Z	2	That's nice and all but can anyone unironicall...	True
...	...	...	...	...	...
35585	@saragomaa4852	2021-11-03T02:00:20Z	2	🙌🖤🖤	True
35586	@maicaelissetuvillara7811	2021-11-03T02:00:20Z	1	Ang aga ko	True
35587	@kristinerobin25	2021-11-03T02:00:20Z	3	FIRST!	True
35588	@buskosin	2021-11-03T02:00:19Z	4	Tempranoo	True
35589	@aungkham3149	2021-11-03T02:00:16Z	14	Hi	True

35590 rows × 5 columns

NOTE: Due to RAM constraint, input data is reduced to top 25000 entries of df3

Sentiment output:

Filter

	NA.	text	sentiment
52754	52750	omgg jisoo??😂❤	irrelevant
52755	52751	Congratulations for being awarded the honorary MBEs…	positive
52756	52752	Please don't leave us 🥺🥺🥺	negative
52757	52753	You 4 aren't exactly helping in any real way	positive
52758	52754	That's nice and all but can anyone unironically tell me…	positive
52759	52755	What a load of bs.. the world target of 1.5° has been r…	irrelevant
52760	52756	And then Lisa took a private jet to go to her boyfriend…	positive
52761	52757	wish you could stop being hypocrites…	irrelevant
52762	52758	Anyone noticed that Jisoo is wearing very similar dres…	irrelevant
52763	52759	Realy Blacpink ı love	positive
52764	52760	This was when I was first beginning to stan Blackpink …	irrelevant
52765	52761	Fácilmente aprendí a diferenciar estas chicas por el as…	irrelevant
52766	52762	❤❤❤❤❤❤❤❤❤❤❤❤❤❤❤❤❤❤❤❤❤❤❤…	irrelevant
52767	52763	لطافت بلاكبينك 🥺💖	irrelevant
52768	52764	Argien der 2023😂	irrelevant

Showing 1 to 15 of 25,000 entries, 3 total columns

Output Summary for Video 3:

Sentiment Analysis For Video 3

irrelevant
64.36 %

negative
4.84 %

positive
30.79 %

■	irrelevant (16091)
■	negative (1211)
■	positive (7698)
■	Total data (25000)

Video 4 : https://www.youtube.com/watch?v=e6rglsLy1Ys&ab_channel=NationalGeographic

Video 4 Dataset:

```
df4
```

Out[5]:

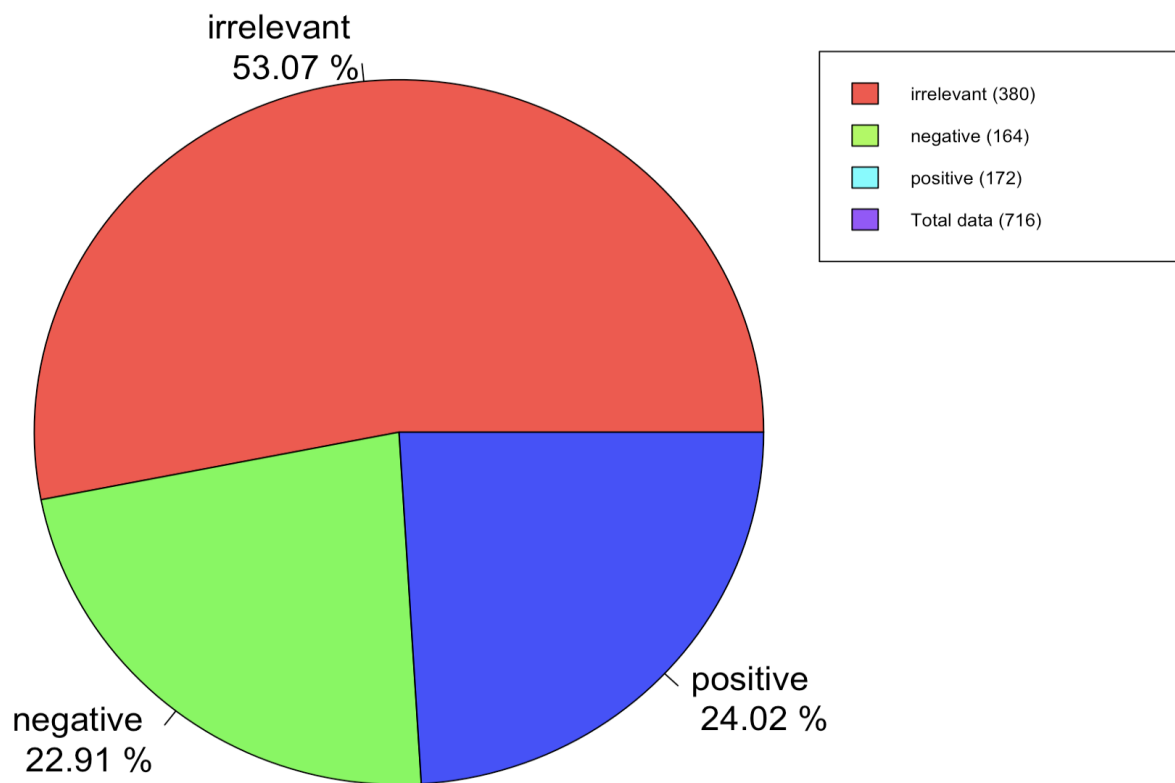|	author	updated_at	like_count	text	public
0	@Draakie100	2023-12-27T14:09:44Z	0	In most "1st world" countries we live in a soc...	True
1	@finaledits11	2023-12-25T12:41:07Z	0	very good	True
2	@RAS_SI_Strategy_Motivation	2023-12-17T13:22:23Z	0	Very reliable 👍	True
3	@user-qp8we3ss9x	2023-12-02T07:43:27Z	0	Great informative video, but I have to reitera...	True
4	@user-mu3iy8fq3d	2023-11-27T08:46:25Z	0	The Paris Agreement stands as a testament to g...	True
...	...	...	...	...	...
711	@ashasshussain7269	2017-10-16T12:10:45Z	25	Hope that pollution will decrease on earth whe...	True
712	@gurjotvirk5257	2017-10-16T12:08:49Z	19	We all know that air pollution is very importa...	True
713	@deep4630	2017-10-16T12:08:01Z	32	We should reduce all type pollution...	True
714	@artgirl1578	2017-10-16T12:04:19Z	13	Hpoe the earth can be better place to breath ! 🙏	True
715	@Kenny-ii1kt	2017-10-16T12:02:26Z	3	Oh n	True

716 rows × 5 columns

Sentiment output:

	Dataset preparation.R ×		Sentiment_Analysis(Calculus Carbo... * ×		df4_result ×				

⇐ ⇒ | 🔃 | ▽ Filter | | 🔍

	NA.	text	sentiment
88344	88340	In most "1st world" countries we live in a society that ...	positive
88345	88341	very good	positive
88346	88342	Very reliable 👍	positive
88347	88343	Great informative video, but I have to reiterate a previ...	positive
88348	88344	The Paris Agreement stands as a testament to global ...	positive
88349	88345	The government's need to ban pencils if its pollution i...	negative
88350	88346	Big drones in the airzone with big fans could give cle...	negative
88351	88347	cloud seeding or artificial raining is solution..must try...	positive
88352	88348	Humans wearing face mask to protect your air polluted	positive
88353	88349	Air pollution is really bad now, I mean, when you ope...	negative
88354	88350	Blame china and some of America	negative
88355	88351	My students are using a small part of the video for th...	negative
88356	88352	I'm from VietNam. Thanks for giving me the useful ne...	positive
88357	88353	Increasing electric cars and solar power generation is ...	positive
88358	88354	New how small tiny parts getting into our blood strea...	positive

Showing 1 to 16 of 716 entries, 3 total columns

Output Summary for Video 4:

Sentiment Analysis For Video 4

irrelevant
53.07 %

■	irrelevant (380)
■	negative (164)
■	positive (172)
■	Total data (716)

positive
24.02 %

negative
22.91 %

Video 5 : https://www.youtube.com/watch?v=TZ0j6kr4ZJ0&ab_channel=DWDocumentary

Video 5 Dataset:

```
df5
```

Out[6]:

	author	updated_at	like_count	text	public
0	@jiggyquinn5635	2023-12-28T00:58:09Z	0	This is not the first place to be swallowed by...	True
1	@erikpolmann4308	2023-12-20T20:07:41Z	0	🙏🙏🤷	True
2	@jojobaker1764	2023-12-19T22:18:16Z	0	Well first off if this was caused by global wa...	True
3	@triplemoon9563	2023-12-19T16:25:41Z	0	Idea… why not just move the house back?	True
4	@julieb3808	2023-12-19T04:43:27Z	0	While taking away the Carl to help them live t...	True
...	...	...	...	...	...
6364	@jan-br3pl	2017-11-08T21:23:38Z	392	I really enjoyed this video. This is the first...	True
6365	@geroestetumor	2017-11-08T21:12:51Z	9	Do you expect me, as a European, to care about...	True
6366	@minutesagoedited9761	2017-11-08T20:41:31Z	39	Thing is, even I could design and build a wate...	True
6367	@SDTNionios	2017-11-08T20:14:21Z	137	Great! Please do keep uploading such documenta...	True
6368	@minutesagoedited9761	2017-11-08T20:13:08Z	13	Grow some MJ. Send to me for $200/oz? lol my p...	True

6369 rows × 5 columns

Sentiment output:

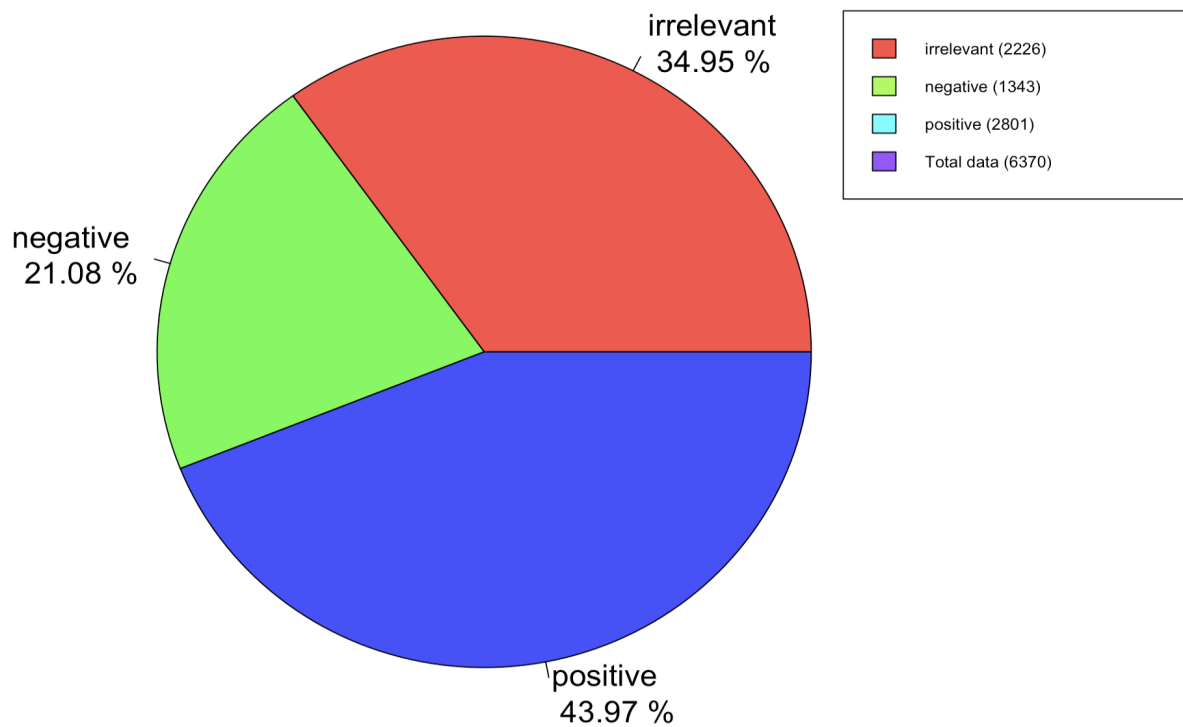|	Dataset preparation.R ×		Sentiment_Analysis(Calculus Carbo... * ×		df5_result ×			
	← →	🔁	▼ Filter				🔍	

	NA.	text	sentiment
89060	89056	This is not the first place to be swallowed by the sea. ...	positive
89061	89057	🙏🙏🤷	irrelevant
89062	89058	Well first off if this was caused by global warming eve...	irrelevant
89063	89059	Idea… why not just move the house back?	irrelevant
89064	89060	While taking away the Carl to help them live they're ta...	irrelevant
89065	89061	Propaganda, western counties? The whole world burn...	negative
89066	89062	This documentary would have been so much better if ...	positive
89067	89063	It's hard to watch this nonsense with people pretendi...	negative
89068	89064	Give me a break… bill gates and Obama and Zuckerb...	irrelevant
89069	89065	More globalist propaganda	negative
89070	89066	Where's the follow-up?	irrelevant
89071	89067	The interstellar music. -_-	positive
89072	89068	with all that's said, these young island girls should k...	positive
89073	89069	Wait, so these guys blame us while they make 5-6 kid...	positive
89074	89070	Stop blamming the climate, you live on an attoll, they ...	irrelevant

Showing 1 to 16 of 6,370 entries, 3 total columns

Output Summary for Video 5:

Sentiment Analysis For Video 5

irrelevant
34.95 %

■	irrelevant (2226)
■	negative (1343)
■	positive (2801)
■	Total data (6370)

negative
21.08 %

positive
43.97 %

P.S:
FOR CODE FILES AND DATASETS REFER TO PROVIDED GITHUB LINK,

GITHUB LINK:

https://github.com/VidyaSagar-RVSR/Sentiment-Analysis-Calculus-Carbon