

Cruv TextRank Summarization

January 24, 2023

```
[1]: import numpy as np
import pandas as pd
import warnings
warnings.filterwarnings('ignore')
```

1 Importing the provided CSV file.

```
[2]: df1 = pd.read_csv("./news.csv")
df1.head()
```

```
[2]:
```

	title \
0	BTS: RM is reminded of Bon Voyage as he travel...
1	RM recalls wondering if he 'made right decisio...
2	BTS: J-Hope and RM go bonkers at Billie Eilish...
3	BTS: J-Hope proudly states he raised Jungkook,...
4	BTS: Jin aka Kim Seokjin takes us through the ...

	content	published_at \
0	After reaching his hotel in the city, RM revea...	2022-07-30T07:00:00Z
1	RM aka Kim Namjoon was the first member to joi...	2022-12-22T15:57:55Z
2	Billie Eilish's concert was held in Seoul, Sou...	2022-08-16T07:00:00Z
3	BTS ARMY y'all would be missing the members a ...	2022-12-18T13:08:40Z
4	BTS member Kim Seokjin aka Jin has the capacit...	2022-11-21T08:00:00Z

	source	topic
0	2	13
1	2	13
2	1	7
3	1	7
4	1	8

```
[3]: df1.shape
```

```
[3]: (810, 5)
```

```
[4]: df1['content'].iloc[401]
```

[4]: 'BTS consists of members Jin, Suga, J-Hope, RM, Jimin, V, and Jungkook—co-writes and co-produces much of their own material. Originally a hip hop group, their musical style has evolved to incorporate a wide range of genres; their lyrics have often discussed mental health, the troubles of school-age youth and coming of age, loss, the journey towards self-love, and individualism. Their work also frequently references literature, philosophy and psychological concepts, and includes an alternate universe storyline.\n\nBTS in the beginning:\n\nAfter launching in 2013 with their single album 2 Cool 4 Skool, BTS respectively released their first Korean-language studio album, Dark & Wild, and Japanese-language studio album, Wake Up, in 2014. The group's second Korean studio album, Wings (2016), was their first to sell one million copies in South Korea. By 2017, BTS had crossed into the global music market, leading the Korean wave into the United States and breaking several sales records. They became the first Korean ensemble to receive a Gold certification from the Recording Industry Association of America (RIAA) for their single Mic Drop, as well as the first act from South Korea to top the Billboard 200 with their studio album Love Yourself: Tear (2018).\n\nBTS's achievements:\n\nBTS became one of the few groups since the Beatles in 1966–1968 with four US number-one albums in less than two years, and Love Yourself: Answer (2018) was the first Korean album certified Platinum by the RIAA. In 2020, BTS became the first all-South Korean act to reach number one on the Billboard Hot 100 and Billboard Global 200 with their Grammy-nominated single Dynamite. Their follow-up releases Savage Love, Life Goes On, Butter, and Permission to Dance made them the quickest act to earn four US number-one singles since Justin Timberlake in 2006.\n\nBTS in 2022:\n\nAs of 2022, BTS is the best-selling artist in South Korean history, having sold in excess of 30 million albums via the Circle Chart, and their studio album Map of the Soul: 7 (2020) is the best-selling album of all time in South Korea. They are the first non-English-speaking and Asian act to hold sold-out concerts at Wembley Stadium and the Rose Bowl (Love Yourself World Tour in 2019), and were named the International Federation of the Phonographic Industry's (IFPI) Global Recording Artist of the Year for both 2020 and 2021.\n\nALSO READ: Watch: BTS teases hilarious RUN BTS TV On-air show: V hosts, SUGA paints, Jungkook on drums and more\n\nStay updated with the latest Hallyu news on: Instagram, YouTube, Twitter, Facebook and Snapchat\n\nWhich BTS member did you get? Let us know in the comments below.\n'

```
[5]: df1.drop(401, axis=0, inplace=True)
```

```
[6]: df1.reset_index()
df1['content'].iloc[404]
```

[6]: 'J-Hope of BTS, earlier released his debut solo album, Jack in the Box, featuring 10 tracks including the lead single 'More'; and the follow-up 'Arson.' Both singles received music videos, and the album was released on July 15th after BTS announced a temporary break to allow for solo pursuits.\n\nJ-Hope- 'Jack in the Box';\n\nJ-Hope's solo album,

Jack in the Box, is a concept album exploring themes of passion, ambition, insecurity, success, and anxiety. The album's central concept revolves around Pandora's box, and the music itself is a blend of old-school hip hop and elements of pop, grunge, and R&B. Fans and critics alike have praised the album for its thought-provoking lyrics and diverse sound. J-Hope's solo debut has been a success, showcasing his unique style and versatility as an artist. Here are his top 5 solo performances of 2022:

1. J-Hope at Lollapalooza music festival in Chicago
J-Hope made history as the first Korean artist to headline a major US music festival, performing at Lollapalooza as a solo act. His powerful vocals and smooth dance moves mesmerised the audience, and he exceeded high expectations with an unforgettable show. J-Hope performed a variety of songs from his first mixtape and latest album, leaving the audience speechless with his confidence.
2. J-Hope's performance at Lollapalooza not only wowed those in attendance, but also drew a massive online audience with over 14.9 million views on the fan app Weverse. He brought out Becky G for a special performance of 'Chicken Noodle Soup' and gave a full hour-long set that left audiences impressed and satisfied. J-Hope's performance was a major milestone in his career and solidified his place as a talented and accomplished artist.

3. J-Hope of BTS was the star of the show at the MAMA Awards 2022, delivering an electrifying performance of his lead tracks 'More' and 'Arson'; from his debut solo album, 'Jack in the Box'. Dressed in all black, J-Hope presented a unique and artistic stage to fans, captivating them with his powerful raps and charming dance moves.
4. During his performance of 'Arson', J-Hope literally set the stage on fire with his dynamic and explosive energy. His skills as a performer were on full display, and he had the entire audience enthralled with his commanding presence. His strong and confident aura shone brightly as he commanded the stage and had everyone bobbing their heads to his songs. In addition to stealing the show with his performance, J-Hope also took home the award for Most Popular Male Artist, making the night even more special for him.
5. J-Hope performing the changed rap part lyrics for 'Cypher Pt 3: KILLER' at BTS Busan Concert

K-pop sensation BTS treated fans to an unforgettable performance in the city of Busan, presenting the show for free in support of the city's bid to host the World Expo 2030. One of the highlights of the concert was a powerful and emotional performance by member J-Hope of the group's hit song 'Cypher Pt. 3: KILLER'. The lyrics of 'Cypher Pt. 3: KILLER' address critics and speak out against the Korean hip-hop scene, and J-Hope used the opportunity to showcase his personal and professional growth through his verse. He made changes to the lyrics that revealed his true personality and conveyed the message that BTS is now on a different level and too busy to pay attention to negative comments. Fans were in awe of J-Hope's outstanding performance, which revealed BTS's ongoing popularity and progress.

6. J-Hope performing at Singer Crush's concert in Seoul

\n

J-Hope's collaboration with the talented singer-songwriter Crush on the track 'Rush Hour' was highly anticipated by fans, and it did not disappoint. The two artists complemented each other perfectly, with their energy and talent perfectly matching and enhancing one another's. \n\nOn December 25, 2022, Crush held a concert in Seoul called 'Crush Hour,' during which he treated fans to a selection of his hits and newer tracks. As the concert came to a close, Crush surprised the audience by inviting J-Hope to join him on stage for a special performance of 'Rush Hour.' The duo electrified the crowd with their mad energy and brilliant dance moves, with J-Hope once again showcasing his impressive performance skills even for a brief appearance. \n\n5. J-Hope to perform at Dick Clark's New Year's Rockin' Eve\nBTS member J-Hope is set to take the stage as a solo artist at Dick Clark's New Year's Rockin' Eve this year. Today, J-Hope was spotted departing from South Korea and heading to New York City for the performance. The singer seemed thrilled at the opportunity, smiling and waving for the cameras as he made his way to the airport. J-Hope also took to Instagram to share his excitement, posting a story with flight emojis to let his fans know he was on his way. \nBTS fans are eagerly anticipating J-Hope's solo performance at Dick Clark's New Year's Rockin' Eve, and many are hoping to see more solo projects from the talented artist in the future. Whether performing with his bandmates or on his own, J-Hope is sure to put on a memorable and energetic show that will have fans rocking in the New Year. \n\nStay updated with the latest Hallyu news on: Instagram, YouTube, Twitter, Facebook and Snapchat\nALSO READ: Birthday boy Ok Taeyeon in 2PM, Dream High, Save Me, Vincenzo and more'

1.0.1 Realizing that other columns other than published_at, source, topic are useless.

```
[7]: df1.drop(['published_at', 'source', 'topic'], axis=1)
```

```
[7]:
```

	title \
0	BTS: RM is reminded of Bon Voyage as he travel...
1	RM recalls wondering if he 'made right decisio...
2	BTS: J-Hope and RM go bonkers at Billie Eilish...
3	BTS: J-Hope proudly states he raised Jungkook,...
4	BTS: Jin aka Kim Seokjin takes us through the ...
..	...
805	BTS' SUGA's Suchwita Ep 2 Teaser OUT: Top 3 so...
806	BTS ARMY celebrate 700 days of Jin's special s...
807	BTS: 'I am not a baby,' says Jungkook as an AR...
808	BTS' Jin shares 1st pics after joining militar...
809	Bad Decisions: BTS' Jin, Jimin, V & Jungkook j...

	content
0	After reaching his hotel in the city, RM revea...
1	RM aka Kim Namjoon was the first member to joi...

```

2   Billie Eilish's concert was held in Seoul, Sou...
3   BTS ARMY y'all would be missing the members a ...
4   BTS member Kim Seokjin aka Jin has the capacit...
..
805  BTS has conquered the world with their group r...
806  Today marks 700 days since BTS' worldwide hand...
807  BTS' youngest member Jungkook came online on W...
808  BTS' eldest member Jin has shared pictures and...
809  After a lot of teasing, Benny Blanco's collabo...

```

[809 rows x 2 columns]

```
[8]: df1.drop(['published_at', 'source', 'topic'], axis=1, inplace=True)
df1.head()
```

```
[8]:
                                     title \
0   BTS: RM is reminded of Bon Voyage as he travel...
1   RM recalls wondering if he 'made right decisio...
2   BTS: J-Hope and RM go bonkers at Billie Eilish...
3   BTS: J-Hope proudly states he raised Jungkook,...
4   BTS: Jin aka Kim Seokjin takes us through the ...

                                     content
0   After reaching his hotel in the city, RM revea...
1   RM aka Kim Namjoon was the first member to joi...
2   Billie Eilish's concert was held in Seoul, Sou...
3   BTS ARMY y'all would be missing the members a ...
4   BTS member Kim Seokjin aka Jin has the capacit...

```

1.0.2 Checking if there are any null values in the dataframe.

```
[9]: df1.isna().sum()
```

```
[9]: title      0
content      4
dtype: int64
```

1.0.3 As there are null values and are very less (4 « < 810), its better to drop them.

```
[10]: df1.dropna(axis=0, inplace=True)
df1.head()
```

```
[10]:
                                     title \
0   BTS: RM is reminded of Bon Voyage as he travel...
1   RM recalls wondering if he 'made right decisio...
2   BTS: J-Hope and RM go bonkers at Billie Eilish...
3   BTS: J-Hope proudly states he raised Jungkook,...

```

4 BTS: Jin aka Kim Seokjin takes us through the ...

content

```
0 After reaching his hotel in the city, RM revea...
1 RM aka Kim Namjoon was the first member to joi...
2 Billie Eilish's concert was held in Seoul, Sou...
3 BTS ARMY y'all would be missing the members a ...
4 BTS member Kim Seokjin aka Jin has the capacit...
```

```
[11]: df1.shape
```

```
[11]: (805, 2)
```

2 Discovering irregularities in data

```
[12]: df1['content'].iloc[0]
```

```
[12]: 'After reaching his hotel in the city, RM revealed that his stay would be for
four days and added that he would step out for dinner. As he sat at a roadside
open-air restaurant, RM feasted on beer, burgers and fries. He said, "I\'m
starving right now. I\'m out to grab some food. It\'s much quieter than I
expected and feels like a rural town. I like the familiar atmosphere." RM
attended Art Basel and explained on camera the details of the art fair. He also
gave a glimpse as he had noodles and beer which was followed by soup noodles and
wrap. Showing the pattern of a ping pong table, RM said, "The table looks like
our (BTS) symbol." He also spoke about the art pieces as he viewed them. After
that, RM took a tram to visit the Foundation Beyeler, a museum. He later took a
walk through the city. On his third day, RM visited the Kunstmuseum Basel, the
Vitra Design Museum and the gallery. As he walked around, RM showed a chair to
his fans and said, "I have breaking news for you guys. Coldplay\'s Chris Martin
made a chair and it\'s displayed in the Vitra Design Museum. If you see this
Chris, give me a call. You\'re amazing." RM next visited Lucerne and hiked to
Mount Rigi. Recalling his previous visit to Lucerne, RM added, "I remember the
day of crossing that bridge and buying souvenirs." He was also reminded of Bon
Voyage, a reality show featuring BTS members RM, Jin, Suga, J-Hope, Jimin, V and
Jungkook. Speaking to the camera, RM said, "I rode the SSB train to Lucerne,
rode a boat, rode the mountain train, walked down the track road, rode the cable
cars, and now I\'m on a boat planning to go ride the SSB again." RM\'s travel in
Switzerland ended with a visit to the Museum Tinguely. Next, RM flew to Paris to
attend the Pinault Collection and to visit Musee d\'Orsay. He then went to
Centre Georges-Pompidou and Orsay Museum. RM\'s vlog ended with him enjoying a
Korean meal and then heading back to Seoul.'
```

```
[13]: df1['content'].iloc[365]
```

[13]: 'BTS’ Jin is good at a lot of things, singing, being ‘worldwide handsome’, cooking for Bangtan, and now even making alcohol The second week of his own program, Drunken Truth, which starred well-known chef Baek Jong Won and saw a guest appearance from actor Kim Nam Gil, unfolds in unique ways
 \n\nDrunken Truth Episode 3\nAs Jin and Baek Jong Won reach a spot to take care of the final steps of his first try at making makgeolli (rice wine), the two show another attempt at their flourishing synergy Jin is known to be good with elders as he can joke around and that is clearly visible in the show The two head out with their own bottles of self-made alcohol, their destination is a traditional market Being a chef, it is Baek Jong Won’s territory who is recognized by the vendors which consist of mostly older women and is asked for photos, while funnily enough, they are unaware of BTS member Jin A neck-and-neck competition between them shows the difference in preferences between the older and younger generations \n\nBrxoby158IDrunken Truth Episode 4\nBaek Jong Won and Jin are just getting ready to make some side dishes to go along with their drinking session when a guest makes himself known Popular actor Kim Nam Gil’s casual presence adds fun to the episode The older ones take turns poking fun at Jin who in turn is his cheeky self around them, displaying his closeness with the two They enjoy the food and alcohol made by the two and Jin’s keeps cracking jokes throughout as is his personality The show focuses on the chef’s aim to keep the integrity and bring back the popularity of traditional Korean alcohol- soju Kim Nam Gil is a silent observer who was invited impromptu by Jin but adds his meaningful remarks throughout As Jin gifts a set of his famous pyjamas to Baek Jong Won, the night runs deeper The next morning, they are back at it again, appreciating Jin’s skills
 \n\nJin heads to meet the alcohol maestro Park Rokdam to ask for giving him the recipe to him so he can be able to share it with the rest of the BTS members The artisan names this alcohol ‘Brothers of Four Seas’ Jin is congratulated for his fantastic first effort \n\nzVKX6VRMzxsStay updated with the latest Hallyu news on: Instagram, YouTube, Twitter, Facebook and Snapchat\n\nALSO READ: Emergency Declaration Review: Im Siwan's stirring portrayal makes way for Kim Nam Gil & Lee Byung Hun's flight'

```
[14]: text = df1['content'].iloc[369]
      text
```

[14]: '2022 came to a celebratory end for BTS, especially member J-Hope, who was in New York, making another fabulous display of his skills While away from the members, he seemed to have enjoyed it to the fullest with a solo performance at Dick Clark’s New Year’s Rockin’ Eve where he did a live stage of 3 songs, solo track ‘= (Equal Sign)’, ‘Chicken Noodle Soup’ his collaboration track with Becky G, and BTS’ ‘Butter’ (Holiday Remix) He officially became only the second South Korean soloist to perform at the event, following PSY This was also J-Hope’s third time at Dick Clark’s New Year’s Rockin’ Eve after group stages with BTS in 2017 and 2019 \n\nJ-Hope’s live\n\nAfter returning from Times Square where he performed as the penultimate

act alongside multiple other singers from around the world and spoke to Ryan Seacrest, the host of the show, J-Hope was back at his hotel room and turned on a live broadcast to speak with his fans. As they congratulated him on his sparkling performance, being the perfectionist that he is, J-Hope expressed his sadness about being unable to use his voice to the fullest and talked about the slipping incident due to the rainy weather during his rehearsal stage.

J-Hope on other BTS members's New Year wishes

As soon as it turned 12, and the New Year began in South Korea, BTS members Jungkook, Jimin, and V took to the fan community platform Weverse to share their wishes with the fans as well as discuss their plans for the coming year. Jungkook and V kept it brief, wishing for a successful and happy year ahead while Jimin wrote a big letter to his fans about the feelings he had seemingly bottled for a long time about his wishes to release new music soon for which he has been meeting up with composers.

Thousands of miles away, member J-Hope was his cheeky self as he unleashed all his love for his fellow BTS members. He began commenting 'love you' on their posts with cute words and received laughs from them in return.

Jin's call to J-Hope

On being asked about BTS's oldest member Jin who became the first from the group to enlist in the military on December 13, J-Hope recalled how he was called by the member with a different phone number and almost missed it. J-Hope said, "Right before I was about to sleep on the 31st, I got a call from Jin and asked him how he was doing to which he said J-Hope, pick up your phone. I told him I did not know this number, how would I know it was his?"

The 'Arson' hitmaker spoke with a smile on his face about how he felt happy hearing Jin's voice. It comforted him and J-Hope mentioned how he remembered all those moments he spent with Jin. He assured the fans by saying that Jin seemed to be healthy and doing well in the military. So in place of Jin, he shared that he was well and asked the fans to not worry. The lovely duo, BTS's Jin and J-Hope nicknamed 2seok have always lightened the fans's hearts with this interaction.

<https://twitter.com/nightstar1201/status/1609428759454822401>

Jimin X Taeyang

BTS member Jimin and BIGBANG member Taeyang are a collaboration nobody would have expected. However, in December 2022, it was reported by industry officials that the two are working together on a release that will soon be revealed to the world. To this, Taeyang's then-agency, YG Entertainment, which he has since departed, replied by saying that they cannot confirm anything at the moment and asked fans to anticipate Taeyang's activities.

On January 1, Taeyang who is currently a free agent, not having signed with anyone for his solo activities as he continues to be with YG Entertainment for any content related to BIGBANG, shared a rare update on his Instagram account. 2 photos were shared with a caption of the hashtag #2023. In black and white, 2 people could be seen in the photos, their backs to the camera. While one could easily be spotted as Taeyang himself, the other one was not tagged.

Fans went into action and soon began investigating every detail of the photo. The other person seemed to be Jimin, as his hands, fluffy hair, and earrings seem to match the BTS member. Though no official announcement was made from

either side, it seems that the singers have decided to give a green signal from their ends Taeyang is rumoured to have been preparing for a January 2023 comeback so it seems as though we can expect the reports any time now Once confirmed, this could very well be the biggest release of the year and one of the most hyped collaborations in K-pop history!\n\nMeanwhile, BIGBANG’s G-Dragon also announced that he is working on an album and hopes to release it in 2023 \n\nStay updated with the latest Hallyu news on: Instagram, YouTube, Twitter, Facebook and Snapchat\n\nALSO READ: BTS' J-Hope, TXT wow at Dick Clark's New Year's Rockin' Eve: 5 highlights from their stages'

```
[15]: df1['content'].iloc[450]
```

```
[15]: 'On May 27 at 8:30 pm IST (May 28 at midnight KST), BTS dropped the first set of concept photos for their upcoming anthology album, 'Proof' The stunning photos feature the members standing in front of the door of a vault, decked out in all-black outfits Meanwhile, the individual photos of the members show them standing in front of a shield that is riddled with bullets, similar to the door of the vault, showing signs of wear and tear, but still standing strong \xa0Check out the first set of concept photos for BTS' 'Proof', below:https://twitter.com/BIGHIT\_MUSIC/status/1530202174600675328?ref\_src=twsrc%5EtfwAs BTS is an acronym for their Korean name, 'Bangtan Sonyeondan', which translates to 'Bulletproof Boys' / 'Bulletproof Boy Scouts', this concept takes us back to the very beginning of BTS The group's name represents their desire to "block out stereotypes, criticisms, and expectations that aim on adolescents like bullets "\xa0With the group's upcoming anthology set to "reflect the thoughts and ideas of the members on the past, present and future of BTS," and represent the group's history while offering a way to relive it, this first set of teaser photos comes about as the perfect visual representation of the concept of 'Proof' \xa0Another detail that especially touches the heart, is the fact that with BTS' youngest member Jungkook standing in the middle of the group photo, his hand with the word 'ARMY' (BTS' fans) is positioned right in the front and centre of the image This also reminds us of RM's "hint" about the upcoming release during BTS' concert in Los Angeles in November 2021, when he had said, "So, we were the bullet and you guys were our proof So, we're now truly bulletproof "BTS' anthology album 'Proof' drops on June 10, 2022, ahead of their debut anniversary on June 13 Stay tuned for more updates!Join the biggest community of K-Pop fans live on Pinkvilla Rooms to get one step closer to your favourite K-Celebs! Click here to join \xa0ALSO READ: BTS' Jungkook heads to the US ahead of White House visit; Meets up with Pink Sweat$ at Seoul Jazz Festival?'
```

2.0.1 There are HTML Tags and Entities, which needs to be cleaned from data.

```
[16]: from bs4 import BeautifulSoup
```

```
[17]: def removeHTMLTagsAndEntities(s):  
      return BeautifulSoup(BeautifulSoup(s, "lxml").text, "html.parser")
```

```
[18]: removeHTMLTagsAndEntities(df1['content'].iloc[365])
```

[18]: BTS' Jin is good at a lot of things, singing, being 'worldwide handsome', cooking for Bangtan, and now even making alcohol. The second week of his own program, Drunken Truth, which starred well-known chef Baek Jong Won and saw a guest appearance from actor Kim Nam Gil, unfolds in unique ways.

Drunken Truth Episode 3

As Jin and Baek Jong Won reach a spot to take care of the final steps of his first try at making makgeolli (rice wine), the two show another attempt at their flourishing synergy. Jin is known to be good with elders as he can joke around and that is clearly visible in the show. The two head out with their own bottles of self-made alcohol, their destination is a traditional market. Being a chef, it is Baek Jong Won's territory who is recognized by the vendors which consist of mostly older women and is asked for photos, while funnily enough, they are unaware of BTS member Jin. A neck-and-neck competition between them shows the difference in preferences between the older and younger generations.

xBrxoby158IDrunken Truth Episode 4

Baek Jong Won and Jin are just getting ready to make some side dishes to go along with their drinking session when a guest makes himself known. Popular actor Kim Nam Gil's casual presence adds fun to the episode. The older ones take turns poking fun at Jin who in turn is his cheeky self around them, displaying his closeness with the two. They enjoy the food and alcohol made by the two and Jin's keeps cracking jokes throughout as is his personality. The show focuses on the chef's aim to keep the integrity and bring back the popularity of traditional Korean alcohol- soju. Kim Nam Gil is a silent observer who was invited impromptu by Jin but adds his meaningful remarks throughout. As Jin gifts a set of his famous pyjamas to Baek Jong Won, the night runs deeper. The next morning, they are back at it again, appreciating Jin's skills.

Jin heads to meet the alcohol maestro Park Rokdam to ask for giving him the recipe to him so he can be able to share it with the rest of the BTS members. The artisan names this alcohol 'Brothers of Four Seas'. Jin is congratulated for his fantastic first effort.

zVKX6VRMzxsStay updated with the latest Hallyu news on: Instagram, YouTube, Twitter, Facebook and Snapchat

ALSO READ: Emergency Declaration Review: Im Siwan's stirring portrayal makes way for Kim Nam Gil & Lee Byung Hun's flight

```
[19]: text = removeHTMLTagsAndEntities(text)  
      text
```

[19]: 2022 came to a celebratory end for BTS, especially member J-Hope, who was in New York, making another fabulous display of his skills. While away from the

members, he seemed to have enjoyed it to the fullest with a solo performance at Dick Clark's New Year's Rockin' Eve where he did a live stage of 3 songs, solo track '= (Equal Sign)', 'Chicken Noodle Soup' his collaboration track with Becky G, and BTS' 'Butter' (Holiday Remix) He officially became only the second South Korean soloist to perform at the event, following PSY This was also J-Hope's third time at Dick Clark's New Year's Rockin' Eve after group stages with BTS in 2017 and 2019

J-Hope's live

After returning from Times Square where he performed as the penultimate act alongside multiple other singers from around the world and spoke to Ryan Seacrest, the host of the show, J-Hope was back at his hotel room and turned on a live broadcast to speak with his fans As they congratulated him on his sparkling performance, being the perfectionist that he is, J-Hope expressed his sadness about being unable to use his voice to the fullest and talked about the slipping incident due to the rainy weather during his rehearsal stage

J-Hope on other BTS members' New Year wishes

As soon as it turned 12, and the New Year began in South Korea, BTS members Jungkook, Jimin, and V took to the fan community platform Weverse to share their wishes with the fans as well as discuss their plans for the coming year Jungkook and V kept it brief, wishing for a successful and happy year ahead while Jimin wrote a big letter to his fans about the feelings he had seemingly bottled for a long time about his wishes to release new music soon for which he has been meeting up with composers

Thousands of miles away, member J-Hope was his cheeky self as he unleashed all his love for his fellow BTS members He began commenting 'love you' on their posts with cute words and received laughs from them in return

Jin's call to J-Hope

On being asked about BTS' oldest member Jin who became the first from the group to enlist in the military on December 13, J-Hope recalled how he was called by the member with a different phone number and almost missed it J-Hope said, "Right before I was about to sleep on the 31st, I got a call from Jin and asked him how he was doing to which he said J-Hope, pick up your phone I told him I did not know this number, how would I know it was his?"

The 'Arson' hitmaker spoke with a smile on his face about how he felt happy hearing Jin's voice It comforted him and J-Hope mentioned how he remembered all those moments he spent with Jin He assured the fans by saying that Jin seemed to be healthy and doing well in the military So in place of Jin, he shared that he was well and asked the fans to not worry The lovely duo, BTS' Jin and J-Hope nicknamed 2seok have always lightened the fans' hearts with this interaction

<https://twitter.com/nightstar1201/status/1609428759454822401> Jimin X Taeyang

BTS member Jimin and BIGBANG member Taeyang are a collaboration nobody would have expected. However, in December 2022, it was reported by industry officials that the two are working together on a release that will soon be revealed to the world. To this, Taeyang's then-agency, YG Entertainment, which he has since departed, replied by saying that they cannot confirm anything at the moment and asked fans to anticipate Taeyang's activities.

On January 1, Taeyang who is currently a free agent, not having signed with anyone for his solo activities as he continues to be with YG Entertainment for any content related to BIGBANG, shared a rare update on his Instagram account. 2 photos were shared with a caption of the hashtag #2023. In black and white, 2 people could be seen in the photos, their backs to the camera. While one could easily be spotted as Taeyang himself, the other one was not tagged.

Fans went into action and soon began investigating every detail of the photo. The other person seemed to be Jimin, as his hands, fluffy hair, and earrings seem to match the BTS member. Though no official announcement was made from either side, it seems that the singers have decided to give a green signal from their ends. Taeyang is rumoured to have been preparing for a January 2023 comeback so it seems as though we can expect the reports any time now. Once confirmed, this could very well be the biggest release of the year and one of the most hyped collaborations in K-pop history!

Meanwhile, BIGBANG's G-Dragon also announced that he is working on an album and hopes to release it in 2023.

Stay updated with the latest Hallyu news on: Instagram, YouTube, Twitter, Facebook and Snapchat
ALSO READ: BTS' J-Hope, TXT wow at Dick Clark's New Year's Rockin' Eve: 5 highlights from their stages

2.0.2 There are many spaces, new line characters and also some sentences are not even having full stops.

```
[20]: import re
```

```
[21]: def addingFullstops(s):  
    s = re.sub("\xa0", ' ', str(s))  
    s = re.sub("(\n)+", '.', str(s))  
    s = re.sub("(\t)+", '', str(s))  
    s = re.sub(" ", '', str(s))  
    s = re.sub("( )", '.', str(s))  
    return s
```

```
[22]: text = addingFullstops(text)
text
```

[22]: '2022 came to a celebratory end for BTS, especially member J-Hope, who was in New York, making another fabulous display of his skills. While away from the members, he seemed to have enjoyed it to the fullest with a solo performance at Dick Clark's New Year's Rockin' Eve where he did a live stage of 3 songs, solo track '= (Equal Sign)', 'Chicken Noodle Soup' his collaboration track with Becky G, and BTS' 'Butter' (Holiday Remix). He officially became only the second South Korean soloist to perform at the event, following PSY. This was also J-Hope's third time at Dick Clark's New Year's Rockin' Eve after group stages with BTS in 2017 and 2019. J-Hope's live. After returning from Times Square where he performed as the penultimate act alongside multiple other singers from around the world and spoke to Ryan Seacrest, the host of the show, J-Hope was back at his hotel room and turned on a live broadcast to speak with his fans. As they congratulated him on his sparkling performance, being the perfectionist that he is, J-Hope expressed his sadness about being unable to use his voice to the fullest and talked about the slipping incident due to the rainy weather during his rehearsal stage. J-Hope on other BTS members' New Year wishes. As soon as it turned 12, and the New Year began in South Korea, BTS members Jungkook, Jimin, and V took to the fan community platform Weverse to share their wishes with the fans as well as discuss their plans for the coming year. Jungkook and V kept it brief, wishing for a successful and happy year ahead while Jimin wrote a big letter to his fans about the feelings he had seemingly bottled for a long time about his wishes to release new music soon for which he has been meeting up with composers. Thousands of miles away, member J-Hope was his cheeky self as he unleashed all his love for his fellow BTS members. He began commenting 'love you' on their posts with cute words and received laughs from them in return. Jin's call to J-Hope. On being asked about BTS' oldest member Jin who became the first from the group to enlist in the military on December 13, J-Hope recalled how he was called by the member with a different phone number and almost missed it. J-Hope said, "Right before I was about to sleep on the 31st, I got a call from Jin and asked him how he was doing to which he said J-Hope, pick up your phone. I told him I did not know this number, how would I know it was his?". The 'Arson' hitmaker spoke with a smile on his face about how he felt happy hearing Jin's voice. It comforted him and J-Hope mentioned how he remembered all those moments he spent with Jin. He assured the fans by saying that Jin seemed to be healthy and doing well in the military. So in place of Jin, he shared that he was well and asked the fans to not worry. The lovely duo, BTS' Jin and J-Hope nicknamed 2seok have always lightened the fans' hearts with this interaction. <https://twitter.com/nightstar1201/status/1609428759454822401> Jimin X Taeyang. BTS member Jimin and BIGBANG member Taeyang are a collaboration nobody would have expected. However, in December 2022, it was reported by industry officials that the two are working together on a release that will soon be revealed to the world. To this, Taeyang's then-agency, YG Entertainment, which he has since departed, replied by saying that they cannot confirm anything at the moment and asked fans to anticipate Taeyang's activities. On January 1, Taeyang who is

currently a free agent, not having signed with anyone for his solo activities as he continues to be with YG Entertainment for any content related to BIGBANG, shared a rare update on his Instagram account. 2 photos were shared with a caption of the hashtag #2023. In black and white, 2 people could be seen in the photos, their backs to the camera. While one could easily be spotted as Taeyang himself, the other one was not tagged... Fans went into action and soon began investigating every detail of the photo. The other person seemed to be Jimin, as his hands, fluffy hair, and earrings seem to match the BTS member. Though no official announcement was made from either side, it seems that the singers have decided to give a green signal from their ends. Taeyang is rumoured to have been preparing for a January 2023 comeback so it seems as though we can expect the reports any time now. Once confirmed, this could very well be the biggest release of the year and one of the most hyped collaborations in K-pop history!. Meanwhile, BIGBANG's G-Dragon also announced that he is working on an album and hopes to release it in 2023. Stay updated with the latest Hallyu news on: Instagram, YouTube, Twitter, Facebook and Snapchat. ALSO READ: BTS' J-Hope, TXT wow at Dick Clark's New Year's Rockin' Eve: 5 highlights from their stages'

2.0.3 Even after cleaning, there are some repeated full stops without any sentence. Pre-processing it

```
[23]: from nltk.tokenize import sent_tokenize
```

```
[24]: def addFullStops(s):
    modified_sentences = []
    sentences = s.split(".")
    for i in sentences:
        i = i.strip()
        if len(i) != 0:
            i += '.'
        modified_sentences.append(i)
    return " ".join(modified_sentences)
```

```
[25]: text = addFullStops(text)
text
```

```
[25]: '2022 came to a celebratory end for BTS, especially member J-Hope, who was in New York, making another fabulous display of his skills. While away from the members, he seemed to have enjoyed it to the fullest with a solo performance at Dick Clark's New Year's Rockin' Eve where he did a live stage of 3 songs, solo track '= (Equal Sign)', 'Chicken Noodle Soup' his collaboration track with Becky G, and BTS' 'Butter' (Holiday Remix). He officially became only the second South Korean soloist to perform at the event, following PSY. This was also J-Hope's third time at Dick Clark's New Year's Rockin' Eve after group stages with BTS in 2017 and 2019. J-Hope's live. After returning from Times Square where he performed as the penultimate act alongside multiple other singers from around the world and spoke to Ryan Seacrest, the host of the show, J-Hope was back at
```

his hotel room and turned on a live broadcast to speak with his fans. As they congratulated him on his sparkling performance, being the perfectionist that he is, J-Hope expressed his sadness about being unable to use his voice to the fullest and talked about the slipping incident due to the rainy weather during his rehearsal stage. J-Hope on other BTS members' New Year wishes. As soon as it turned 12, and the New Year began in South Korea, BTS members Jungkook, Jimin, and V took to the fan community platform Weverse to share their wishes with the fans as well as discuss their plans for the coming year. Jungkook and V kept it brief, wishing for a successful and happy year ahead while Jimin wrote a big letter to his fans about the feelings he had seemingly bottled for a long time about his wishes to release new music soon for which he has been meeting up with composers. Thousands of miles away, member J-Hope was his cheeky self as he unleashed all his love for his fellow BTS members. He began commenting 'love you' on their posts with cute words and received laughs from them in return. Jin's call to J-Hope. On being asked about BTS' oldest member Jin who became the first from the group to enlist in the military on December 13, J-Hope recalled how he was called by the member with a different phone number and almost missed it. J-Hope said, "Right before I was about to sleep on the 31st, I got a call from Jin and asked him how he was doing to which he said J-Hope, pick up your phone. I told him I did not know this number, how would I know it was his?". The 'Arson' hitmaker spoke with a smile on his face about how he felt happy hearing Jin's voice. It comforted him and J-Hope mentioned how he remembered all those moments he spent with Jin. He assured the fans by saying that Jin seemed to be healthy and doing well in the military. So in place of Jin, he shared that he was well and asked the fans to not worry. The lovely duo, BTS' Jin and J-Hope nicknamed 2seok have always lightened the fans' hearts with this interaction. <https://twitter.com/nightstar1201/status/1609428759454822401> Jimin X Taeyang. BTS member Jimin and BIGBANG member Taeyang are a collaboration nobody would have expected. However, in December 2022, it was reported by industry officials that the two are working together on a release that will soon be revealed to the world. To this, Taeyang's then-agency, YG Entertainment, which he has since departed, replied by saying that they cannot confirm anything at the moment and asked fans to anticipate Taeyang's activities. On January 1, Taeyang who is currently a free agent, not having signed with anyone for his solo activities as he continues to be with YG Entertainment for any content related to BIGBANG, shared a rare update on his Instagram account. 2 photos were shared with a caption of the hashtag #2023. In black and white, 2 people could be seen in the photos, their backs to the camera. While one could easily be spotted as Taeyang himself, the other one was not tagged. Fans went into action and soon began investigating every detail of the photo. The other person seemed to be Jimin, as his hands, fluffy hair, and earrings seem to match the BTS member. Though no official announcement was made from either side, it seems that the singers have decided to give a green signal from their ends. Taeyang is rumoured to have been preparing for a January 2023 comeback so it seems as though we can expect the reports any time now. Once confirmed, this could very well be the biggest release of the year and one of the most hyped collaborations in K-pop history!. Meanwhile, BIGBANG's G-Dragon also announced that he is working on an album and

hopes to release it in 2023. Stay updated with the latest Hallyu news on: Instagram, YouTube, Twitter, Facebook and Snapchat. ALSO READ: BTS' J-Hope, TXT wow at Dick Clark's New Year's Rockin' Eve: 5 highlights from their stages.'

2.0.4 For certain words, I saw contractions were used, so decided to replace them with actual word.

```
[26]: import contractions
```

```
[27]: def removingContractions(s):  
      words = []  
      for i in s.split(" "):  
          words.append(contractions.fix(i))  
      s = " ".join(words)  
      return s
```

```
[28]: removingContractions("I'd")
```

```
[28]: 'I would'
```

```
[29]: text = removingContractions(text)  
text
```

```
[29]: '2022 came to a celebratory end for BTS, especially member J-Hope, who was in  
New York, making another fabulous display of his skills. While away from the  
members, he seemed to have enjoyed it to the fullest with a solo performance at  
Dick Clark's New Year's Rockin' Eve where he did a live stage of 3 songs, solo  
track '= (Equal Sign)', 'Chicken Noodle Soup' his collaboration track with Becky  
G, and BTS' 'Butter' (Holiday Remix). He officially became only the second South  
Korean soloist to perform at the event, following PSY. This was also J-Hope's  
third time at Dick Clark's New Year's Rockin' Eve after group stages with BTS in  
2017 and 2019. J-Hope's live. After returning from Times Square where he  
performed as the penultimate act alongside multiple other singers from around  
the world and spoke to Ryan Seacrest, the host of the show, J-Hope was back at  
his hotel room and turned on a live broadcast to speak with his fans. As they  
congratulated him on his sparkling performance, being the perfectionist that he  
is, J-Hope expressed his sadness about being unable to use his voice to the  
fullest and talked about the slipping incident due to the rainy weather during  
his rehearsal stage. J-Hope on other BTS members' New Year wishes. As soon as it  
turned 12, and the New Year began in South Korea, BTS members Jungkook, Jimin,  
and V took to the fan community platform Weverse to share their wishes with the  
fans as well as discuss their plans for the coming year. Jungkook and V kept it  
brief, wishing for a successful and happy year ahead while Jimin wrote a big  
letter to his fans about the feelings he had seemingly bottled for a long time  
about his wishes to release new music soon for which he has been meeting up with  
composers. Thousands of miles away, member J-Hope was his cheeky self as he  
unleashed all his love for his fellow BTS members. He began commenting 'love
```


you' on their posts with cute words and received laughs from them in return. Jin's call to J-Hope. On being asked about BTS' oldest member Jin who became the first from the group to enlist in the military on December 13, J-Hope recalled how he was called by the member with a different phone number and almost missed it. J-Hope said, "Right before I was about to sleep on the 31st, I got a call from Jin and asked him how he was doing to which he said J-Hope, pick up your phone. I told him I did not know this number, how would I know it was his?". The 'Arson' hitmaker spoke with a smile on his face about how he felt happy hearing Jin's voice. It comforted him and J-Hope mentioned how he remembered all those moments he spent with Jin. He assured the fans by saying that Jin seemed to be healthy and doing well in the military. So in place of Jin, he shared that he was well and asked the fans to not worry. The lovely duo, BTS' Jin and J-Hope nicknamed 2seok have always lightened the fans' hearts with this interaction. <https://twitter.com/nightstar1201/status/1609428759454822401> Jimin X Taeyang. BTS member Jimin and BIGBANG member Taeyang are a collaboration nobody would have expected. However, in December 2022, it was reported by industry officials that the two are working together on a release that will soon be revealed to the world. To this, Taeyang's then-agency, YG Entertainment, which he has since departed, replied by saying that they cannot confirm anything at the moment and asked fans to anticipate Taeyang's activities. On January 1, Taeyang who is currently a free agent, not having signed with anyone for his solo activities as he continues to be with YG Entertainment for any content related to BIGBANG, shared a rare update on his Instagram account. 2 photos were shared with a caption of the hashtag #2023. In black and white, 2 people could be seen in the photos, their backs to the camera. While one could easily be spotted as Taeyang himself, the other one was not tagged. Fans went into action and soon began investigating every detail of the photo. The other person seemed to be Jimin, as his hands, fluffy hair, and earrings seem to match the BTS member. Though no official announcement was made from either side, it seems that the singers have decided to give a green signal from their ends. Taeyang is rumoured to have been preparing for a January 2023 comeback so it seems as though we can expect the reports any time now. Once confirmed, this could very well be the biggest release of the year and one of the most hyped collaborations in K-pop history!. Meanwhile, BIGBANG's G-Dragon also announced that he is working on an album and hopes to release it in 2023. Stay updated with the latest Hallyu news on: Instagram, YouTube, Twitter, Facebook and Snapchat. ALSO READ: BTS' J-Hope, TXT wow at Dick Clark's New Year's Rockin' Eve: 5 highlights from their stages.'

2.0.5 Till here combining what we did

```
[30]: def initial_preprocessing(s):
      s = removeHTMLTagsAndEntities(s)
      s = addingFullstops(s)
      s = removingContractions(s)
      return s

[31]: initial_preprocessing(df1['content'].iloc[369])
```

[31]: '2022 came to a celebratory end for BTS, especially member J-Hope, who was in New York, making another fabulous display of his skills. While away from the members, he seemed to have enjoyed it to the fullest with a solo performance at Dick Clark's New Year's Rockin' Eve where he did a live stage of 3 songs, solo track '= (Equal Sign)', 'Chicken Noodle Soup' his collaboration track with Becky G, and BTS' 'Butter' (Holiday Remix). He officially became only the second South Korean soloist to perform at the event, following PSY. This was also J-Hope's third time at Dick Clark's New Year's Rockin' Eve after group stages with BTS in 2017 and 2019. J-Hope's live. After returning from Times Square where he performed as the penultimate act alongside multiple other singers from around the world and spoke to Ryan Seacrest, the host of the show, J-Hope was back at his hotel room and turned on a live broadcast to speak with his fans. As they congratulated him on his sparkling performance, being the perfectionist that he is, J-Hope expressed his sadness about being unable to use his voice to the fullest and talked about the slipping incident due to the rainy weather during his rehearsal stage. J-Hope on other BTS members' New Year wishes. As soon as it turned 12, and the New Year began in South Korea, BTS members Jungkook, Jimin, and V took to the fan community platform Weverse to share their wishes with the fans as well as discuss their plans for the coming year. Jungkook and V kept it brief, wishing for a successful and happy year ahead while Jimin wrote a big letter to his fans about the feelings he had seemingly bottled for a long time about his wishes to release new music soon for which he has been meeting up with composers. Thousands of miles away, member J-Hope was his cheeky self as he unleashed all his love for his fellow BTS members. He began commenting 'love you' on their posts with cute words and received laughs from them in return. Jin's call to J-Hope. On being asked about BTS' oldest member Jin who became the first from the group to enlist in the military on December 13, J-Hope recalled how he was called by the member with a different phone number and almost missed it. J-Hope said, "Right before I was about to sleep on the 31st, I got a call from Jin and asked him how he was doing to which he said J-Hope, pick up your phone. I told him I did not know this number, how would I know it was his?". The 'Arson' hitmaker spoke with a smile on his face about how he felt happy hearing Jin's voice. It comforted him and J-Hope mentioned how he remembered all those moments he spent with Jin. He assured the fans by saying that Jin seemed to be healthy and doing well in the military. So in place of Jin, he shared that he was well and asked the fans to not worry. The lovely duo, BTS' Jin and J-Hope nicknamed 2seok have always lightened the fans' hearts with this interaction. <https://twitter.com/nightstar1201/status/1609428759454822401> Jimin X Taeyang. BTS member Jimin and BIGBANG member Taeyang are a collaboration nobody would have expected. However, in December 2022, it was reported by industry officials that the two are working together on a release that will soon be revealed to the world. To this, Taeyang's then-agency, YG Entertainment, which he has since departed, replied by saying that they cannot confirm anything at the moment and asked fans to anticipate Taeyang's activities. On January 1, Taeyang who is currently a free agent, not having signed with anyone for his solo activities as he continues to be with YG Entertainment for any content related to BIGBANG, shared a rare update on his Instagram account. 2 photos were shared with a

caption of the hashtag #2023. In black and white, 2 people could be seen in the photos, their backs to the camera. While one could easily be spotted as Taeyang himself, the other one was not tagged... Fans went into action and soon began investigating every detail of the photo. The other person seemed to be Jimin, as his hands, fluffy hair, and earrings seem to match the BTS member. Though no official announcement was made from either side, it seems that the singers have decided to give a green signal from their ends. Taeyang is rumoured to have been preparing for a January 2023 comeback so it seems as though we can expect the reports any time now. Once confirmed, this could very well be the biggest release of the year and one of the most hyped collaborations in K-pop history!. Meanwhile, BIGBANG's G-Dragon also announced that he is working on an album and hopes to release it in 2023. Stay updated with the latest Hallyu news on: Instagram, YouTube, Twitter, Facebook and Snapchat. ALSO READ: BTS' J-Hope, TXT wow at Dick Clark's New Year's Rockin' Eve: 5 highlights from their stages'

```
[32]: initial_preprocessing(df1['content'].iloc[0])
```

```
[32]: 'After reaching his hotel in the city, RM revealed that his stay would be for four days and added that he would step out for dinner. As he sat at a roadside open-air restaurant, RM feasted on beer, burgers and fries. He said, "I am starving right now. I am out to grab some food. It is much quieter than I expected and feels like a rural town. I like the familiar atmosphere." RM attended Art Basel and explained on camera the details of the art fair. He also gave a glimpse as he had noodles and beer which was followed by soup noodles and wrap. Showing the pattern of a ping pong table, RM said, "The table looks like our (BTS) symbol." He also spoke about the art pieces as he viewed them. After that, RM took a tram to visit the Foundation Beyeler, a museum. He later took a walk through the city. On his third day, RM visited the Kunstmuseum Basel, the Vitra Design Museum and the gallery. As he walked around, RM showed a chair to his fans and said, "I have breaking news for you guys. Coldplay\'s Chris Martin made a chair and it is displayed in the Vitra Design Museum. If you see this Chris, give me a call. You are amazing." RM next visited Lucerne and hiked to Mount Rigi. Recalling his previous visit to Lucerne, RM added, "I remember the day of crossing that bridge and buying souvenirs." He was also reminded of Bon Voyage, a reality show featuring BTS members RM, Jin, Suga, J-Hope, Jimin, V and Jungkook. Speaking to the camera, RM said, "I rode the SSB train to Lucerne, rode a boat, rode the mountain train, walked down the track road, rode the cable cars, and now I am on a boat planning to go ride the SSB again." RM\'s travel in Switzerland ended with a visit to the Museum Tinguely. Next, RM flew to Paris to attend the Pinault Collection and to visit Musee d\'Orsay. He then went to Centre Georges-Pompidou and Orsay Museum. RM\'s vlog ended with him enjoying a Korean meal and then heading back to Seoul.'
```

2.0.6 Removing the punctuations and Stopwords and coverting entire sentence into lower case.

```
[33]: from nltk.corpus import stopwords
```

```
[34]: engStopwords = stopwords.words("english")
print(engStopwords)
```

```
['i', 'me', 'my', 'myself', 'we', 'our', 'ours', 'ourselves', 'you', "you're",
"you've", "you'll", "you'd", 'your', 'yours', 'yourself', 'yourselves', 'he',
'him', 'his', 'himself', 'she', "she's", 'her', 'hers', 'herself', 'it', "it's",
'its', 'itself', 'they', 'them', 'their', 'theirs', 'themselves', 'what',
'which', 'who', 'whom', 'this', 'that', "that'll", 'these', 'those', 'am', 'is',
'are', 'was', 'were', 'be', 'been', 'being', 'have', 'has', 'had', 'having',
'do', 'does', 'did', 'doing', 'a', 'an', 'the', 'and', 'but', 'if', 'or',
'because', 'as', 'until', 'while', 'of', 'at', 'by', 'for', 'with', 'about',
'against', 'between', 'into', 'through', 'during', 'before', 'after', 'above',
'below', 'to', 'from', 'up', 'down', 'in', 'out', 'on', 'off', 'over', 'under',
'again', 'further', 'then', 'once', 'here', 'there', 'when', 'where', 'why',
'how', 'all', 'any', 'both', 'each', 'few', 'more', 'most', 'other', 'some',
'such', 'no', 'nor', 'not', 'only', 'own', 'same', 'so', 'than', 'too', 'very',
's', 't', 'can', 'will', 'just', 'don', "don't", 'should', "should've", 'now',
'd', 'll', 'm', 'o', 're', 've', 'y', 'ain', 'aren', "aren't", 'couldn',
"couldn't", 'didn', "didn't", 'doesn', "doesn't", 'hadn', "hadn't", 'hasn',
"hasn't", 'haven', "haven't", 'isn', "isn't", 'ma', 'mightn', "mightn't",
'mustn', "mustn't", 'needn', "needn't", 'shan', "shan't", 'shouldn',
"shouldn't", 'wasn', "wasn't", 'weren', "weren't", 'won', "won't", 'wouldn',
"wouldn't"]
```

```
[35]: def removePunctuations(s):
    words = []
    s = re.sub(r"<[>()|&@#%[\]\'\\"";?~*!]", ' ', str(s))
    for i in s.split(" "):
        if i not in engStopwords:
            words.append(i)
    s = " ".join(words)
    return s
```

```
[36]: removePunctuations(text)
```

```
[36]: '2022 came celebratory end BTS especially member J-Hope New York making
another fabulous display skills. While away members seemed enjoyed fullest solo
performance Dick Clark's New Year's Rockin' Eve live stage 3 songs solo track
'= Equal Sign ' 'Chicken Noodle Soup' collaboration track Becky G BTS'
'Butter' Holiday Remix . He officially became second South Korean soloist
perform event following PSY. This also J-Hope's third time Dick Clark's New
Year's Rockin' Eve group stages BTS 2017 2019. J-Hope's live. After returning
Times Square performed penultimate act alongside multiple singers around world
```

spoke Ryan Seacrest host show J-Hope back hotel room turned live broadcast speak fans. As congratulated sparkling performance perfectionist J-Hope expressed sadness unable use voice fullest talked slipping incident due rainy weather rehearsal stage. J-Hope BTS members' New Year wishes. As soon turned 12 New Year began South Korea BTS members Jungkook Jimin V took fan community platform Weverse share wishes fans well discuss plans coming year. Jungkook V kept brief wishing successful happy year ahead Jimin wrote big letter fans feelings seemingly bottled long time wishes release new music soon meeting composers. Thousands miles away member J-Hope cheeky self unleashed love fellow BTS members. He began commenting 'love you' posts cute words received laughs return. Jin's call J-Hope. On asked BTS' oldest member Jin became first group enlist military December 13 J-Hope recalled called member different phone number almost missed it. J-Hope said "Right I sleep 31st I got call Jin asked said J-Hope pick phone. I told I know number would I know ". The 'Arson' hitmaker spoke smile face felt happy hearing Jin's voice. It comforted J-Hope mentioned remembered moments spent Jin. He assured fans saying Jin seemed healthy well military. So place Jin shared well asked fans worry. The lovely duo BTS' Jin J-Hope nicknamed 2seok always lightened fans' hearts interaction. <https://twitter.com/nightstar1201/status/1609428759454822401> Jimin X Taeyang. BTS member Jimin BIGBANG member Taeyang collaboration nobody would expected. However December 2022 reported industry officials two working together release soon revealed world. To Taeyang's then-agency YG Entertainment since departed replied saying cannot confirm anything moment asked fans anticipate Taeyang's activities. On January 1 Taeyang currently free agent signed anyone solo activities continues YG Entertainment content related BIGBANG shared rare update Instagram account. 2 photos shared caption hashtag #2023. In black white 2 people could seen photos backs camera. While one could easily spotted Taeyang one tagged. Fans went action soon began investigating every detail photo. The person seemed Jimin hands fluffy hair earrings seem match BTS member. Though official announcement made either side seems singers decided give green signal ends. Taeyang rumoured preparing January 2023 comeback seems though expect reports time now. Once confirmed could well biggest release year one hyped collaborations K-pop history . Meanwhile BIGBANG's G-Dragon also announced working album hopes release 2023. Stay updated latest Hallyu news on: Instagram YouTube Twitter Facebook Snapchat. ALSO READ: BTS' J-Hope TXT wow Dick Clark's New Year's Rockin' Eve: 5 highlights stages.'

3 Trying to combine all of these pre-processing in a single function

```
[37]: def preprocessingRow(s):
    s = BeautifulSoup(BeautifulSoup(s, "lxml").text, "html.parser")
    s = re.sub("\xa0", ' ', str(s))
    s = re.sub("(\n)+", '.', str(s))
    s = re.sub("(\t)+", '', str(s))
    s = re.sub(" ", '', str(s))
    s = re.sub("( )", '.', str(s))
```

```

modified_sentences = []
sentences = s.split(".")
for i in sentences:
    i = i.strip()
    if len(i) != 0:
        i += '.'
        modified_sentences.append(i)
s = " ".join(modified_sentences)
words = []
for i in s.split(" "):
    words.append(contractions.fix(i))
s = " ".join(words)
words = []
s = re.sub(r"[<>()|&@ø\[\]\'\\"";?~*!]", ' ', str(s)).lower()
for i in s.split(" "):
    if i not in engStopwords and i not in words:
        words.append(i)
s = " ".join(words)
s = re.sub(' +', ' ', s)
return s

```

```
[38]: preprocessingRow(df1['content'].iloc[369])
```

```
[38]: '2022 came celebratory end bts especially member j-hope new york making another
fabulous display skills. away members seemed enjoyed fullest solo performance
dick clark's year's rockin' eve live stage 3 songs track '= equal sign '
'chicken noodle soup' collaboration becky g bts' 'butter' holiday remix .
officially became second south korean soloist perform event following psy. also
j-hope's third time group stages 2017 2019. live. returning times square
performed penultimate act alongside multiple singers around world spoke ryan
seacrest host show back hotel room turned broadcast speak fans. congratulated
sparkling perfectionist expressed sadness unable use voice talked slipping
incident due rainy weather rehearsal stage. members' year wishes. soon 12 began
korea jungkook jimin v took fan community platform weverse share wishes fans
well discuss plans coming year. kept brief wishing successful happy ahead wrote
big letter feelings seemingly bottled long release music meeting composers.
thousands miles cheeky self unleashed love fellow members. commenting 'love you'
posts cute words received laughs return. jin's call j-hope. asked oldest jin
first enlist military december 13 recalled called different phone number almost
missed it. said "right sleep 31st got pick phone. told know would ". 'arson'
hitmaker smile face felt hearing voice. comforted mentioned remembered moments
spent jin. assured saying healthy military. place shared worry. lovely duo
nicknamed 2seek always lightened fans' hearts interaction. https://twitter.com/nightstar1201/status/1609428759454822401jimin x taeyang. bigbang taeyang
nobody expected. however reported industry officials two working together
revealed world. taeyang's then-agency yg entertainment since departed replied
cannot confirm anything moment anticipate activities. january 1 currently free
```

agent signed anyone activities continues content related rare update instagram account. 2 photos caption hashtag #2023. black white people could seen backs camera. one easily spotted tagged. went action investigating every detail photo. person hands fluffy hair earrings seem match member. though official announcement made either side seems decided give green signal ends. rumoured preparing 2023 comeback expect reports now. confirmed biggest hyped collaborations k-pop history meanwhile bigbang's g-dragon announced album hopes 2023. stay updated latest hallyu news on: youtube twitter facebook snapchat. read: txt wow eve: 5 highlights stages.'

3.0.1 Although, I was not able to figure out how to remove a link because, I wasn't able to find any regular expression for removing it. I did research on link removal fuctions through BeautifulSoup but wasnt able to do it.

3.0.2 The above function will work perfectly when provided row by row input but at that time we cannot utilize the multiprocessing functions of NLP. Hence creating same function, but this time we shall pass an entire column rather than a row to function

```
[39]: from tqdm.auto import tqdm
```

```
[40]: def preprocessingColumn(col):
        for row in tqdm(col, total=col.shape[0]):
            cleanedData = preprocessingRow(row)
            yield cleanedData
```

```
[41]: preprocessingColumn(df1['content']) # will result in a Generator object
```

```
[41]: <generator object preprocessingColumn at 0x000001A464E8D0B0>
```

4 Storing cleaned data in the dataframe

```
[42]: cleanedContent = preprocessingColumn(df1['content'])
```

4.0.1 Using spacy to fasten the process of Data Pre-processing

```
[43]: import spacy
        nlp = spacy.load('en_core_web_sm', disable=['ner', 'parser'])
```

4.0.2 Perorming Data Pre-processing on the Dataframe's content column

```
[44]: cleanedText = [str(doc) for doc in nlp.pipe(cleanedContent, batch_size=5000,
        ↪n_process=-1)]
```

```
0%|          | 0/805 [00:00<?, ?it/s]
```

```
[45]: cleanedText[0]
```


[45]: 'reaching hotel city rm revealed stay would four days added step dinner. sat roadside open-air restaurant feasted beer burgers fries. said starving right now. grab food. much quieter expected feels like rural town. familiar atmosphere. attended art basel explained camera details fair. also gave glimpse noodles followed soup wrap. showing pattern ping pong table looks bts symbol. spoke pieces viewed them. took tram visit foundation beyeler museum. later walk city. third day visited kunstmuseum vitra design museum gallery. walked around showed chair fans breaking news guys. coldplay chris martin made displayed see give call. amazing. next lucerne hiked mount rigi. recalling previous remember crossing bridge buying souvenirs. reminded bon voyage reality show featuring members jin suga j-hope jimin v jungkook. speaking rode ssb train boat mountain track road cable cars planning go ride again. travel switzerland ended tingly. flew paris attend pinault collection musee orsay. went centre georges-pompidou orsay vlog enjoying korean meal heading back seoul.'

[46]: cleanedText[369]

[46]: '2022 came celebratory end bts especially member j-hope new york making another fabulous display skills. away members seemed enjoyed fullest solo performance dick clark's year's rockin' eve live stage 3 songs track '= equal sign ' 'chicken noodle soup' collaboration becky g bts' 'butter' holiday remix . officially became second south korean soloist perform event following psy. also j-hope's third time group stages 2017 2019. live. returning times square performed penultimate act alongside multiple singers around world spoke ryan seacrest host show back hotel room turned broadcast speak fans. congratulated sparkling perfectionist expressed sadness unable use voice talked slipping incident due rainy weather rehearsal stage. members' year wishes. soon 12 began korea jungkook jimin v took fan community platform weverse share wishes fans well discuss plans coming year. kept brief wishing successful happy ahead wrote big letter feelings seemingly bottled long release music meeting composers. thousands miles cheeky self unleashed love fellow members. commenting 'love you' posts cute words received laughs return. jin's call j-hope. asked oldest jin first enlist military december 13 recalled called different phone number almost missed it. said "right sleep 31st got pick phone. told know would ". 'arson' hitmaker smile face felt hearing voice. comforted mentioned remembered moments spent jin. assured saying healthy military. place shared worry. lovely duo nicknamed 2seek always lightened fans' hearts interaction. <https://twitter.com/nightstar1201/status/1609428759454822401>jimin x taeyang. bigbang taeyang nobody expected. however reported industry officials two working together revealed world. taeyang's then-agency yg entertainment since departed replied cannot confirm anything moment anticipate activities. january 1 currently free agent signed anyone activities continues content related rare update instagram account. 2 photos caption hashtag #2023. black white people could seen backs camera. one easily spotted tagged. went action investigating every detail photo. person hands fluffy hair earrings seem match member. though official announcement made either side seems decided give green signal ends. rumoured preparing 2023 comeback expect reports now. confirmed biggest hyped

collaborations k-pop history meanwhile bigbang's g-dragon announced album hopes 2023. stay updated latest hallyu news on: youtube twitter facebook snapchat. read: txt wow eve: 5 highlights stages.'

```
[47]: len(cleanedText)
```

```
[47]: 805
```

4.0.3 Checking for any errors in the process of data cleaning, such that our code might have skipped processing an entire row.

```
[48]: flag = 1
for i in range(len(cleanedText)):
    if len(cleanedText[i]) == 0:
        print(f"Error at: {i}")
        flag = 0
if flag == 1:
    print("No error, in process of Data pre-processing")
```

No error, in process of Data pre-processing

5 Trying to summarize one row

```
[49]: from gensim.models import Word2Vec
from scipy import spatial
import networkx as nx
```

```
[50]: text = initial_preprocessing(df1['content'].iloc[405])
text
```

```
[50]: 'BTS member Jimin\'s first-ever OST, With You, for the K-drama Our Blues was
released on Sunday. The track also features Jimin\'s friend and former HOTSHOT
and Want To One member Ha Sung Woon. YamYam Entertainment released the music
video on their YouTube channel. The song is part of episode six of the K-drama.
(Also Read | BTS: Jimin\'s Our Blues OST release date changed; his friend Ha
Sung Woon reveals details about their upcoming song)The song features actors
Shin Min-a and Lee Byung-hun in Jeju Island. The lyrics of the song start with
Jimin singing, "I want to be with you and I want to stay with you, Just like the
stars shining bright." Ha Sung Woon starts singing after Jimin. The duo sing in
both Korean as well as English. A part of Jimin\'s lyrics also goes, "I am
nothing without your love, I promise I will never leave your love my heart is
beating because of you."Reacting to the clip, BTS fandom ARMY dropped comments
praising Jimin, Ha Sung Woon and the song. A person wrote, "Popular opinion:
Jimin\'s voice can bring you to heaven." A fan commented, "PARK JIMIN ,
everything you do is a masterpiece! Your voice touches you deep in your heart.
You are special and indescribable! You deserve the world." "Sung Woon and
Jimin\'s voice together is such a masterpiece," said a fan.A fan also wrote, "Oh
```

my God there is a warm place in Jimin\'s voice. His voice really makes the person meditate on how beautiful life is His voice is positive energy." A comment read, "Jimin you are art, Jimin\'s voice is perfect, it is beautiful, it is unique, you are an angel Park Jimin, it is the best song in the world! Congratulations great voices together!" On Twitter, fans praised both the singers. A fan tweeted, "Jimin\'s voice is starlight incarnate, it is so beautiful." "Jimin\'s voice is like honey, my ears are blessed with his voice," said another person. "Lyrics are just exactly what Jimin will say to ARMYs #WithYouJiminOST #WithYou @BTS_twt. Jimin you are so lovely, we so lovely, lovely, lovely..love you so much," wrote a fan on Twitter. During a recent live session with J-Hope and V, Jimin had replied to a fan who asked about his OST for the series. Jimin had said, "You guys, I am trying out this for the first time, and I want to try out new things in the future too, so I am starting out from here.. I will work hard!!" While J-Hope had revealed that they \'already know\', V had revealed, "I listened to it earlier today." As quoted by Soompi, producer Song Dong Woon had said about the track, "After looking among numerous songs that would suit Jimin\'s and Ha Sung Woon\'s voices and Our Blues, we were able to finish recording a song that greatly suits the drama." A tvN drama, Our Blues premiered on April 9 and the series takes place on Jeju Island. The television series features Shin Min-a, Lee Byung-hun, Cha Seung-won, Lee Jung-eun, Uhm Jung-hwa, Han Ji-min, and Kim Woo-bin. Our Blues has been written by Noh Hee Kyung, who is known for shows such as It is Okay, That is Love; That Winter, The Wind Blows; Dear My Friends, and Live. This is the first time that Jimin has sung for a K drama. However, other BTS members have already lent their voices to t. While V sang Christmas Tree for Our Beloved Summer, Jin released Yours for Jirisan. BTS members--RM, Jin, Suga, J-Hope, Jimin, V and Jungkook recently returned from Las Vegas after four concerts. They will release their upcoming album in June this year.'

```
[51]: sentences = [i.strip()+"." for i in text.split(".") if len(i) != 0]
      sentences
```

```
[51]: ["BTS member Jimin's first-ever OST, With You, for the K-drama Our Blues was
      released on Sunday.",
      "The track also features Jimin's friend and former HOTSHOT and Want To One
      member Ha Sung Woon.",
      'YamYam Entertainment released the music video on their YouTube channel.',
      'The song is part of episode six of the K-drama.',
      "(Also Read | BTS: Jimin's Our Blues OST release date changed; his friend Ha
      Sung Woon reveals details about their upcoming song)The song features actors
      Shin Min-a and Lee Byung-hun in Jeju Island.",
      'The lyrics of the song start with Jimin singing, "I want to be with you and I
      want to stay with you, Just like the stars shining bright.',
      '" Ha Sung Woon starts singing after Jimin.',
      'The duo sing in both Korean as well as English.',
      'A part of Jimin\'s lyrics also goes, "I am nothing without your love, I
      promise I will never leave your love my heart is beating because of you.',
```

"Reacting to the clip, BTS fandom ARMY dropped comments praising Jimin, Ha Sung Woon and the song.',

'A person wrote, "Popular opinion: Jimin\'s voice can bring you to heaven.',

" A fan commented, "PARK JIMIN , everything you do is a masterpiece! Your voice touches you deep in your heart.',

'You are special and indescribable! You deserve the world.',

" "Sung Woon and Jimin\'s voice together is such a masterpiece," said a fan.',

'A fan also wrote, "Oh my God there is a warm place in Jimin\'s voice.',

'His voice really makes the person meditate on how beautiful life is His voice is positive energy.',

" A comment read, "Jimin you are art, Jimin\'s voice is perfect, it is beautiful, it is unique, you are an angel Park Jimin, it is the best song in the world! Congratulations great voices together!"On Twitter, fans praised both the singers.',

'A fan tweeted, "Jimin\'s voice is starlight incarnate, it is so beautiful.',

" "Jimin\'s voice is like honey, my ears are blessed with his voice," said another person.',

"Lyrics are just exactly what Jimin will say to ARMYs #WithYouJiminOST #WithYou @BTS_twt.',

'Jimin you are so lovely, we so lovely, lovely, lovely.',

'love you so much," wrote a fan on Twitter.',

'During a recent live session with J-Hope and V, Jimin had replied to a fan who asked about his OST for the series.',

'Jimin had said, "You guys, I am trying out this for the first time, and I want to try out new things in the future too, so I am starting out from here.',

'I will work hard!!" While J-Hope had revealed that they \'already know\', V had revealed, "I listened to it earlier today.',

"As quoted by Soompi, producer Song Dong Woon had said about the track, "After looking among numerous songs that would suit Jimin\'s and Ha Sung Woon\'s voices and Our Blues, we were able to finish recording a song that greatly suits the drama.',

"A tvN drama, Our Blues premiered on April 9 and the series takes place on Jeju Island.',

'The television series features Shin Min-a, Lee Byung-hun, Cha Seung-won, Lee Jung-eun, Uhm Jung-hwa, Han Ji-min, and Kim Woo-bin.',

'Our Blues has been written by Noh Hee Kyung, who is known for shows such as It is Okay, That is Love; That Winter, The Wind Blows; Dear My Friends, and Live.',

'This is the first time that Jimin has sung for a K drama.',

'However, other BTS members have already lent their voices to t.',

'While V sang Christmas Tree for Our Beloved Summer, Jin released Yours for Jirisan.',

'BTS members--RM, Jin, Suga, J-Hope, Jimin, V and Jungkook recently returned from Las Vegas after four concerts.',

'They will release their upcoming album in June this year.']

```
[52]: sentence_tokens = []
      for sentence in sentences:
```

```

for word in preprocessingRow(sentence).split(" "):
    if word not in sentence_tokens:
        sentence_tokens.append(word)

```

```

[53]: w2v = Word2Vec(sentence_tokens, vector_size=1, min_count = 1, epochs = 1000)
sentence_embeddings = [[w2v.wv[word][0] for word in words] for words in
    ↪sentence_tokens]
max_len = max([len(tokens) for tokens in sentence_tokens])
sentence_embeddings = [np.pad(embedding,(0,max_len-len(embedding)), 'constant')
    ↪for embedding in sentence_embeddings]

```

```

[54]: similarity_matrix = np.zeros([len(sentence_tokens), len(sentence_tokens)])
for i,row_embedding in enumerate(sentence_embeddings):
    for j,column_embedding in enumerate(sentence_embeddings):
        similarity_matrix[i][j]=1-spatial.distance.
    ↪cosine(row_embedding,column_embedding)

```

```

[55]: nx_graph = nx.from_numpy_array(similarity_matrix)
scores = nx.pagerank(nx_graph)

```

```

[56]: total_summary_sentences = len(sentences) // 4
top_sentence={sentence:scores[index] for index,sentence in enumerate(sentences)}
top=dict(sorted(top_sentence.items(), key=lambda x: x[1], reverse=True)[:
    ↪total_summary_sentences])

```

```

[57]: summary = ""
for sentence in sentences:
    if sentence in top.keys():
        summary += sentence
print(summary)

```

The track also features Jimin's friend and former HOTSHOT and Want To One member Ha Sung Woon. YamYam Entertainment released the music video on their YouTube channel. You are special and indescribable! You deserve the world." "Sung Woon and Jimin's voice together is such a masterpiece," said a fan. "Lyrics are just exactly what Jimin will say to ARMYs #WithYouJiminOST #WithYou @BTS_twt. Jimin you are so lovely, we so lovely, lovely, lovely. During a recent live session with J-Hope and V, Jimin had replied to a fan who asked about his OST for the series. While V sang Christmas Tree for Our Beloved Summer, Jin released Yours for Jirisan.

5.0.1 For below text

```

[58]: text

```

```

[58]: 'BTS member Jimin\'s first-ever OST, With You, for the K-drama Our Blues was
released on Sunday. The track also features Jimin\'s friend and former HOTSHOT
and Want To One member Ha Sung Woon. YamYam Entertainment released the music

```

video on their YouTube channel. The song is part of episode six of the K-drama. (Also Read | BTS: Jimin's Our Blues OST release date changed; his friend Ha Sung Woon reveals details about their upcoming song)The song features actors Shin Min-a and Lee Byung-hun in Jeju Island. The lyrics of the song start with Jimin singing, "I want to be with you and I want to stay with you, Just like the stars shining bright." Ha Sung Woon starts singing after Jimin. The duo sing in both Korean as well as English. A part of Jimin's lyrics also goes, "I am nothing without your love, I promise I will never leave your love my heart is beating because of you."Reacting to the clip, BTS fandom ARMY dropped comments praising Jimin, Ha Sung Woon and the song. A person wrote, "Popular opinion: Jimin's voice can bring you to heaven." A fan commented, "PARK JIMIN , everything you do is a masterpiece! Your voice touches you deep in your heart. You are special and indescribable! You deserve the world." "Sung Woon and Jimin's voice together is such a masterpiece," said a fan.A fan also wrote, "Oh my God there is a warm place in Jimin's voice. His voice really makes the person meditate on how beautiful life is His voice is positive energy." A comment read, "Jimin you are art, Jimin's voice is perfect, it is beautiful, it is unique, you are an angel Park Jimin, it is the best song in the world! Congratulations great voices together!"On Twitter, fans praised both the singers. A fan tweeted, "Jimin's voice is starlight incarnate, it is so beautiful." "Jimin's voice is like honey, my ears are blessed with his voice," said another person. "Lyrics are just exactly what Jimin will say to ARMYs #WithYouJiminOST #WithYou @BTS_twt. Jimin you are so lovely, we so lovely, lovely, lovely..love you so much," wrote a fan on Twitter.During a recent live session with J-Hope and V, Jimin had replied to a fan who asked about his OST for the series. Jimin had said, "You guys, I am trying out this for the first time, and I want to try out new things in the future too, so I am starting out from here.. I will work hard!!" While J-Hope had revealed that they 'already know', V had revealed, "I listened to it earlier today."As quoted by Soompi, producer Song Dong Woon had said about the track, "After looking among numerous songs that would suit Jimin's and Ha Sung Woon's voices and Our Blues, we were able to finish recording a song that greatly suits the drama."A tvN drama, Our Blues premiered on April 9 and the series takes place on Jeju Island. The television series features Shin Min-a, Lee Byung-hun, Cha Seung-won, Lee Jung-eun, Uhm Jung-hwa, Han Ji-min, and Kim Woo-bin. Our Blues has been written by Noh Hee Kyung, who is known for shows such as It is Okay, That is Love; That Winter, The Wind Blows; Dear My Friends, and Live.This is the first time that Jimin has sung for a K drama. However, other BTS members have already lent their voices to t. While V sang Christmas Tree for Our Beloved Summer, Jin released Yours for Jirisan.BTS members--RM, Jin, Suga, J-Hope, Jimin, V and Jungkook recently returned from Las Vegas after four concerts. They will release their upcoming album in June this year.'

5.0.2 We got summary as

```
[59]: summary
```

```
[59]: 'The track also features Jimin\'s friend and former HOTSHOT and Want To One member Ha Sung Woon.YamYam Entertainment released the music video on their YouTube channel.You are special and indescribable! You deserve the world.'"Sung Woon and Jimin\'s voice together is such a masterpiece," said a fan."Lyrics are just exactly what Jimin will say to ARMYs #WithYouJiminOST #WithYou @BTS_twt.Jimin you are so lovely, we so lovely, lovely, lovely.During a recent live session with J-Hope and V, Jimin had replied to a fan who asked about his OST for the series.While V sang Christmas Tree for Our Beloved Summer, Jin released Yours for Jirisan.'
```

6 Time to implement it on actual Dataframe

```
[60]: def summarizerRow(s):
    # Applying only the necessary Pre-processing
    text = initial_preprocessing(s)
    # print("1")

    # Generating the tokens
    sentences = [i.strip()+ "." for i in text.split(".") if len(i) != 0]
    sentence_tokens = [[word for word in preprocessingRow(sentence).split(" ")]
    ↪for sentence in sentences]
    # print("2")

    # Calculating the embedding for each word
    w2v = Word2Vec(sentence_tokens, vector_size=1, min_count = 1, epochs = 1000)
    sentence_embeddings = [[w2v.wv[word][0] for word in words] for words in
    ↪sentence_tokens]
    max_len = max([len(tokens) for tokens in sentence_tokens])
    sentence_embeddings = [np.
    ↪pad(embedding,(0,max_len-len(embedding)),'constant') for embedding in
    ↪sentence_embeddings]
    # print("3")

    # Developing the similarity matrix based on cosine similarity
    similarity_matrix = np.zeros([len(sentence_tokens), len(sentence_tokens)])
    for i,row_embedding in enumerate(sentence_embeddings):
        for j,column_embedding in enumerate(sentence_embeddings):
            similarity_matrix[i][j]=1-spatial.distance.
    ↪cosine(row_embedding,column_embedding)
    # print("4")

    # Implementing the TextRank
```

```

nx_graph = nx.from_numpy_array(similarity_matrix)
try:
    scores = nx.pagerank(nx_graph, max_iter=100000, tol=1.0e-2)

    # Sorting out the importing sentences
    total_summary_sentences = len(sentences) // 3
    top_sentence={sentence:scores[index] for index,sentence in
↪enumerate(sentences)}
    top=dict(sorted(top_sentence.items(), key=lambda x: x[1],
↪reverse=True)[:total_summary_sentences])

    # Joining the most important sentences
    summary = ""
    for sentence in sentences:
        if sentence in top.keys():
            summary += sentence + " "
    # print("5")
    return summary
except:
    summary = ""
    # print("6")
    return summary

```

```
[61]: print(summarizerRow(df1['content'].iloc[402]))
```

At the centre of this plan stand the many official YouTube channels that partake in building a community on the platform for them. This is led primarily by BANGTANTV, the group's behind-the-scenes deal breaker and content carrier. The biggest contributor to these channels continues to be boy group BTS that has now achieved a massive milestone. 7 billion views across all their official channels, BTS has become the most viewed artist in the history of the platform. The difference between the first and the tenth acts on the list, BTS and Ariana Grande, is over 6 billion views.

6.0.1 Developing a similar fuction column-wise

```
[62]: def summarizerColumn(col):
    for row in tqdm(col, total=col.shape[0]):
        summary = summarizerRow(row)
        yield summary

```

```
[63]: summaryGenerator = summarizerColumn(df1['content'])
```

```
[64]: summary = [str(doc) for doc in nlp.pipe(summaryGenerator, batch_size=5000,
↪n_process=-1)]

```

```
0%|          | 0/805 [00:00<?, ?it/s]
```

```
[65]: summary[0]
```

```
[65]: 'As he sat at a roadside open-air restaurant, RM feasted on beer, burgers and fries. " RM attended Art Basel and explained on camera the details of the art fair. He also gave a glimpse as he had noodles and beer which was followed by soup noodles and wrap. After that, RM took a tram to visit the Foundation Beyeler, a museum. As he walked around, RM showed a chair to his fans and said, "I have breaking news for you guys. Coldplay\'s Chris Martin made a chair and it is displayed in the Vitra Design Museum. " RM next visited Lucerne and hiked to Mount Rigi. " RM\'s travel in Switzerland ended with a visit to the Museum Tinguely. '
```

```
[66]: df1['content'].shape
```

```
[66]: (805,)
```

```
[67]: len(summary)
```

```
[67]: 805
```

```
[68]: df1['content'].iloc[804]
```

```
[68]: 'After a lot of teasing, Benny Blanco’s collaborative song with BTS and Snoop Dogg is finally here! ‘Bad Decisions’ features BTS’ vocal line members Jin, Jimin, V and Jungkook along with rapper Snoop Dogg, and is a pre-release single from Benny Blanco’s upcoming full-length album \xa0Released on August 5 at 9:30 am IST, the light beat of ‘Bad Decisions’ becomes the perfect background for its lyrics that express honest feelings to a loved person The dance track carries a cool feeling, which helps the listener feel refreshed on a hot summer day Meanwhile, the hilarious music video sees Benny Blanco getting ready to attend a BTS concert, only to find out that he has got the date wrong, and to be told that he\'s made a "bad decision"!Check out the fun music video for ‘Bad Decisions’, below:\xa0https://www.youtube.com/embed/BGNkkVrJZksPrior to the release, two music video teasers were dropped, adopting a style reminiscent of that of a movie trailer With a classic narration accompanying BTS’ shots from their previous music video, the teasers hilariously teased the then-upcoming song, without actually giving away any hint about the song itself or its music video Further, Benny Blanco also kicked off a hunt for ‘Bad Decisions’, by sharing a video which showed him hiding a USB drive containing the song "somewhere in LA" \xa0https://twitter.com/ItsBennyBlanco/status/1555041143183708160?ref\_src=twsrc%5EtfwThe next few hours saw ARMYs coming together to search for the USB drive, with #FindBadDecisions trending on Twitter With an hour and a half to go before the song’s release, Benny Blanco took to Twitter again, this time dropping an address for the USB drive’s location, and asking fans to meet up there!From the teasers to a hunt for a USB drive and more, ‘Bad Decisions’ has been a series of chaotically fun events!\xa0Following today’s music video premiere and release of ‘Bad Decisions’ on streaming platforms, the lyric video of the song will be out
```


on BTS' YouTube channel tomorrow, at 9:30 am IST on August 6 This will be followed by a 'Visualizer' released through Benny Blanco's YouTube channel on August 8, and the 'BTS Recording Sketch' video through BTS' YouTube channel on August 16 What did you think about 'Bad Decisions'? Share your thoughts with us through the comments!

Join the biggest community of K-Pop fans live on Pinkvilla Rooms to get one step closer to your favourite K-Celebs! Click here to join

ALSO READ: TWICE drops exciting timetable for 11th mini album 'BETWEEN 1&2'; Teases opening trailer, album preview & more'

```
[69]: summary[804]
```

```
[69]: 'With a classic narration accompanying BTS' shots from their previous music video, the teasers hilariously teased the then-upcoming song, without actually giving away any hint about the song itself or its music video Further, Benny Blanco also kicked off a hunt for 'Bad Decisions', by sharing a video which showed him hiding a USB drive containing the song "somewhere in LA". With an hour and a half to go before the song's release, Benny Blanco took to Twitter again, this time dropping an address for the USB drive's location, and asking fans to meet up there!From the teasers to a hunt for a USB drive and more, 'Bad Decisions' has been a series of chaotically fun events! Following today's music video premiere and release of 'Bad Decisions' on streaming platforms, the lyric video of the song will be out on BTS' YouTube channel tomorrow, at 9:30 am IST on August 6. '
```

```
[70]: df2 = pd.DataFrame()
```

```
[71]: df2['Original Content'] = df1['content'].apply(initial_preprocessing)
df2.head()
```

```
[71]:
```

	Original Content
0	After reaching his hotel in the city, RM revea...
1	RM aka Kim Namjoon was the first member to joi...
2	Billie Eilish's concert was held in Seoul, Sou...
3	BTS ARMY you all would be missing the members ...
4	BTS member Kim Seokjin aka Jin has the capacit...

```
[72]: df2.shape
```

```
[72]: (805, 1)
```

```
[73]: df2['Original Content'].iloc[369]
```

```
[73]: '2022 came to a celebratory end for BTS, especially member J-Hope, who was in New York, making another fabulous display of his skills.While away from the members, he seemed to have enjoyed it to the fullest with a solo performance at Dick Clark's New Year's Rockin' Eve where he did a live stage of 3 songs, solo track '= (Equal Sign)', 'Chicken Noodle Soup' his collaboration track with Becky G, and BTS' 'Butter' (Holiday Remix).He officially became only the second South
```

Korean soloist to perform at the event, following PSY. This was also J-Hope's third time at Dick Clark's New Year's Rockin' Eve after group stages with BTS in 2017 and 2019. J-Hope's live. After returning from Times Square where he performed as the penultimate act alongside multiple other singers from around the world and spoke to Ryan Seacrest, the host of the show, J-Hope was back at his hotel room and turned on a live broadcast to speak with his fans. As they congratulated him on his sparkling performance, being the perfectionist that he is, J-Hope expressed his sadness about being unable to use his voice to the fullest and talked about the slipping incident due to the rainy weather during his rehearsal stage. J-Hope on other BTS members' New Year wishes. As soon as it turned 12, and the New Year began in South Korea, BTS members Jungkook, Jimin, and V took to the fan community platform Weverse to share their wishes with the fans as well as discuss their plans for the coming year. Jungkook and V kept it brief, wishing for a successful and happy year ahead while Jimin wrote a big letter to his fans about the feelings he had seemingly bottled for a long time about his wishes to release new music soon for which he has been meeting up with composers. Thousands of miles away, member J-Hope was his cheeky self as he unleashed all his love for his fellow BTS members. He began commenting 'love you' on their posts with cute words and received laughs from them in return. Jin's call to J-Hope. On being asked about BTS' oldest member Jin who became the first from the group to enlist in the military on December 13, J-Hope recalled how he was called by the member with a different phone number and almost missed it. J-Hope said, "Right before I was about to sleep on the 31st, I got a call from Jin and asked him how he was doing to which he said J-Hope, pick up your phone. I told him I did not know this number, how would I know it was his?". The 'Arson' hitmaker spoke with a smile on his face about how he felt happy hearing Jin's voice. It comforted him and J-Hope mentioned how he remembered all those moments he spent with Jin. He assured the fans by saying that Jin seemed to be healthy and doing well in the military. So in place of Jin, he shared that he was well and asked the fans to not worry. The lovely duo, BTS' Jin and J-Hope nicknamed 2seok have always lightened the fans' hearts with this interaction. <https://twitter.com/nightstar1201/status/1609428759454822401> Jimin X Taeyang. BTS member Jimin and BIGBANG member Taeyang are a collaboration nobody would have expected. However, in December 2022, it was reported by industry officials that the two are working together on a release that will soon be revealed to the world. To this, Taeyang's then-agency, YG Entertainment, which he has since departed, replied by saying that they cannot confirm anything at the moment and asked fans to anticipate Taeyang's activities. On January 1, Taeyang who is currently a free agent, not having signed with anyone for his solo activities as he continues to be with YG Entertainment for any content related to BIGBANG, shared a rare update on his Instagram account. 2 photos were shared with a caption of the hashtag #2023. In black and white, 2 people could be seen in the photos, their backs to the camera. While one could easily be spotted as Taeyang himself, the other one was not tagged. Fans went into action and soon began investigating every detail of the photo. The other person seemed to be Jimin, as his hands, fluffy hair, and earrings seem to match the BTS member. Though no official announcement was made from either side, it seems that the singers have

decided to give a green signal from their ends. Taeyang is rumoured to have been preparing for a January 2023 comeback so it seems as though we can expect the reports any time now. Once confirmed, this could very well be the biggest release of the year and one of the most hyped collaborations in K-pop history!. Meanwhile, BIGBANG's G-Dragon also announced that he is working on an album and hopes to release it in 2023. Stay updated with the latest Hallyu news on: Instagram, YouTube, Twitter, Facebook and Snapchat. ALSO READ: BTS' J-Hope, TXT wow at Dick Clark's New Year's Rockin' Eve: 5 highlights from their stages'

```
[74]: df2['New Content'] = summary
      df2.head()
```

```
[74]:                                     Original Content \
0  After reaching his hotel in the city, RM revea...
1  RM aka Kim Namjoon was the first member to joi...
2  Billie Eilish's concert was held in Seoul, Sou...
3  BTS ARMY you all would be missing the members ...
4  BTS member Kim Seokjin aka Jin has the capacit...

                                     New Content
0  As he sat at a roadside open-air restaurant, R...
1  RM aka Kim Namjoon was the first member to joi...
2  They really enjoyed the concert as the audienc...
3  The boys are going to complete their projects ...
4  Some days back, we saw the Sea of Jin concept ...
```

```
[75]: df2.drop(df2[df2['New Content'] == ''].index[0], axis=0, inplace=True)
```

```
[76]: df2[df2['New Content'] == '']
```

```
[76]: Empty DataFrame
      Columns: [Original Content, New Content]
      Index: []
```

```
[77]: df2.isna().sum()
```

```
[77]: Original Content      0
      New Content         0
      dtype: int64
```

```
[78]: df2.shape
```

```
[78]: (804, 2)
```

```
[79]: df2[df2['New Content'].isna()]
```

```
[79]: Empty DataFrame
      Columns: [Original Content, New Content]
```

Index: []

```
[80]: try:
      df2.to_csv('./Summarized News.csv', index=False)
      print("Your file has been created !!!")
    except:
      print("Failed to create fail. Might be due to permission related errors.")
```

Your file has been created !!!

7 Finally working on our Summarized File

```
[81]: df3 = pd.read_csv('./Summarized News.csv')
      df3.head()
```

```
[81]:
```

	Original Content \
0	After reaching his hotel in the city, RM revea...
1	RM aka Kim Namjoon was the first member to joi...
2	Billie Eilish's concert was held in Seoul, Sou...
3	BTS ARMY you all would be missing the members ...
4	BTS member Kim Seokjin aka Jin has the capacit...

	New Content
0	As he sat at a roadside open-air restaurant, R...
1	RM aka Kim Namjoon was the first member to joi...
2	They really enjoyed the concert as the audienc...
3	The boys are going to complete their projects ...
4	Some days back, we saw the Sea of Jin concept ...

```
[82]: df3.isna().sum()
```

```
[82]: Original Content    0
      New Content        0
      dtype: int64
```

```
[83]: content = df3['Original Content'].iloc[370]
      content
```

```
[83]: "On January 23, JYP Entertainment released new concept images for Stray Kids'
second fan meeting 'SKZ's Chocolate Factory'.Each member looks amazing in the
soft pastel clothes and pretty accessories as they work in the chocolate
factory.The fan meeting will be held on February 12 and 13.On February 13th, an
offline fan meeting and an online paid live broadcast on the Beyond Live
platform will be held at the same time, and precious memories will be made with
domestic and foreign fans.Tickets for offline performances were pre-purchased
for the fan club from 8:00 pm to 11:59 pm on January 17th for members of the 2nd
period of the official fan club STAY, and all seats were sold out at the same
```

time as they opened. Thanks to such enthusiastic support, JYP Entertainment opened additional seats available for viewing at 8 pm on January 19th, and this also sold out quickly, realizing the power of Stray Kids' tickets. This fan meeting is the first in about a year since the first official fan meeting on February 20, 2021, and Stray Kids is expected to meet fans around the world for a long time and will spend sweet time together that transcends time and space. Stray Kids, who broke their best scores in various indicators last year and achieved 'Career High', were honored with 4 crowns, including the first grand prize in their debut, at the awards ceremony to close the year. At the '2021 The Fact Music Awards', '2021 Asian Artist Awards', and '2021 Mnet Asian Music Awards' respectively, they won the Artist of the Year Award, Performance Award of the Year, and Worldwide Fans' Choice Top Ten respectively. At the Golden Disc Awards, they added a trophy in the album category to show off their 'K-Pop 4th generation representative group' ALSO READ: The official trailer of Netflix's 'We Are Dead' surpasses THIS number of views on YouTube The magnificent celebration of K-world culminates with The HallyuTalk Awards, watch here What do you think of the concept photos? Let us know in the comments below."

```
[84]: summary = df3['New Content'].iloc[370]
      summary
```

```
[84]: "Tickets for offline performances were pre-purchased for the fan club from 8:00
      pm to 11:59 pm on January 17th for members of the 2nd period of the official fan
      club STAY, and all seats were sold out at the same time as they opened. Thanks
      to such enthusiastic support, JYP Entertainment opened additional seats
      available for viewing at 8 pm on January 19th, and this also sold out quickly,
      realizing the power of Stray Kids' tickets. At the '2021 The Fact Music Awards',
      '2021 Asian Artist Awards', and '2021 Mnet Asian Music Awards' respectively,
      they won the Artist of the Year Award, Performance Award of the Year, and
      Worldwide Fans' Choice Top Ten respectively. "
```

```
[85]: content_split = []
      for i in content.split("."):
          i = i.strip()
          if len(i) != 0 and i != " ":
              content_split.append(i)
      len(content_split)
```

```
[85]: 10
```

```
[86]: summary_split = []
      for i in summary.split("."):
          i = i.strip()
          if len(i) != 0 and i != " ":
              summary_split.append(i)
      len(summary_split)
```

```
[86]: 3
```

```
[87]: removed_lines = []
      for i in content_split:
          if i not in summary_split:
              removed_lines.append(i)
      len(removed_lines)
```

```
[87]: 7
```

7.0.1 Combining all above process into a function

```
[88]: def removed_lines(content, summary):
      content_split = []
      for i in content.split("."):
          i = i.strip()
          if len(i) != 0 and i != " ":
              content_split.append(i)

      summary_split = []
      for i in summary.split("."):
          i = i.strip()
          if len(i) != 0 and i != " ":
              summary_split.append(i)

      removed_line = []
      for i in content_split:
          if i not in summary_split:
              removed_line.append(i)

      removedLines = ". ".join(removed_line)
      return removedLines
```

```
[89]: removed_line = df3.apply(lambda x: removed_lines(x['Original Content'], x['New_
      ↪Content']), axis=1)
```

```
[90]: removed_line
```

```
[90]: 0      After reaching his hotel in the city, RM revea...
      1      The group released their debut single album 2 ...
      2      Billie Eilish's concert was held in Seoul, Sou...
      3      BTS ARMY you all would be missing the members ...
      4      BTS member Kim Seokjin aka Jin has the capacit...
           ...
      799     BTS has conquered the world with their group r...
      800     Since it was released, the meaningful song rec...
      801     After checking out his live video, BTS' Jungko...
```

```
802     BTS' eldest member Jin has shared pictures and...
803     After a lot of teasing, Benny Blanco's collabo...
Length: 804, dtype: object
```

```
[91]: len(removed_line)
```

```
[91]: 804
```

```
[92]: df3.shape
```

```
[92]: (804, 2)
```

```
[93]: df3['Removed Lines'] = removed_line
df3.head()
```

```
[93]:                                     Original Content \
0  After reaching his hotel in the city, RM revea...
1  RM aka Kim Namjoon was the first member to joi...
2  Billie Eilish's concert was held in Seoul, Sou...
3  BTS ARMY you all would be missing the members ...
4  BTS member Kim Seokjin aka Jin has the capacit...
```

```
                                     New Content \
0  As he sat at a roadside open-air restaurant, R...
1  RM aka Kim Namjoon was the first member to joi...
2  They really enjoyed the concert as the audienc...
3  The boys are going to complete their projects ...
4  Some days back, we saw the Sea of Jin concept ...
```

```
                                     Removed Lines
0  After reaching his hotel in the city, RM revea...
1  The group released their debut single album 2 ...
2  Billie Eilish's concert was held in Seoul, Sou...
3  BTS ARMY you all would be missing the members ...
4  BTS member Kim Seokjin aka Jin has the capacit...
```

```
[94]: len(df3['Original Content'].iloc[0].split("."))-len(df3['New Content'].iloc[0].
      ↪split("."))
```

```
[94]: 17
```

```
[95]: len(df3['Removed Lines'].iloc[0].split("."))
```

```
[95]: 17
```

```
[96]: from sentence_transformers import SentenceTransformer, util
```

```
[97]: model = SentenceTransformer('all-MiniLM-L6-v2')
```

```
[98]: originalC = df3['Original Content'].iloc[1]
      newC = df3['New Content'].iloc[1]
```

```
[99]: en_1 = model.encode(originalC)
      en_2 = model.encode(newC)
      print(type(en_1))
```

```
<class 'numpy.ndarray'>
```

7.0.2 Cosine Similarity

```
[100]: result = util.cos_sim(en_1, en_2)

      print(float(result))
      print(type(result))
```

```
0.7904312014579773
```

```
<class 'torch.Tensor'>
```

```
[101]: result_float = result.item()
      print(result_float)
      print(type(result_float))
```

```
0.7904312014579773
```

```
<class 'float'>
```

```
[102]: def metrics(originalC, newC):
      en_1 = model.encode(originalC)
      en_2 = model.encode(newC)
      result = util.cos_sim(en_1, en_2)
      # print(float(result))
      return float(result)
```

```
[103]: cosine = df3.apply(lambda x: metrics(x['Original Content'], x['New Content']),
      ↪axis=1)
```

```
[104]: cosine.shape
```

```
[104]: (804,)
```

```
[105]: df3['Cosine Similarity'] = cosine
```

```
[106]: df3.head()
```

```
[106]:
```

	Original Content \
0	After reaching his hotel in the city, RM revea...
1	RM aka Kim Namjoon was the first member to joi...
2	Billie Eilish's concert was held in Seoul, Sou...
3	BTS ARMY you all would be missing the members ...

4 BTS member Kim Seokjin aka Jin has the capacit...

New Content \

0 As he sat at a roadside open-air restaurant, R...
1 RM aka Kim Namjoon was the first member to joi...
2 They really enjoyed the concert as the audienc...
3 The boys are going to complete their projects ...
4 Some days back, we saw the Sea of Jin concept ...

	Removed Lines	Cosine Similarity
0	After reaching his hotel in the city, RM revea...	0.796752
1	The group released their debut single album 2 ...	0.790431
2	Billie Eilish's concert was held in Seoul, Sou...	0.840863
3	BTS ARMY you all would be missing the members ...	0.709553
4	BTS member Kim Seokjin aka Jin has the capacit...	0.719757

```
[107]: min(cosine)
```

```
[107]: 0.36477231979370117
```

```
[108]: max(cosine)
```

```
[108]: 0.9571455717086792
```

```
[109]: np.mean(cosine)
```

```
[109]: 0.7547245741409448
```

```
[110]: np.median(cosine)
```

```
[110]: 0.7777727246284485
```

```
[111]: nlp1 = spacy.load("en_core_web_lg")
```

```
[112]: doc1 = nlp(originalC)
doc2 = nlp(newC)
print(doc1.similarity(doc2))
```

```
0.9229698009279503
```

```
[113]: def nlptextsimilarity(originalC, newC):
    doc1 = nlp(originalC)
    doc2 = nlp(newC)
    return doc1.similarity(doc2)
```

```
[114]: nlptextsimi = df3.apply(lambda x: nlptextsimilarity(x['Original Content'],
↪x['New Content']), axis=1)
```

```
[115]: df3['NLP Text Similarity'] = nlptextsimi
```

```
[116]: df3.head()
```

```
[116]:
```

	Original Content \
0	After reaching his hotel in the city, RM revea...
1	RM aka Kim Namjoon was the first member to joi...
2	Billie Eilish's concert was held in Seoul, Sou...
3	BTS ARMY you all would be missing the members ...
4	BTS member Kim Seokjin aka Jin has the capacit...

	New Content \
0	As he sat at a roadside open-air restaurant, R...
1	RM aka Kim Namjoon was the first member to joi...
2	They really enjoyed the concert as the audienc...
3	The boys are going to complete their projects ...
4	Some days back, we saw the Sea of Jin concept ...

	Removed Lines	Cosine Similarity \
0	After reaching his hotel in the city, RM revea...	0.796752
1	The group released their debut single album 2 ...	0.790431
2	Billie Eilish's concert was held in Seoul, Sou...	0.840863
3	BTS ARMY you all would be missing the members ...	0.709553
4	BTS member Kim Seokjin aka Jin has the capacit...	0.719757

	NLP Text Similarity
0	0.947731
1	0.922970
2	0.938945
3	0.923807
4	0.909008

```
[117]: min(nlptextsimi)
```

```
[117]: 0.5577889218703898
```

```
[118]: max(nlptextsimi)
```

```
[118]: 0.9906746332886787
```

```
[119]: np.mean(nlptextsimi)
```

```
[119]: 0.9239102736903291
```

```
[120]: np.median(nlptextsimi)
```

```
[120]: 0.9329814334145199
```

```
[121]: def harmonicMean(x, y):
        return (2*x*y)/(x+y)
```

```
[122]: harmean = df3.apply(lambda x: harmonicMean(x['Cosine Similarity'], x['NLP Text_
        ↳Similarity'])), axis=1)
```

```
[123]: df3['Harmonic Mean'] = harmean
```

```
[124]: df3.head()
```

```
[124]:                                     Original Content \
```

```
0 After reaching his hotel in the city, RM revea...
1 RM aka Kim Namjoon was the first member to joi...
2 Billie Eilish's concert was held in Seoul, Sou...
3 BTS ARMY you all would be missing the members ...
4 BTS member Kim Seokjin aka Jin has the capacit...
```

```
                                     New Content \
```

```
0 As he sat at a roadside open-air restaurant, R...
1 RM aka Kim Namjoon was the first member to joi...
2 They really enjoyed the concert as the audienc...
3 The boys are going to complete their projects ...
4 Some days back, we saw the Sea of Jin concept ...
```

```
Removed Lines Cosine Similarity \
```

```
0 After reaching his hotel in the city, RM revea... 0.796752
1 The group released their debut single album 2 ... 0.790431
2 Billie Eilish's concert was held in Seoul, Sou... 0.840863
3 BTS ARMY you all would be missing the members ... 0.709553
4 BTS member Kim Seokjin aka Jin has the capacit... 0.719757
```

	NLP Text Similarity	Harmonic Mean
0	0.947731	0.865708
1	0.922970	0.851574
2	0.938945	0.887202
3	0.923807	0.802628
4	0.909008	0.803388

```
[125]: df3.to_csv('./Final Result.csv')
```

```
[126]: min(harmean)
```

```
[126]: 0.49884301558564753
```

```
[127]: max(harmean)
```

```
[127]: 0.9623399165870538
```

```
[128]: np.mean(harmean)
```

```
[128]: 0.8265396330480197
```

```
[129]: np.median(harmean)
```

```
[129]: 0.8414175524334191
```