

# **NLP Topic & Sentiment Analysis on the HK National Security Law**

**Author | Rand Sobczak Jr.**

**Date | xx September 2021**

# Table of Contents

- 1.0 [Problem Identification Overview](#)
- 2.0 [Generated Deliverables](#)
- 3.0 [Data Pre-Processing Steps](#)
  - 3.1 [Data Cleaning](#)
    - 3.1.1 [News](#)
    - 3.1.2 [Twitter](#)
  - 3.2 [Exploratory Data Analysis & Pre-processing](#)
    - 3.2.1 [News](#)
    - 3.2.2 [Twitter](#)
- 4.0 [Model Description](#)
- 5.0 [Model Findings](#)
  - 5.1 [News](#)
  - 5.2 [Twitter](#)
- 6.0 [Reading Between the Lines](#)
- 7.0 [Next Steps](#)

## 1.0 Problem Identification Overview

After the end of the First Opium War in 1841, the Qing Dynasty ceded Hong Kong ("HK") to the United Kingdom ("UK"). After World War 2 & still under British rule, HK became a global financial center & its population developed a hybrid culture of maintaining their Chinese ethos while adopting British principles; notably Common Law.

In 1984, the UK & China signed the Sino-British Joint Declaration whereby the UK agreed to cede HK back to China in 1997 as "one country, two systems" ("OcTs"). The mutually agreed OcTs outlined that from 1997 to 2047, HK would not participate in China's socialist system & HK's capitalist system, Common Law & way of life to remain unchanged. In 1997, the transfer was complete & HK was handed back to China under OcTs.

In 2020, the Congress of China unanimously passed the National Security Law ("NSL") which, for simplicity purposes, criminalizes secession, subversion of state power, terrorism & collusion with foreign entities in HK. This law was not enacted by the HK Legislative Council which in Article 23 of HK Basic Law is defined as the governing body allowed to enact such law. This **dichotomy in perceived legal right led to notable protests in HK** ("HK Protests").

## 2.0 Generated Deliverables

This Natural Language Processing ("NLP") Data Science project uses the Data Science Method ("DSM") to identify variances in sentiment during the HK Protests using a hybrid of automatic & rules-based systems. The source of the data was obtained by scraping selected international newspapers & Twitter. The primary modules & tools for the DSM process are defined below:

- Bag of Words, BeautifulSoup, datetime, Emoji Dictionary, genism, matplotlib, NER, the nltk interface, os, Pandas, Regex, scikit-learn, seaborn, Spacy, statistics, Tokenization, twint, pickle, Word2Vec & WordCloud

The goal of the DSM is not to find or present an answer to an apparent geo-political situation but to pragmatically present the facts as they exist; the facts are created by humans resting in unstructured text in newspapers & social media outlets. Therefore, the initial goal is to build the following:

- Sentiment Analysis & Topic Analysis

To reiterate, the goal of this project is not to establish a position on who's right. The goal is to go through the DSM to see where certain groups stand & on what footing.

With this, **three ( 3 ) deliverables were generated**:

1. The source code for the modeling developed to analyze the aforementioned problem ([link](#))
2. This document outlining the process
3. A PDF presentation with our findings ([link](#)) referenced

## Data Pre-Processing Steps

### Data Cleaning (Master):

1. The data is **broken down into two (2) main groups** defined below with explanations:
  - a. **News**, using *beautifulsoup4* ([link](#))
    - i. Thirty (30) News articles deliberating on the National Security Law & the respective protests were identified.
    - ii. The Articles consist of an equal split between China / HK Articles & International Articles ([Appendix I](#))
    - iii. The ordering of the News Articles does not impact the findings & supported by no particular reason
    - iv. The data definitions & explanations are also found in ([Appendix I \(cont.\)](#))
  - b. **Twitter**, using *Twint* ([link](#))
    - i. 288,416 Tweets from Twitter were pulled from people submitting Tweets ("Tweeters")
    - ii. These cover a three (3) month period between 31 May to 29 Aug 2020
      1. The National Security Law was introduced on 30 June 2020
    - iii. They were pulled based on twelve (12) hashtags ([Appendix II](#))
      1. For transparency purposes, these hashtags exclusively arise from HK and International Tweeters; Chinese social media's contribution is not present herein which will we aim to treat in a fair manner
      2. The full dataset is found on the GitHub site under the data section ([link](#))

We will segment the final steps of Data Cleaning into their respective groups.

### Data Cleaning (News):

1. No null values were presented
2. **Non-ascii Values** ([link](#)) **presented the only cleaning step** for the News; these presented themselves as "large apostrophes" (for lack of a better word) & thus, converted to normal English for NLP handling ([Appendix III](#))
3. **No Duplicate** observations were presented
4. **Sentiment Intensity Analyzer ("SIA")** was then added ([Appendix III](#))
  - a. The SIA is a text analysis method that detects polarity; e.g., a positive or negative opinion
  - b. The scores are composed of four (4) categories:
    - i. Compound (comp): computed by normalizing the other scores with a range of -1 to 1.
    - ii. Negative (neg): **xxx**; with a range of 0 to 1.
    - iii. Neutral (neu): **xxx**; with a range of 0 to 1.
    - iv. Positive (pos): **xxx**; with a range of 0 to 1.
  - c. It's also known as opinion mining, deriving the opinion or attitude of the "tweeter". The probabilities will add up to 100%.
5. **Stop words** were then removed
  - a. Stop words are a set of commonly used words in a language that provide little value ([link](#))
6. **Stemming & lemmatization** were then done on the Articles & Titles. The **outputs were assigned to their own variables** to maintain the original Articles & Titles
7. **Term Frequency Inverse Document Frequency** ("TfidfVectorizer") a common algorithm to transform text into a meaningful representation of numbers which we can use to fit machine algorithms for prediction was the final step

### Data Cleaning (Twitter):

1. Using a language variable obtained in scraping, the Twitter data came in **x51 languages with 67% in English & 24% in Unknown**. It was discovered that **Unknown are Tweets with only hashtags or a link**. In conjunction with **other languages (~9%)**, **Unknown** observations **were dropped** resulting in a **revised data frame of 194,087 observations** ([Appendix IV](#))
2. **-30k duplicates** were identified & then dropped **resulting in a data frame of 163,699 observations**
3. With the help of Regex, **Mentions of and to Other Users** in Tweets were then defined & dropped ([Appendix IV](#))
4. Regex was also used to **remove links** to websites within Tweets ([Appendix IV](#))
5. Emoji's are also found in Tweets. Using an Emoji dictionary ([link](#)), **Emoji's were converted to words** ([Appendix IV](#))
6. **Profanity** was converted to the word "**NEGATIVE**" so the SIA would work correctly ([Appendix IV](#))
7. The **SIA** was then added ([Appendix IV](#))
8. **Stop Words** were then removed from the Tweets
9. **Stemming & lemmatization** were then done & the outputs were **assigned to their own variables** to maintain the original Tweet
10. **TfidfVectorizer** was the final step

## Exploratory Data Analysis (EDA) & Pre-processing:

In the Exploratory data analysis (**EDA**) section is **aimed to summarize the main characteristics of a News Article or Tweet & identify what the data may be telling us.**

### EDA & Pre-processing | News

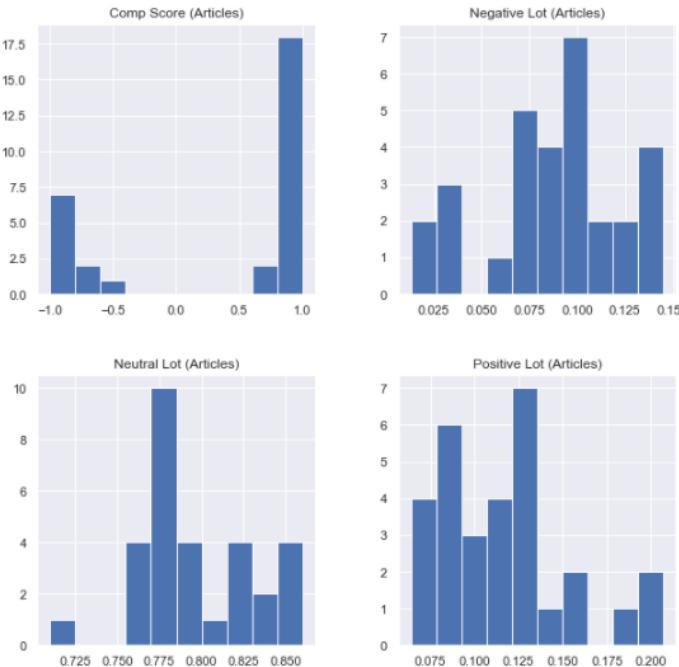
We split the analysis between the Titles & Articles; there was no split between Chinese & International news sources.

**Titles** | The sequential order of the process & their findings listed below:

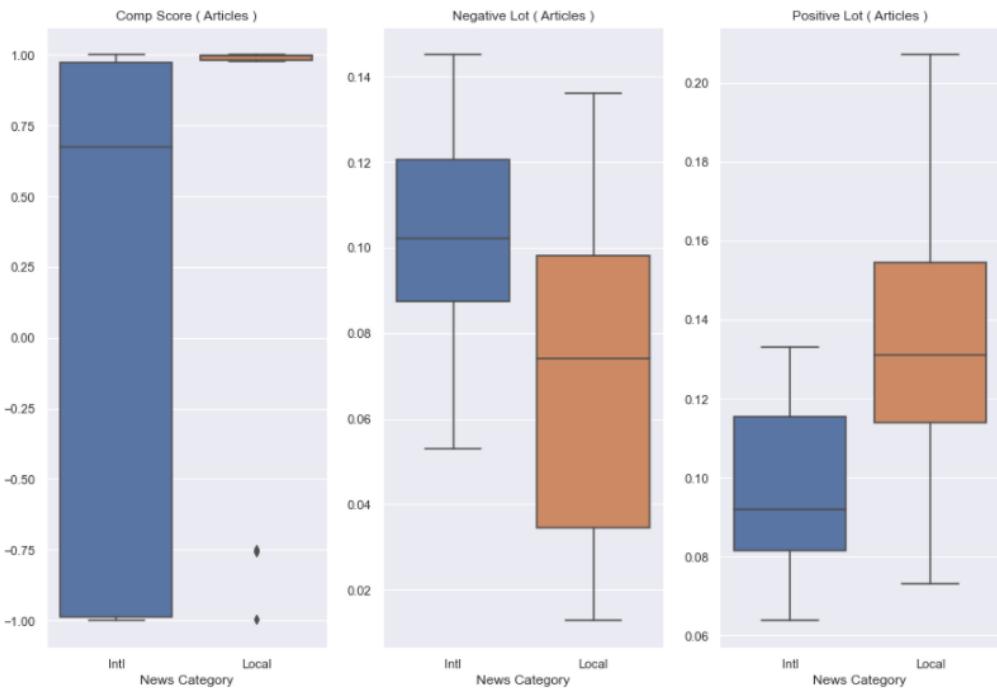
1. Distribution of the SIA scores via **Histograms**
  - a. Skewness observed to the Left on Comp Scores & Right on Negative & Positives Scores
2. Distribution of the SIA scores via **Box Plots**
  - a. **China** | The noticeable observation is the concentration of China's News Titles at Zero on the Negative category. The Compound score also showed outliers towards negative one (-1) while a predominate concentration towards one (1) on the Compound Score. **This suggests that Chinese News are predominately positive with some outliers.**
  - b. **International** | International News Titles were fairly concentrated in the middle with some outliers on the Positive side. **No objective takeaway was received from these results.**
3. The **Topic & Bottom five (5) Comp scores** via a **Bar Chart**
  - a. **Top 5:** These are titles suggesting **positivity** as they gravitate towards 1 on the Comp Score.
    - i. **Chinese news dominated** with a common thread that the NSL will **protect democracy & freedom in HK & lead to stability & prosperity**
  - b. **Bottom 5:** These are titles suggesting **negativity** as they gravitate towards -1 on the Comp Score.
    - i. **International news held the majority** with a common thread that the NSL is leading to anger, is a more definitive stance by China & HK's future looks volatile
    - ii. **Chinese News** was present with a common thread of the international news being misinformed & biased
4. Visual representation of the most frequently used Nouns & Adjective / Verbs via WordCloud's
  - a. No substantial takeaway was achieved with the WordCloud's
  - b. Interesting observations were safeguarding, democracy, protests, books freedom & imposed

The supporting charts & WordCloud's are not presented here as Titles had an average of 9.6 words; they can be found in [Appendix V](#).

**Articles** | The sequential order of the process remains the same albeit Articles contain 643 words on average; thus we will include the EDA output here as well as [Appendix VI](#). We hope our findings can be easier to understand next to the charts & word clouds.



The **most distinct observation** on the **Distribution of the SIA on the Histograms is the polarity in the Comp Scores**. There does not seem to be a "middle ground" as the thirty (30) articles deviate to one side; either one (1) or negative one (-1) being Positive or Negative respectively. Further to this, **the Positive side of the equation commands more than the Negative side**.



upper & lower quartile ranges deviate from near 1 to near -1. The key takeaway would be:

- Chinese News are primarily Positive, &
- International News are diverse

The Box Plots highlight a deeper understanding of how this polarity may be presenting itself.

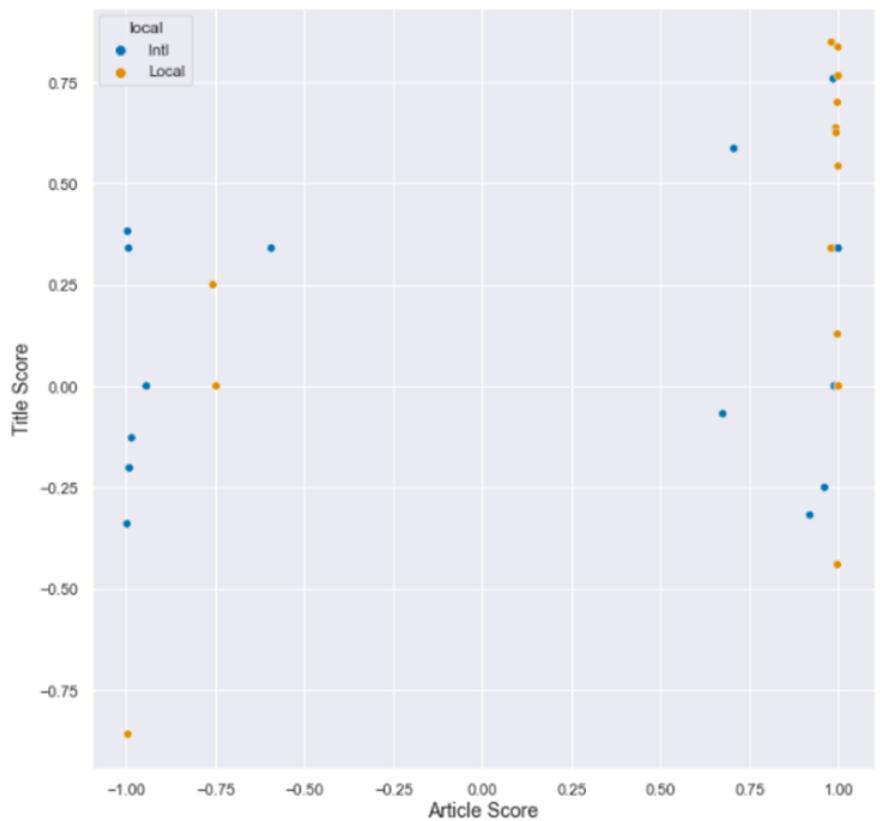
From right to left, **the Positive Lot appears to be supported by the Chinese New Articles**. It's Median (~0.13) is above the upper quartile range of their International counterparts.

**The Negative Lot**, meanwhile, **is more concentrated with the International Articles**. Its median (~0.11) lies above the upper quartile range of their China counterparts.

**The Comp Score**, which is a combination of them all, **shows a high concentration of China News in the Positive section (1) with some (x2) outliers in the Negative section below Zero**.

International News appears more diverse in Positive & Negative as it's

Comparing the Comp Scores between Articles & Titles



This observation inspired the construction of a Scatter Plot with Comp Scores of both the Articles & Titles on the x & Y-axis respectively.

The apparent **concentration of China News** for both on the **Positive sides (upper right)** is the **main takeaway**

While nearly half of the Negative International News Articles had a Positive or Neutral Title.

Nevertheless, **the break down is truly a ±50% on the Articles**. The last takeaway from this section would be that **while the China Articles control the Positive camp, it's not a domineering**.

Conversely, **the International Articles control the Negative camp, it's also not domineering**.

20 | Articles above 50%  
10 | Articles below -50%

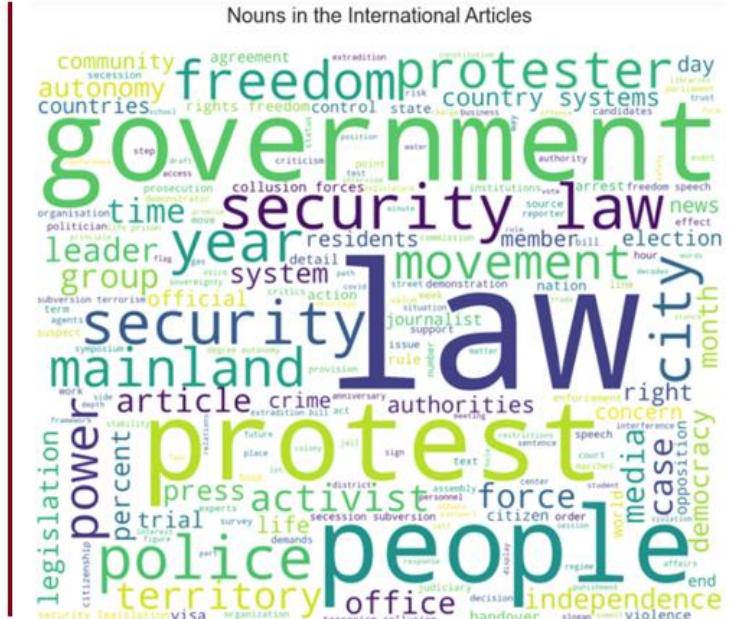
The breakdown between groups +50%:  
China | 60.0 percent  
Intl | 40.0 percent

The breakdown between groups -50%:  
China | 30.0 percent  
Intl | 70.0 percent

A clearer understanding of what the Titles for the Positive & Negative Articles in the [Appendix](#) but a slightly deeper understanding of how those broke down can be found after Modeling. Key takeaways for the potential polarity are as below:

- **International Negative** | political activists quitting, censorship in reading materials & confrontations with police
- **China Positive** | the Law protects & supports HK's democracy & freedom, will lead to prosperity & the wide support in HK for it
- **International Positive** | The Titles don't convey much positivity
- **China Negative** | The BBC are misinformed & present a comprehensive / objective report from the Chief Executive of HK

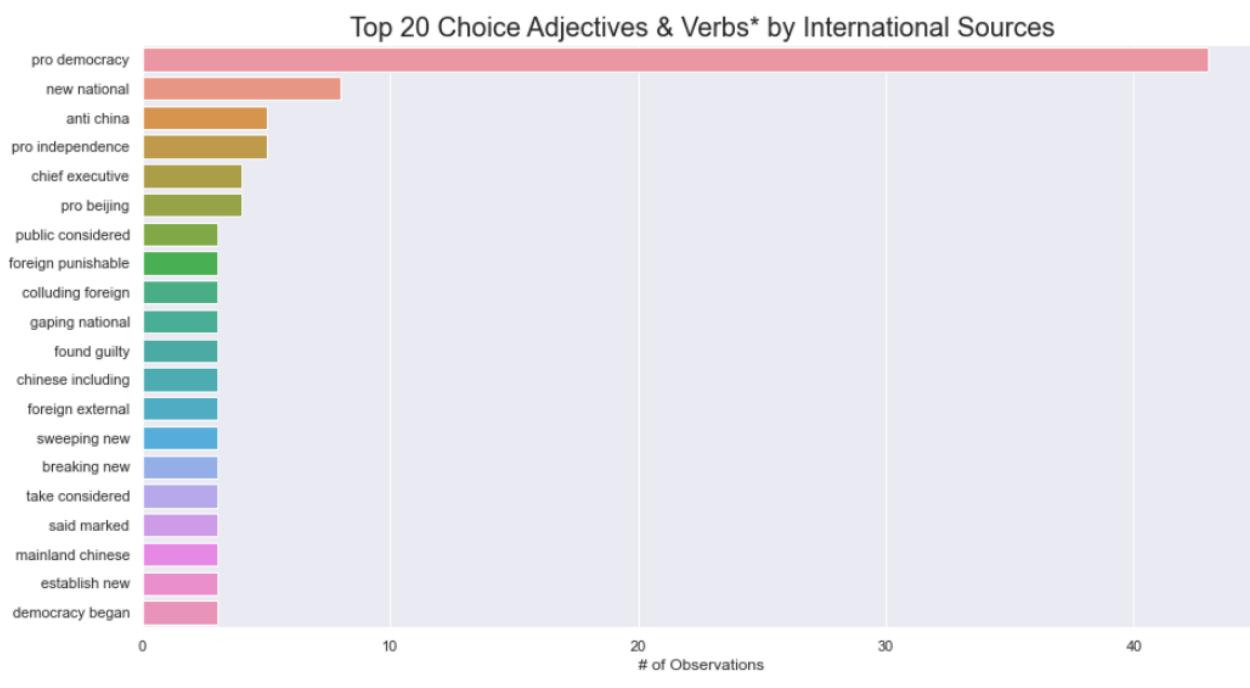
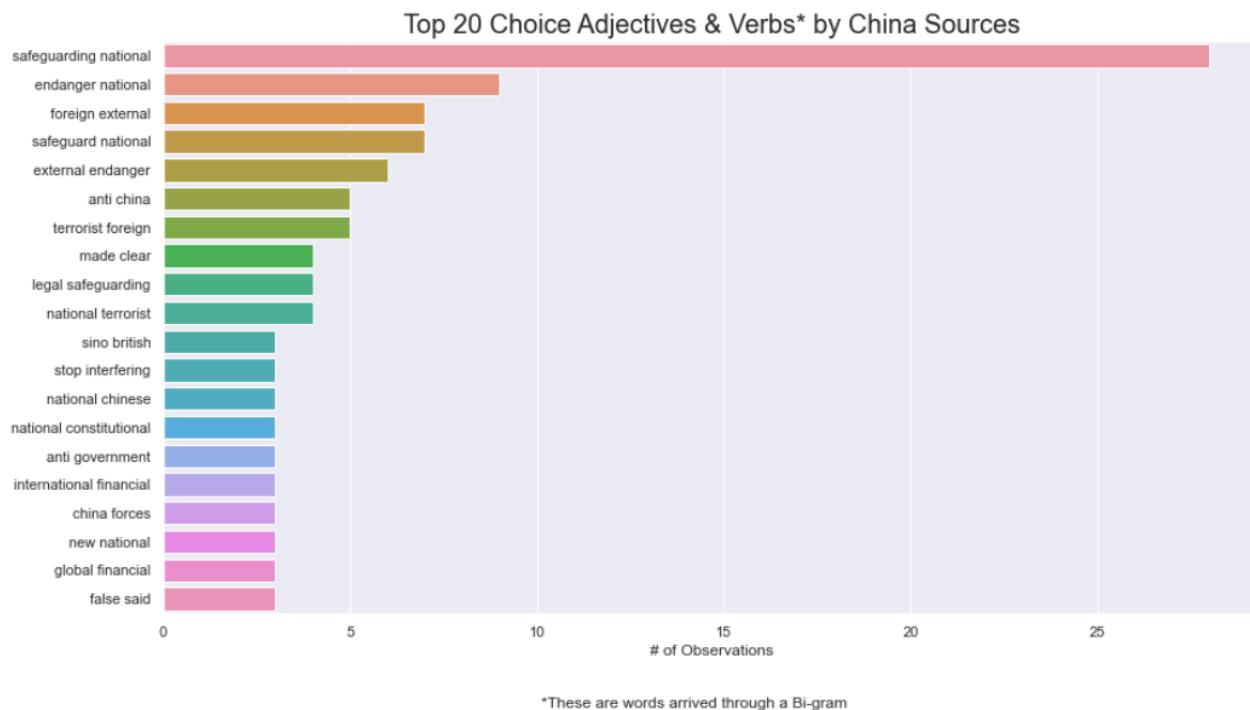
Moving over into the WordCloud's we can see what may be within this polarity. The categories were dissected into to show what words apply to each category for a simpler side-by-side.



The WordCloud's are difficult to digest but nice on the eyes. Some observations below:

- China uses words like stability (only) & motherland conveying unity. International articles use of the word mainland isn't unity
- China use of the words **safeguarding** & **peaceful** are a polar opposite of the International articles use of the words **inciting** & **arrested**
- International articles contained **pro democracy** which aren't present in the China articles

Counting the number of frequently mentioned bi-gram words with adjectives & verbs for each category further highlights the respective categories message.



The final notes for News EDA are as below:

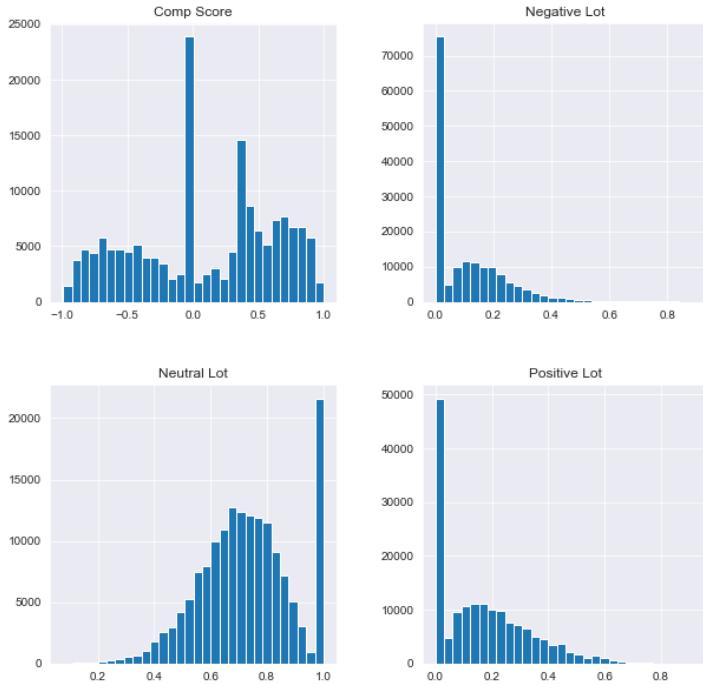
- **China Positive** | China & the Law's intent is to protect & support the freedom & prosperity of HK as a unified motherland
- **International Positive** | No clear indication for a positive position to be confirmed in later stages
- **China Negative** | The West may be misinformed or not see the longer term as well as endangering HK by interfering
- **International Negative** | Concern over the freedom of the HK people, particularly those as protestors

## EDA & Pre-processing | Twitter

Twitter EDA is similar to the News albeit isn't broken into Title or Article groups as well as Chinese or International Tweets.

As mentioned, it is prudent to assume that the entirety of Tweets are not from within China & expected to be disgruntled with the NSL. The inclusion of only non-China based social media is not the desire albeit the consequence of not having known Chinese social media to scrape as well as the Chinese populous does not have a desire to express their political opinions on social media.

The supporting charts & visuals will also be included herein albeit they will all be found in [Appendix VII](#) for reference.



The first step of the EDA process was to investigate the **SIA on a Histogram** for each of its categories.

The Negative & Positive Lots both present skewness. The Neutral Low is seen on all as well.

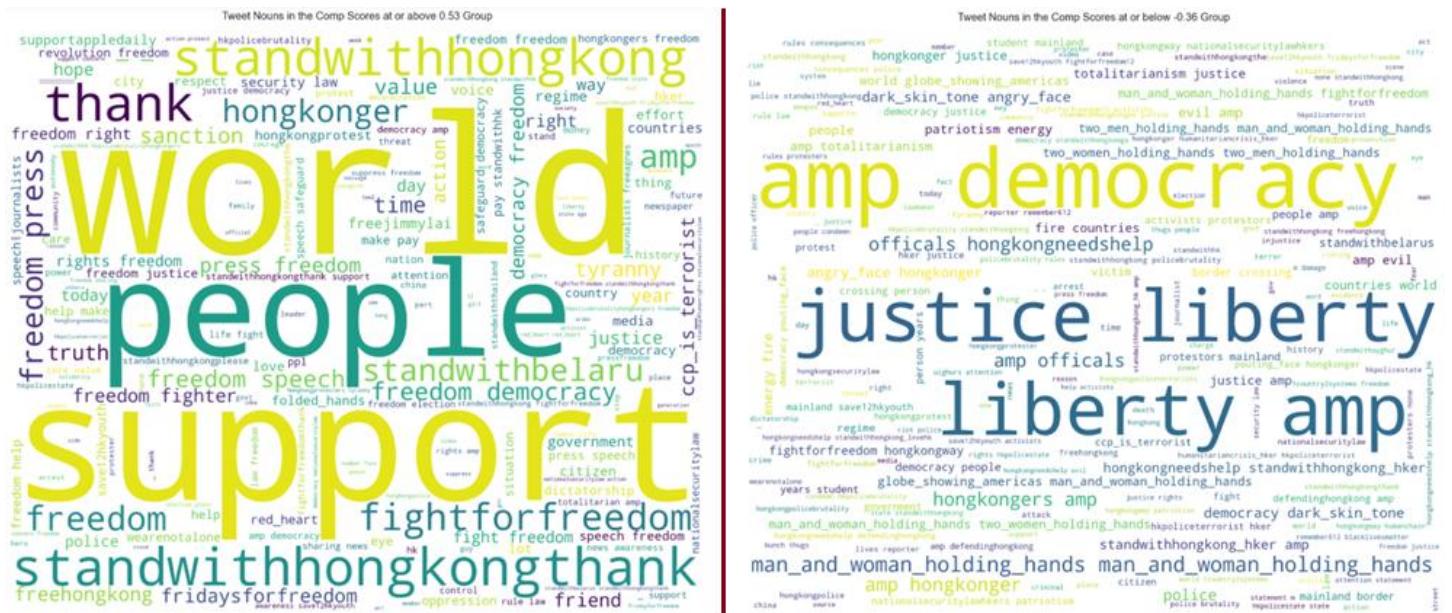
The notable observation is the **Comp Score** itself; it **appears Bimodal**. This **suggests** there to be a **grouping of Positive & Negative Tweets**. In other words, the Tweeters don't only convey negativity in their social media engagements. In fact, the number of Positive Tweets seem to outnumber the Negative ones.

With in the Bimodal nature of the Comp Score, we decided to investigate how to best split the Tweets into the two (2) categories; i.e., Positive & Negative.

Using Descriptive Statistics, we determined that **Tweets with a Comp Score at or above 0.539 to be Positive & Tweets at or below -0.361 to be Negative**.

count	163699.000000
mean	0.095108
std	0.534434
min	-0.990100
25%	-0.361200
50%	0.077200
75%	0.539900
max	0.998900
Name:	comp, dtype: float64

The WordCloud's present a fair introduction to the segments of the groups. World, support & standwithhongkong align to my expectations of being Positive. The odd observation is that democracy, liberty & justice fall into the Below -0.36 camp; i.e., Negative. Those come across as Positive words.



The further look into these groups can be seen on the Bigrams of the Top 30 Words for the following groups with their respective Appendix:

- Positive | Comp Score at or Above 0.53 | [Appendix](#)
- Negative | Comp Score at or Below -0.36 | [Appendix](#)

We decided to save them for the Appendix to save space. The key observations of Bigrams used for each are listed below:

- **Negative Summary:**

- Human Rights
- Evil CCP
- References to Freedom / Democracy
- Police Brutality
- Fighting
- Holdings Hands ( 🤝 )

- **Positive Summary:**

- Thanks much / support
- Human Rights
- Free World
- Please standwithhongkong / support
- folded\_hands ( 🙏 )
- freedom press / democracy

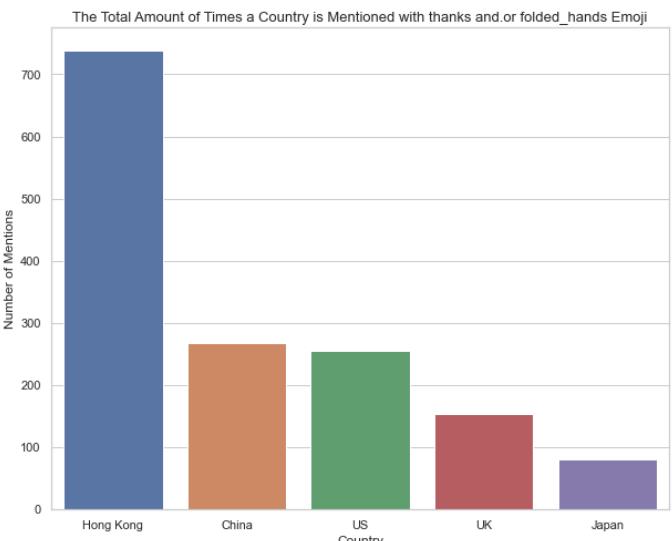
We also took a look at whether or not there are any notable **mentions to geopolitical entities associated with thanks & folded\_hands** ( 🙏 ); inspired by an interest to help determine what the “Free World” is referring too.

Hong Kong is where we believe most of the Tweeters are from. The observation is that China comes in as second place.

We took a look at Tweets which used the word China (below) & determined there may be sarcasm but also mentions of the international community.

The next top countries were the US, the UK & Japan. This was ironic as the three (3) countries used to pull News from were the same. No planning was associated with this outcome.

Mentions of the US was also looked at & it appears that this may be one country referred to as the Free World.



Examples of observations with both thanks & folded\_hands for China are below.

Thank #StandWithHongKong !!! We sincerely hope free world stand firm Communist China's human rights abuses Hong Kong. Many thanks folded\_hands light\_skin\_tone

Deeply touched heartfelt thanks done #HK folded\_hands light\_skin\_tone folded\_hands light\_skin\_tone folded\_hands light\_skin\_tone Always grateful international support fighting evil #CCPChina regime. #StandWithHongKong

Thank much #StandWithHongKong . folded\_hands Although don't know actually plans, still thanks UK helps. #newwestberlin #5De mandsNot1Less #BoycottChina

Examples of observations with both thanks & folded\_hands for the US are below.

Heartfelt thanks standing solidarity Hongkongers folded\_hands Pls consider implementation #MagnitskyAct sanction CCP & HK officials violate "1country2systems" & "high autonomy" promises Joint Declaration fire Let JUSTICE done UK #StandwithHongKong

\_nozomi Thank much US UK,thanks always supportive us red\_heart folded\_hands light\_skin\_tone #StandWithHongKong

As hongkonger, cannot save home without help US. We believe democracy freedom important everywhere world. HKers appreciate please see help ~ please save US HK folded\_hands light\_skin\_tone #freehongkong #thankTrump #thankUS

The final notes for Twitter EDA are as below:

- **Positivity** | Thanks much / support, Free World, help & folded\_hands ( 🙏 )
- **Negativity** | Evil CCP, Police Brutality, Fighting & Holdings Hands ( 🚫 )
- Mentioned in both | Human Rights & freedom press / democracy
- There appears to be a **Pro-China** or **Pro-HK split**
- Investigated links to **geopolitical entities** as well as potential sarcasm directed towards them
  - Positive mentions to the US while mentions towards China may be sarcastic

### 3.0 Model Description

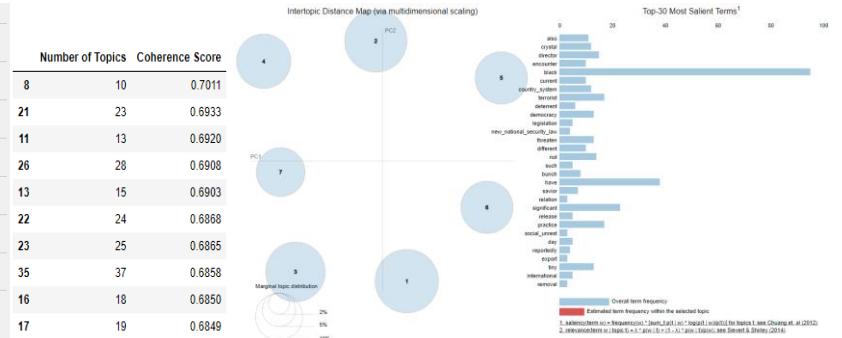
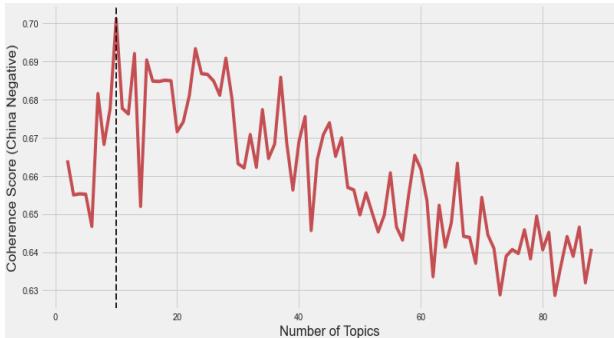
The modelling approach for Twitter & News (both Titles & Articles) are ubiquitous; they both use Latent Dirichlet Allocation (“LDA”) to classify text in the Tweets & News to particular topics. Before they got to LDA, they went through another “cleaning” process; some of which is repetition albeit it was done for reassurance. The **pre-LDA process is outlined below** with examples for each:

- **Lemmatization**
  - The process of grouping together the inflected forms of a word so they can be analyzed as a single item identified by the word’s lemma or its dictionary form.
- **Stop Word Removal**
  - Stop words are a set of commonly used words in a language. English examples are “a”, “the”, “is”, “are”, “and” etc. They don’t provide much benefit to NLP & thus removed.
- **Bag of Words**
  - A representation of text that describes the occurrence of words within a document. It’s called “bag” as it disregards the order or structure of the sentence; the occurrence of the word is what matters.
- **Bigrams & Trigrams**
  - Bigrams & Trigrams are sequences of two (2) or three (3) words respectively. It is more efficient to identify common occurrences of groups of words that appear together.

**Once completed, they were sent to LDA & subsequently the pyLDAvis** which is an interactive topic model visualization tool. pyLDAvis helps users interpret the topics in a topic model that has been fit to a corpus of text data.

LDA & pyLDAvis give an initial glimpse of how many topics there are & which words are assigned to them. To determine the optimal number of topics within the documents, a genism **Coherence Score is calculated** which measures the scores of a single topic by measuring the degree of semantic similarity between high scoring words in the topic. In other words, **a set of statements or facts that support each other are said to be coherent**.

Once the Coherence Score is resolute, the LDA & pyLDAvis process is done again & an investigation of each those topics was reviewed. This process is called “**Reading between the Lines**” (“RBL”) & represents the **final step**.



## **4.0 Model Findings**

### **Model Findings | News**

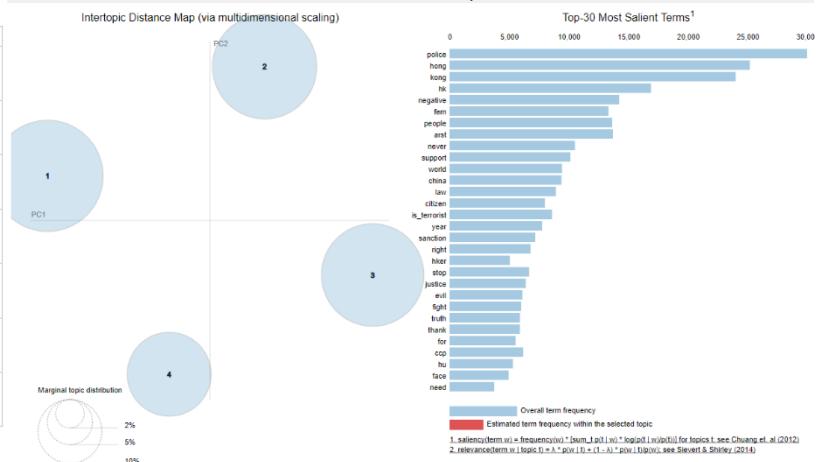
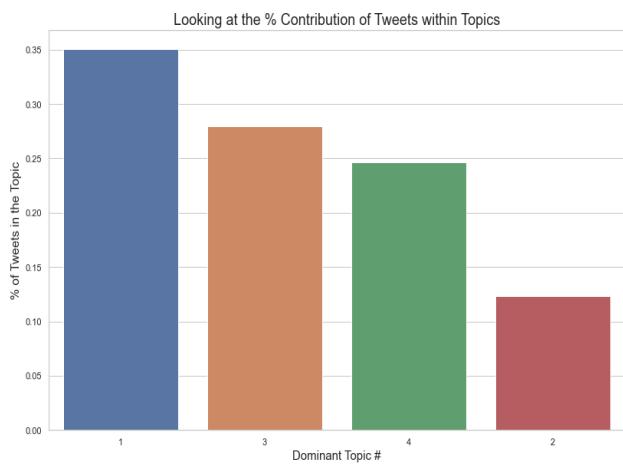
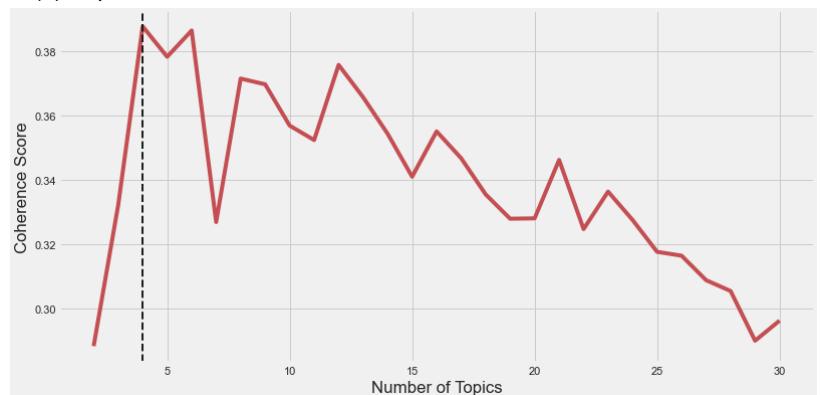
XX  
XX

## Model Findings | Twitter

The Coherence Score on the 163,699 Tweets presented four (4) Topics; we will outline what was seen:

- The **sizes** look to be **reasonably similar**
- The **distance between the center** for all indicates they are **distinct**
- The **top words & trigrams to be the following for each topic** with supporting potential rational:
  1. police, arst, law & is\_terrorist
  2. negative, people, never, world & China
  3. support, justice, evil, fight, truth & thank
  4. need, time, protest & puppet

The contribution for each topic ranged from ~15-35%.



The real meaning of what these top words convey does not present itself clearly until we dig deeper. **The RBL section** is just that, **we read the full sentences**; in this case, Tweets.

We could not read all the Tweets so we plotted it on a histogram by their contribution & took a section of the top & bottom of the contribution percentages. This works off a modius operandi that if the top & bottom contributors of the topic are aligned, then the middle will most likely will be aligned as well.

The RBL is truly a fraction of the total number of Tweets in each topic. Topic 1, for example, had 57,394 Tweets in its set while 80 Tweets which were moved to RBL; thus 0.01% of the set.

## Twitter RBL | Topic 1

A more detailed breakdown of the below can be found in [Appendix IX](#).

The **summary is highlighted below:**

- A notable amount of appreciation for support from the world for helping fight for freedom, democracy & human rights
- Mentions of the free world which may be classified as the UK, Australia, Taiwan & the USA with personal mentions towards:
  1. rep riggleman
  2. sec pompeo
  3. Aderholt
  4. rep gooden
  5. sen haley

81436 thank rep riggleman hong kong aspiring predatory fascistic s statement intent get away atrocities one stop anymore the socialimperialisticis waging war free world amp theus

104258 britian showing world fulfilling historical responsibility hong kong showing world fighting must safeguard democracy human rights  
104768 thank senators while hong kong people greatly appreciate us providing refuge place earth truly free long exists aspiring predatoty nationalistic tyrranical cc p hk frontlineis to

81431 thank sen haley hong kong aspiring predatory fascistic s statement intent get away atrocities one stop anymore the socialimperialisticis waging war free world amp theus

30934 thanks sec pompeo the solidarity among freedomloving nations support hong kong heartening lets stay strong keep fighting triumph democracy comes

## Twitter RBL | Topic 2

A more detailed breakdown of the below can be found in [Appendix IX](#).

The **summary is highlighted below:**

- A sense of comradery towards their fight
- Conveying a urge to not give up on their fight
- The usage of positivity to achieve the aforementioned goals:
  1. Candle
  2. musical\_notes / sing hallelujah
  3. folded\_hands
  4. Different configurations of people holding hands

8337 \_kwai i really love song loudly\_crying face folded\_hands hkers inexhaustible patriotism energy fire unite countries world globe\_showing\_americas one double\_exc lamation\_mark may negative bless hk hkers folded\_hands two\_men\_holding\_hands man\_and\_woman\_holding\_hands two\_men\_holding\_hands man\_and\_woman\_holding\_hands two\_men\_holdi ng\_hands man\_and\_woman\_holding\_hands two\_women\_holding\_hands two\_men\_holding\_hands man\_and\_woman\_holding\_hands two\_men\_holding\_hands man\_and\_woman\_holding\_hands

64655 musical\_notes sing hallelujah lord sing hallelujah lord musical\_notes

35663 \_news \_hker no matter evil no matter brutal still fighting justice liberty amp democracy flexed\_biceps dark\_skin\_tone angry\_face giving up

65804 \_spalding \_hker no matter evil no matter brutal is giving up still fighting justice liberty amp democracy pls help

## Twitter RBL | Topic 3

A more detailed breakdown of the below can be found in [Appendix IX](#).

The **summary is highlighted below:**

- Tweeters vocalize their negative opinion towards what they believe China's goal is; notably:
  1. Treat HK the way they believe China treats Tibet & Xinjiang (the Uighurs)
  2. Limitations towards freedom & Press Freedom
  3. The suppression of democracy
  4. Political suppression via arresting protestors
  5. Authoritarian rule & tyranny
- Belief that the HK police are puppets

92563 i spotted van parked outside chinese embassy portland place london afternoon communist china cruel humans tibetans uighurs hongkongers dog cruelty china matter  
s too

29611 how police broke media office morning they destroying reputation freedom speech makes public lives white terror more people deserve sanctioned

35711 hong kongs rule law nonexistent the city totally seized authoritarian rule now govt everything silence democracy activists hk always asset globe us

26508 the enactment tool ccp supress freedom press hong kong it shows ccp afraid power press put white terror hk hk police puppet ccp \_is\_terrorist

## Twitter RBL | Topic 4

A more detailed breakdown of the below can be found in [Appendix IX](#).

The **summary is highlighted below:**

- Confrontations with the police with notable mentions of
  1. People being arrested & sentenced
  2. Hurting police with a laser pointer
  3. Beatings at various locations
  4. Assaulting a police officer by kicking their back
  5. A lady carrying a US flag in a shopping mall
  6. Possession of kg ketamine
- Confrontations with legal authorities with charges of:
  1. Attempting to split the nation via posts on Facebook
  2. Banning of books
  3. Possession of a flick knife deemed to be an offensive weapon
  4. Gatherings banned with mentions of Tiananmen Square
  5. A judge sentenced an old teacher for assaulting a police officer
- Belief that China is working with the Triads to form a "Chinese secret police"
- The protestors belief that the media confuses their intentions
  1. They believe they aren't asking for violence; they are asking for justice & the protests are one way to vocalize this intent for making a change

862      a lady sentenced jail months assault police laser pen no laser pen found arrested popo submit another laser pen exhibit pro ccp negative judge still accept ev  
idence \_is\_terrorist

6945      similar ccp lies tiananmen massacre hongkong government try fraudulent history incident police disappeared hours let gangsters attack protesters

42111      months ago riot police beaten passengers prince edward station arbitrarily it suspected passengers beaten death related cctv disclosed till nw

67433      this lady carried us flag shopping mall n dragged n arrested by many riot po\_ice here obviously nation flags r

45534      illegal thing say also group students arrested charge attempting split yhe nation posting facebook post

69576      on contrary hkgov made books famous now more people going find ways read them like mainlanders used come hk read banned books

85428      a uni student possession pipe month sentence a girl related possession kg ketamine walk away with absolutely nothing

72368      oh liberate hk revolution time liberate hk rev  
olution time

## 6.0 Reading Between the Lines

We will outline the key findings for both separately & then finish with an overall summary of the Sentiment & Topic Analysis for the HK NSL found in the thirty (30) News Articles & 163,699 Tweets.

**News** | The findings for each of the four (4) categories on the equally split News is below:

- **China Negative** | 10 Topics
  - The violence of protestors
  - Foreign interference in China & HK's internal affairs
  - A misunderstanding of the pure intent of the National Security Law by the West
  - China's firm belief that the National Security Law will lead to long term stability & prosperity in HK
    - *Which China believes the West perceives this as incorrect, ignoring or misunderstands it*
- **International Negative** | 10 Topics
  - Their belief of the Law's infringement on Common Law
  - Their credence towards the vagueness of the language used in the Law
  - Their position on the Law's potential ramifications
    - *They outline China's opinion as well*
  - The actions & tactics of police & protestors respectively
- **China Positive** | 11 Topics
  - The Law is for safeguarding the majority of HK people
  - Although some of the international community are misinformed & making false accusations
    - *Carrie Lam, HK business & 53 countries are in support of the Law*
  - The HK people are open to the Law & celebrating it
    - *They outline China's opinion as well*
- **International Positive** | 17 Topics
  - Dominated with mentions of China's opinion & words which may be classified as Positive
  - Support for the Law from HK's legislator & Public
  - China's diplomatic ties with Japan have improved
  - Solidarity amongst protestors

We will **outline the four (4) sub-categories** all together **in a few sentences** for each:

- **China Positive**
  - China believes that the NSL will safeguard the majority of the people & bring long-term prosperity to HK as well. The HK govt & businesses support the Law as well as the HK people who are celebrating it
- **China Negative**
  - The International community both misunderstands the purity of the Law & are interfering in China's internal affairs
- **International Negative**
  - They believe the Law infringes on Common Law, is written in a vague enough way to allow Beijing to manipulate the Law in a way that represents their best interests while infringing on the freedom of the HK people
- **International Positive**
  - The International News mentions China's position & a great number of positive words are used; thus, they will come up as a positive sentence albeit are not interpreted as positive when reading. The objectively positive comments are mentions of the protestor's solidarity as well as mentions of the HK govt's support for the Law

Finally, **a few sentences for the two (2) main categories**:

- **China News** believes the Law will safeguard the majority of the people & bring long-term prosperity to HK; these are things the HK govt, HK businesses, & 53 countries support; as well as the HK people who are celebrating it. The West misunderstands China's intent & China wishes they stay out of China's internal affairs; i.e., not the affairs of the West
- **International News** believes the Law infringes on Common Law, is written in a vague enough way to allow Beijing to manipulate the Law in a way that represents their best interests while infringing on the freedom of the HK people. Aside from mentioning China's position & the HK govt's support for the Law, there was not much positivity towards the Law itself

**Twitter** | The 163,699 Tweets fell into four (4) categories listed below:

- Appreciation for the support coming from abroad
- Comradery towards their fight against the law & perceived ramifications of the NSL
- Distrust towards China & their perceived intentions for the law
- Dissatisfaction of how the police & the HK govt are handling it

Finally, a few sentences to summarize the Tweets:

- Tweeters believe that China may one day treat HK in a way similar to Tibet, Xinjiang (the Uighurs) & the Tiananmen Massacre; which they don't speak positively about. To mitigate this perceived risk, they appreciate the support of the free world, which may be classified as the UK, Australia, Taiwan & the USA. The Tweeters believe they should band together in unity to overcome this perceived risk

### The International News & Tweets Connection

As mentioned in the beginning of this report, the **Tweeters are expected to be almost exclusively from those outside of China**. Their alignment with their categories & those seen with the International News was expected & outlined below:

- They **both mention the comradery of the protestors** in a positive manner
- Believe that **the NSL infringes on Common Law & will infringe on the freedom of the HK people**
- **Tweeters explicitly express distrust towards China** & their intentions of the NSL while the International **News alludes to this idea** by mentioning the vagueness in which it was written
- Tweeters mention the protestors engagement with police a notable amount more than the International News

### Reading Between the Lines for Both

The narrative which connects them all is:

- China believes that the NSL will safeguard the people of HK & bring long-term prosperity them while conversely the International community, including the Tweeters, believe the NSL infringes on Common Law & will infringe on the freedom of the HK people
- China recognizes the International position but believes they have a misunderstanding of the pure intent of the NSL. International News makes mention of China's position albeit believe China incorporated vague wordage to potentially serve other intentions. The Tweeters also believe that certain media groups misunderstand their intentions & believe protests are the only way to the society they wish to have.
- China believes that the International community should stay out of Chinese affairs & the Tweeters are thankful for the engagement / support from the International community
- China's position on police engagement is justified given the actions of the protestors. International News highlights the engagements while Tweeters convey a belief that the police are going above their call of duty

Please note, this Data Science project was completed in September 2021 but the data is from June - Aug 2020. Much may have changed but that would fall outside the scope of this project.

## 7.0 Next Steps

Inflation is a difficult & highly disputed financial beast but the closer you get to taming it your eyes will open wider.

Throughout the process there were a number of things I either wanted to change or add. I will start with some Variables which were not included due to financial & time constraints:

- Steel
  - o I could only get it back to 2008
- Gasoline
  - o I could only get it back to 2005
- Growth of M2
  - o I put it aside to mitigate any overlap with M2 Velocity
- US Wages Hourly Earnings
  - o Limited data as well
- US Dollar Index: Broad, Goods and Services
  - o Only goes back to 2006 ( discussed below )

Further to these, further consideration may be applicable to the below:

- Reassess the Variables which were chosen in the SS & LG divide; discussed in Pre-processing #5 above
- Although scraping on Variables was investigated ( see "Scraping the Variables as well ( individually )" in the Data Wrangling & EDA [link here](#) ), a pivot to Winsorizing way present better results
- Develop a model to predict Wages CPI itself in order to remove ourselves from the US gov't's reporting
- I believe the US Dollar Index Variable ( DXY ) does not correctly address the situation of the United States either Imports or Exports Inflation ( usually the former ). This may be because the DXY's weighting does not correctly align to the US's trade. In short, it's a weighted geometric mean of the Eurozone ( EUR ), Japan ( JPY ), the United Kingdom ( GBP ), Canada ( CAD ), Sweden ( SEK ) & Switzerland ( CHF ) but that does not correctly align with the US's trade with the world

There are a number of Variables to take into consideration but for now I would recommend keeping an eye on Oil.

## Appendix I

	file_name	date	source	country	local	title		article	word_count
0	bbc_1	2020-06-30	BBC	UK	Intl	Hong Kong security law: What is it and is it w...	China has passed a wide-ranging new security l...	15490	
1	bbc_2	2020-07-01	BBC	UK	Intl	Hong Kong's new security law: Why it scares pe...	China has introduced a new national security l...	6021	
2	nytimes_3	2020-06-28	NY Times	US	Intl	What China's New National Security Law Means f...	Chinese lawmakers have approved a national sec...	4028	
3	forbes_4	2020-07-01	Forbes	US	Intl	Hong Kong Makes First Arrests Under Beijing's ...	Share to FacebookShare to TwitterShare to Link...	4408	
4	insider_5	2020-08-02	Insider	US	Intl	Teenage arrests, blank protest signs, and a ke...	Thursday, July 30, marked one month since Chin...	6326	
5	abc_6	2020-07-01	ABC News	US	Intl	What's in Hong Kong's new national security la...	The full details of the controversial national...	2420	
6	cnn_7	2020-07-03	CNN	US	Intl	Hong Kong's security law could have a chilling...	London (CNN Business)Hong Kong insists its vib...	5683	
7	nypost_8	2020-07-01	NY Post	US	Intl	Hong Kong police arrest over 300 in first prot...	Hong Kong police arrested more than 300 people...	1712	
8	nippon_9	2020-08-14	Nippon	Japan	Intl	Hong Kong's Security Law: How Should Japan Res...	Growing Private and Governmental Support for H...	6367	
9	bbc_10	2020-06-30	BBC	UK	Intl	Hong Kong security law: Anger as China's Xi si...	The UK, EU and Nato have expressed concern and...	11881	
10	bbc_11	2020-06-30	BBC	UK	Intl	Hong Kong security law: Minutes after new law...	On Tuesday morning, the news started to break ...	4712	
11	bbc_12	2020-07-08	BBC	UK	Intl	Hong Kong security law: Beijing security offic...	A new national security office has been offici...	3238	
12	bbc_13	2020-07-05	BBC	UK	Intl	Hong Kong security law: Pro-democracy books pu...	Books by pro-democracy figures have been remov...	4325	
13	bbc_14	2020-06-30	BBC	UK	Intl	Hong Kong security law: Life sentences for bre...	People in Hong Kong could face life in jail fo...	6500	
14	theDiplomat_15	2020-07-01	the Diplomat	US	Intl	Hong Kong Through Water and Fire	Credit: AP Photo/Kin CheungAdvertisementIt has...	9868	
	file_name	date	source	country	local	title		article	word_count
15	peoples_daily_1	2020-07-03	People's Daily	China	Local	Newly passed national security law a seawall s...	The Law of the People's Republic of China on S...	3486	
16	peoples_daily_2	2020-07-02	People's Daily	China	Local	Hong Kong national security law helps ensure I...	On June 30, the Standing Committee of the 13th...	8028	
17	peoples_daily_3	2020-07-02	People's Daily	China	Local	Law and order dawns in Hong Kong as new law ta...	-- The festive mood across Hong Kong on the 23...	9627	
18	peoples_daily_4	2020-07-02	People's Daily	China	Local	National Security Law paves the way for more p...	Hong Kong, the popular Asian city, returned to...	9498	
19	peoples_daily_5	2020-07-02	People's Daily	China	Local	HKSAR national security law to put HK back on ...	Zhang Xiaoming, deputy director of the Hong Ko...	6744	
20	globaltimes_6	2021-04-19	Global Times	China	Local	Biased tone, misinformation 'major mistakes' b...	Victor Gao (center), chair professor of Soocho...	16593	
21	globaltimes_7	2020-07-06	Global Times	China	Local	London's measures meaningless, 'bluff rather t...	A Huawei store stands next to a Globe Telecom ...	2960	
22	globaltimes_8	2020-06-30	Global Times	China	Local	National Security Law to protect HK democracy...	Hong Kong citizens on Tuesday gather to suppor...	3711	
23	min_foreign_aff_9	2020-07-20	Ministry of Foreign Affairs	China	Local	Ambassador Liu Xiaoming Gives Exclusive Live I...	On 19 July 2020, H.E. Ambassador Liu Xiaoming ...	5324	
24	min_foreign_aff_10	2020-07-01	Ministry of Foreign Affairs	China	Local	Foreign Ministry Spokesperson Zhao Lijian's Re...	In recent years, the US government has placed ...	18122	
25	xinhua_11	2020-05-16	Xinhua	China	Local	IPCC's report comprehensive, objective, fact-b...	Video PlayerClose\n\n\nCarrie Lam, chief exe...	806	
26	xinhua_12	2020-07-01	Xinhua	China	Local	China adopts law on safeguarding national secur...	Video PlayerClose\n\n\nLi Zhanshu, chairman...	4315	
27	theStandard_13	2020-08-26	the Standard	China	Local	State media says Hong Kong has a toxic 'Apple'	State media People's Daily today published an ...	3079	
28	china_daily_14	2020-06-20	China Daily	China	Local	Wide support for proposed national security law	Residents sign to support the national securit...	4294	
29	hkgovt_15	2020-06-30	Hong Kong Gov't	Hong Kong	Local	CE welcomes passage of The Law of the People's...	In response to the passage of The Law of the P...	6143	

([back](#))

## Appendix I (cont.)

Column	Description
0	file_name
1	date
2	source
3	country
4	local
5	title
6	article
7	word_count

([back](#))

## Appendix II

	hash	created_at	username	tweet	language	replies_count	retweets_count	likes_count	reply_to	hashtags
7	#FightForHongKong	2020-08-23	cipopolitik	China will never stop. #HongKongProtest #Figh...	en	0	0	0	□	['hongkongprotest', 'fightforhongkong', 'prayf...
8003	#HongKongProtest	2020-06-04	paulisally	Hong Kong Police is MURDERER!!!!!! 총총 경찰은 살인범...	ja	0	0	1	□	['policebrutality', 'hongkongpolice', 'hongkon...
8503	#PrayForHongkong	2020-06-30	cartwheelwombat	Please pray for Hong Kong. You can follow him...	en	1	2	2	□	['prayforhongkong', 'standwithhongkong', 'hong...
9025	#WeAreNotAlone	2020-07-04	whats_a_gurg	There's not much more satisfying than seeing o...	en	0	0	0	□	['wearenotalone', 'questioneverysthing']
9307	#nochinaextradition	2020-05-31	techno_tech4u	'Remove China Apps' crosses 1 million download...	en	0	0	0	□	['china', 'chinaapp', 'antichina', 'nochinaext...
24003	#standwithhongkong	2020-08-17	minsuga1993309	Please #StandWithHongKong!	en	0	0	1	□	['standwithhongkong']
175007	#hkprotests	2020-08-24	juniushogodie	It's also the 1st time MTR closed its station ...	en	0	7	17	□	['hkprotests']
180023	#freehongkong	2020-08-27	pppongo	@StateDept @SecPompeo Thank you U.S so much fo...	en	0	0	0	□	['screen_name': 'StateDept', 'name': 'Departm...', 'standup4humanrights', 'standwithhongkong', ...]
260001	#hkpolicebrutality	2020-06-16	akaribug	They keep telling us to trust a liar while tur...	en	0	1	0	□	['hkpolicebrutality', 'xinjiang', 'standwithho...
280017	#hkpolicestate	2020-06-14	ylye33	A 17 years old girl was attacked by the #HKpol...	en	0	0	3	□	['hkpolice', '612protest', 'hkpolicebrutality']
288002	#hkpolicererrorism	2020-06-08	leung18188868	@foxxthepopoo @koshi2008 They are not police. ...	en	0	1	0	□	['screen_name': 'foxxthepopoo', 'name': 'ianfh...', 'hkpolicebrutality', 'hkpolicererrorism', 'hk...
284002	#HKpolicererrorist	2020-07-25	bethenewyorkers	@PENamerica Journalists are always the target ...	en	0	2	2	□	['screen_name': 'PENamerica', 'name': 'PEN Am...', 'hkpolicererrorist', 'nationalsecuritylaw']

Column	Description
0	hash
1	created_at
2	username
3	tweet
4	language
5	replies_count
6	retweets_count
7	likes_count
8	reply_to
9	hashtags

([back](#))

## Appendix III

### [ -- Non-ascii & Unclean Values | Conversion -- ]

```
df_news_title.title[27]
'State media says Hong Kong has a toxic 'Apple'

df_news_title['title'] = df_news_title['title'].str.replace(r'\'', '\\\'')
df_news_title['title'] = df_news_title['title'].str.replace(r'\\\'', '\'')

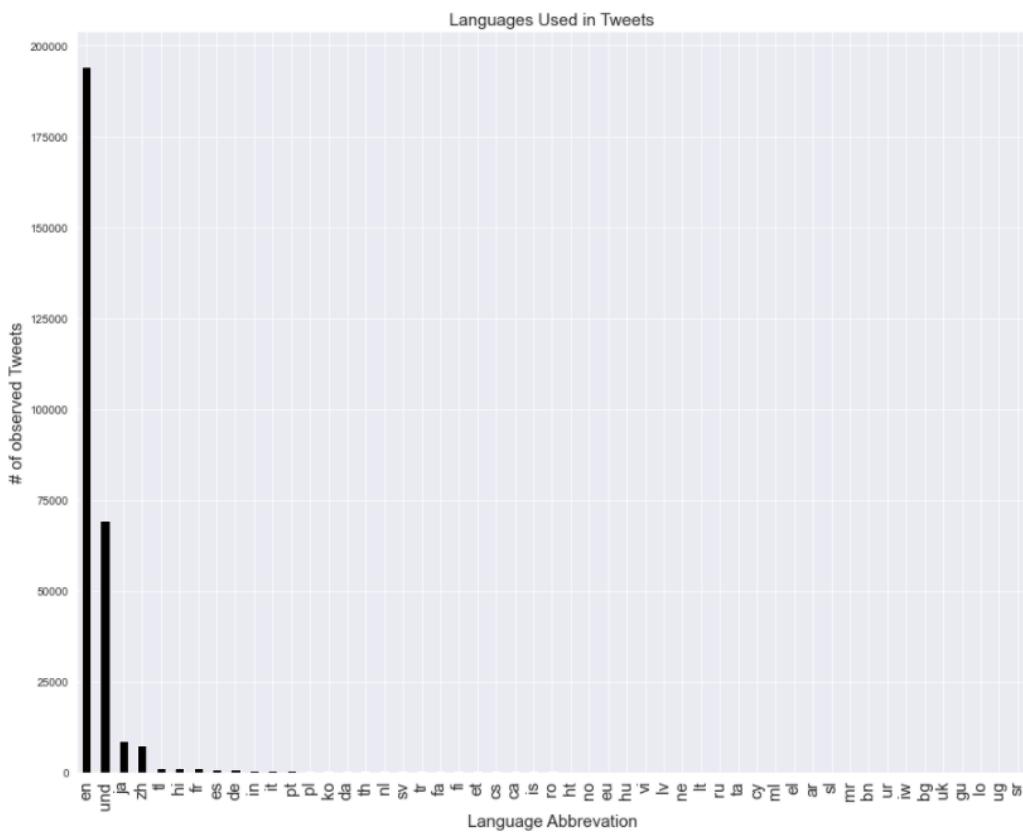
df_news_title.title[27]
"State media says Hong Kong has a toxic 'Apple'"
```

### [ -- Sentiment Intensity Analyzer, News Title's & Article's | Statistics -- ]

	comp_title	neg_title	neu_title	pos_title		comp_article	neg_article	neu_article	pos_article
count	30.000000	30.000000	30.000000	30.000000	count	30.000000	30.000000	30.000000	30.000000
mean	0.204997	0.093133	0.688767	0.218067	mean	0.338153	0.087933	0.795733	0.116567
std	0.428156	0.145032	0.150645	0.157375	std	0.897364	0.035946	0.034230	0.036774
min	-0.859100	0.000000	0.432000	0.000000	min	-0.997900	0.013000	0.710000	0.064000
25%	-0.051450	0.000000	0.593500	0.125750	25%	-0.754575	0.068500	0.772500	0.089000
50%	0.295000	0.000000	0.675000	0.197000	50%	0.979150	0.093500	0.785500	0.115500
75%	0.575000	0.198000	0.784000	0.317500	75%	0.996675	0.110000	0.822750	0.132500
max	0.848100	0.568000	1.000000	0.535000	max	0.999600	0.145000	0.860000	0.207000

([back](#))

## Appendix IV



```

hash           12      en    194087
created_at     91      und   69375
username       45111    ja    8715
tweet          221205    zh    7327
language        51      tl    1254
replies_count   217     hi    1207
retweets_count  881     fr    1046
likes_count     1110    es    908
reply_to        27555    de    892
hashtags        77715    in    507
dtype: int64
                                         Name: language, dtype: int64
                                         Name: language, dtype: float64

```

```

df_tweets_w_dup[df_tweets_w_dup['language'].str.contains('und')]['tweet']
9      https://t.co/8lgfzmGtav #lukashenko #Tikhan...
17     #FreeAgness #FightForFreedom #StandWithHongKon...
18     #Hongkongers #hongkong #beijing #china #HongK...
20     #boycottchina #hongkong #freehk #boycottmulan ...
21     https://t.co/B5KWeYZE6t #boycottchina #hongko...
...
288397  #HKPoliceState #HKPoliceBrutality #HKTotallita...
288402  @IndyWorld #HKPoliceState #HKPoliceTerrorism
288406  #香港警察 #港警是施虐者 #hkpolicebrutality #HKPoliceTerr...
288408  @SpokespersonCHN #ICantBreathe #Icannotbreathe...
288415  #HongKongProtest #StandwithHK #HKPoliceTerrori...
Name: tweet, Length: 69375, dtype: object

```

([back](#))

## [ -- Mentions to other Users, Removal-- ]

df[['hash', 'created_at', 'tweet']].head(1)			
hash	created_at	tweet	
0 #FightForHongKong	2020-08-28	@benedictrogers @NOW4humanity Thanks for speak...	
df['tweet'] = df['tweet'].str.replace("@[A-Za-z0-9]+", "")			
df[['hash', 'created_at', 'tweet']].head(1)			
hash	created_at	tweet	
0 #FightForHongKong	2020-08-28	Thanks for speaking up for us #fridaysforfre...	

## [ -- Links to Website, Removal -- ]

```
df['tweet'][9]
'Joshua Wong: 'Forms of resistance need to be fluid and flexible' #FightForHongKong #HongKongProtest https://t.co/COPnczKp4'

df['tweet'] = df['tweet'].str.replace(r'http\S+', '')

df['tweet'][9]
'Joshua Wong: 'Forms of resistance need to be fluid and flexible' #FightForHongKong #HongKongProtest '
```

## [ -- Emoji, Conversion -- ]

```
df.tweet[12]
'Heartbreaking...😢 we have to do whatever we can do to #FightForHongKong #StandWithHongKong'

df['tweet'] = df.tweet.apply(lambda x: convert_emojis_to_word(x))

df.tweet[12]
'Heartbreaking... loudly_crying_face we have to do whatever we can do to #FightForHongKong #StandWithHongKong'
```

## [ -- Profanity, Handling -- ]

```
df.tweet[58]
' CORRUPT AS NEGATIVE - [ ] #PL #PremierLeague #PremierLeagueIsCorrupt #NUFC #NUFCTakeover #Hongkongprotest #Uighurs #budfoo
tball #BudweiserUK #Nike #nikefootball #CarlsbergUK #AveryDennison #OfficialPanini #CocaCola_GB #CocaCola #barclaysuknews
#BarclaysFooty #EA #EASPORTS #TAGHeuer '
```

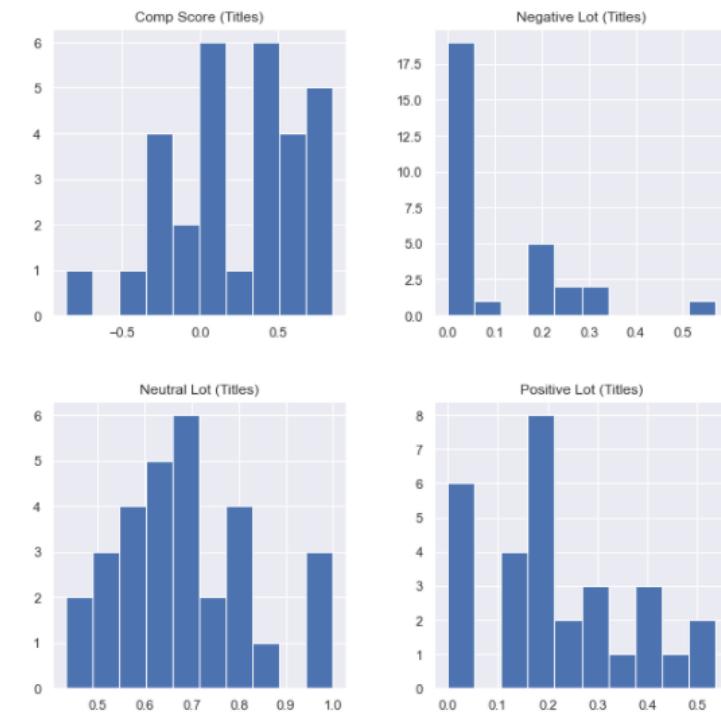
## [ -- Sentiment Intensity Analyzer, Twitter | Statistics -- ]

	comp	neg	neu	pos
count	163699.000000	163699.000000	163699.000000	163699.000000
mean	0.094896	0.102728	0.728009	0.169262
std	0.534622	0.124993	0.164952	0.161322
min	-0.990100	0.000000	0.078000	0.000000
25%	-0.361200	0.000000	0.616000	0.000000
50%	0.077200	0.065000	0.725000	0.144000
75%	0.539900	0.179000	0.835000	0.270000
max	0.998900	0.897000	1.000000	0.922000

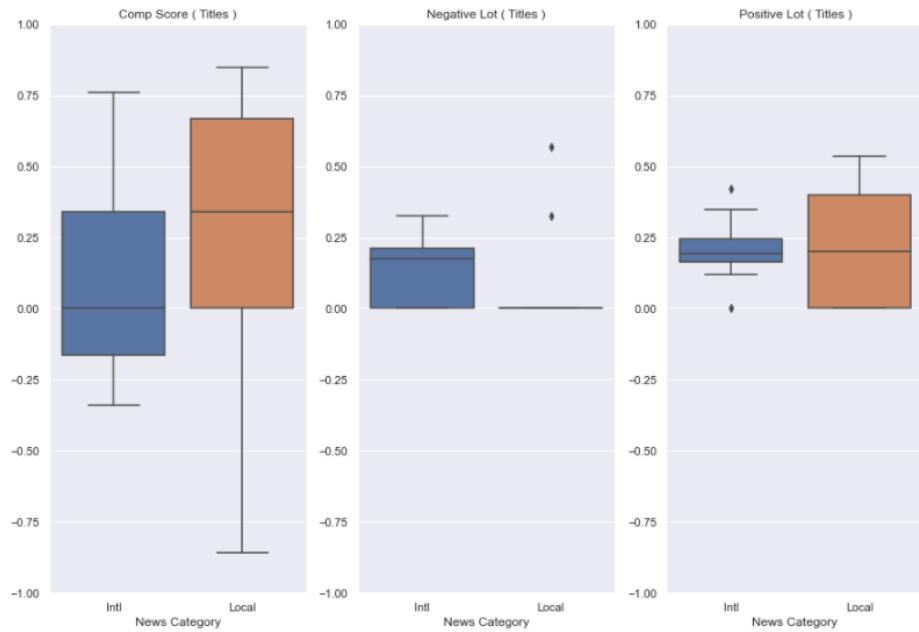
([back](#))

## Appendix V

### [ Sentiment Intensity Analyzer Histogram, Titles ]

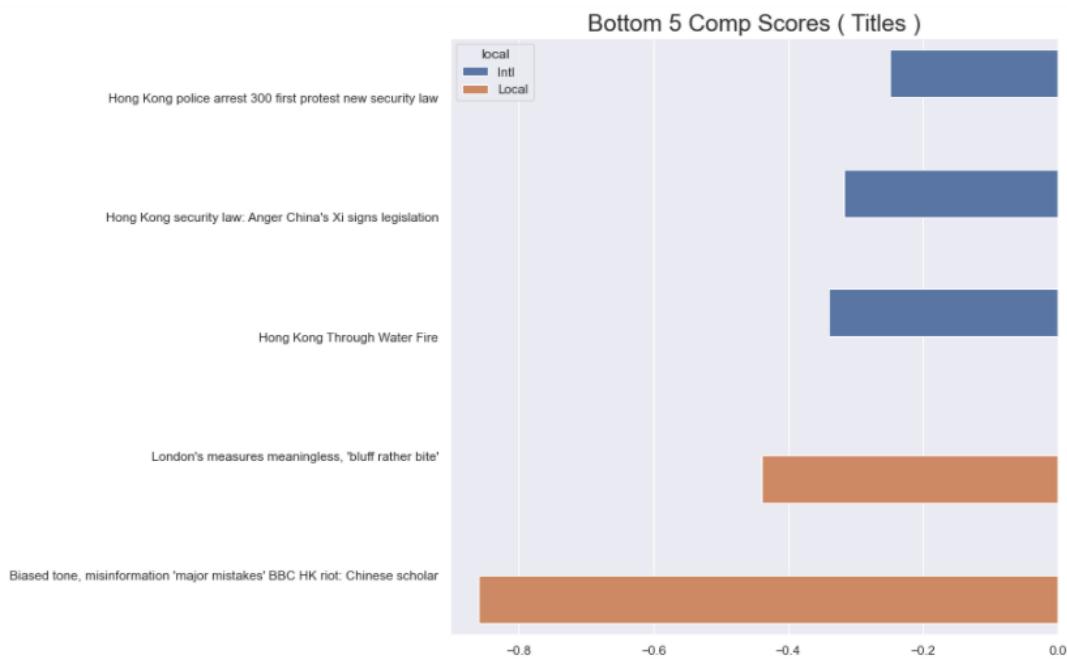
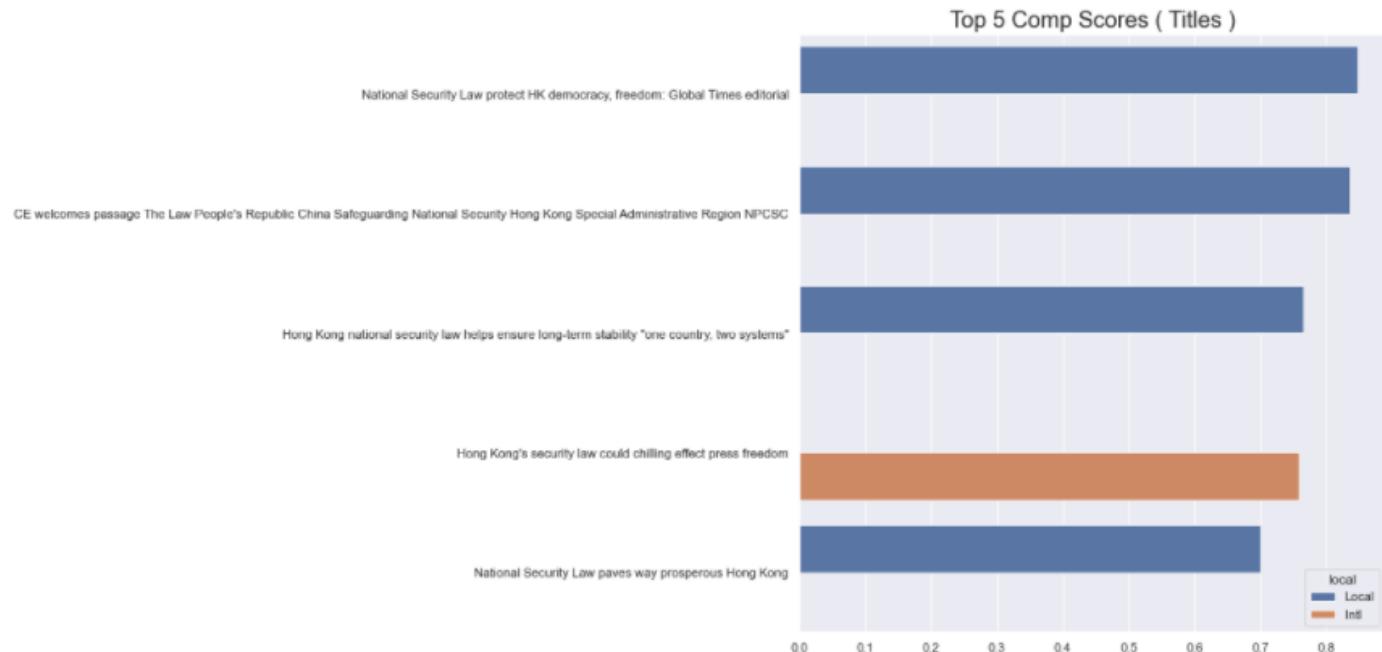


### [ Sentiment Intensity Analyzer Box Plot, Titles ]



([back](#))

## [ Sentiment Intensity Analyzer Bar Charts, Titles ]



([back](#))

## [ Word Clouds, Titles ]

Nouns in the Titles



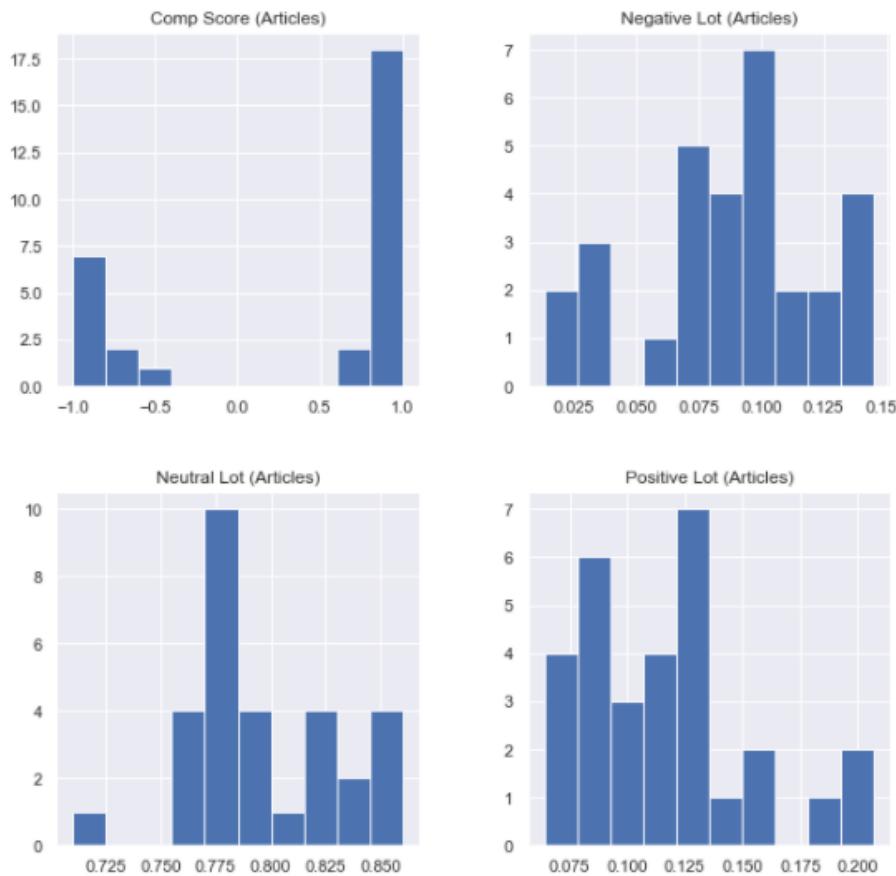
Adjectives & Verbs in the Titles



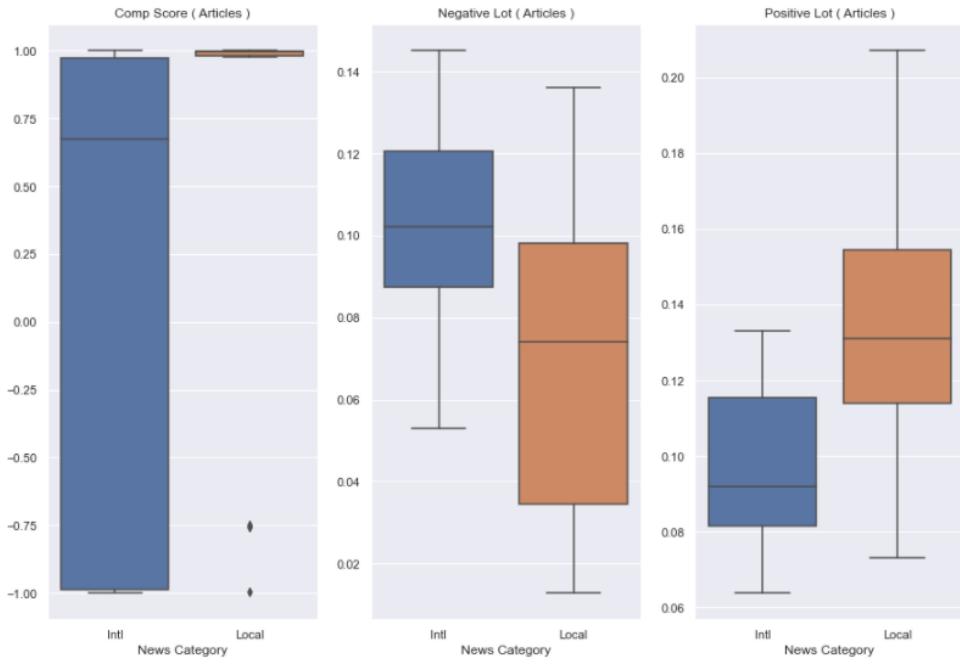
[\(back\)](#)

## Appendix VI

### [ Sentiment Intensity Analyzer Histogram, Articles ]



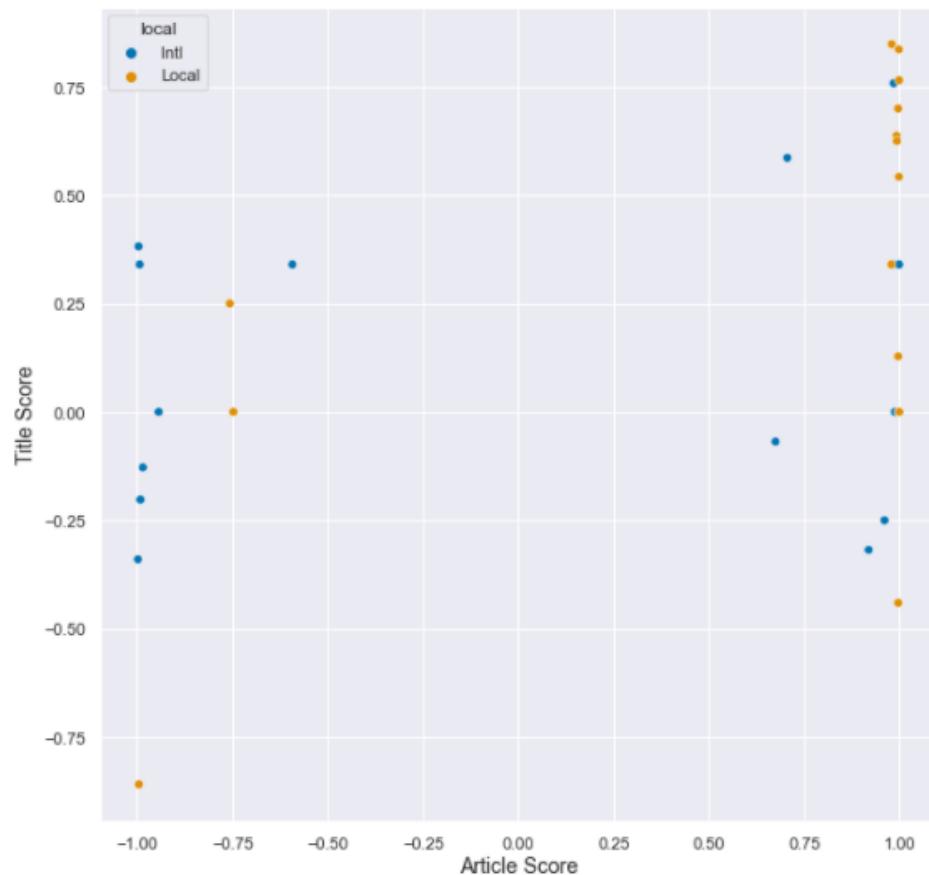
### [ Sentiment Intensity Analyzer Box Plot, Articles ]



([back](#))

## [ Sentiment Intensity Analyzer Scatter Plot, Articles ]

Comparing the Comp Scores between Articles & Titles



## [ Sentiment Intensity Analyzer Breakdown, Articles ]

20 | Articles above 50%  
10 | Articles below -50%

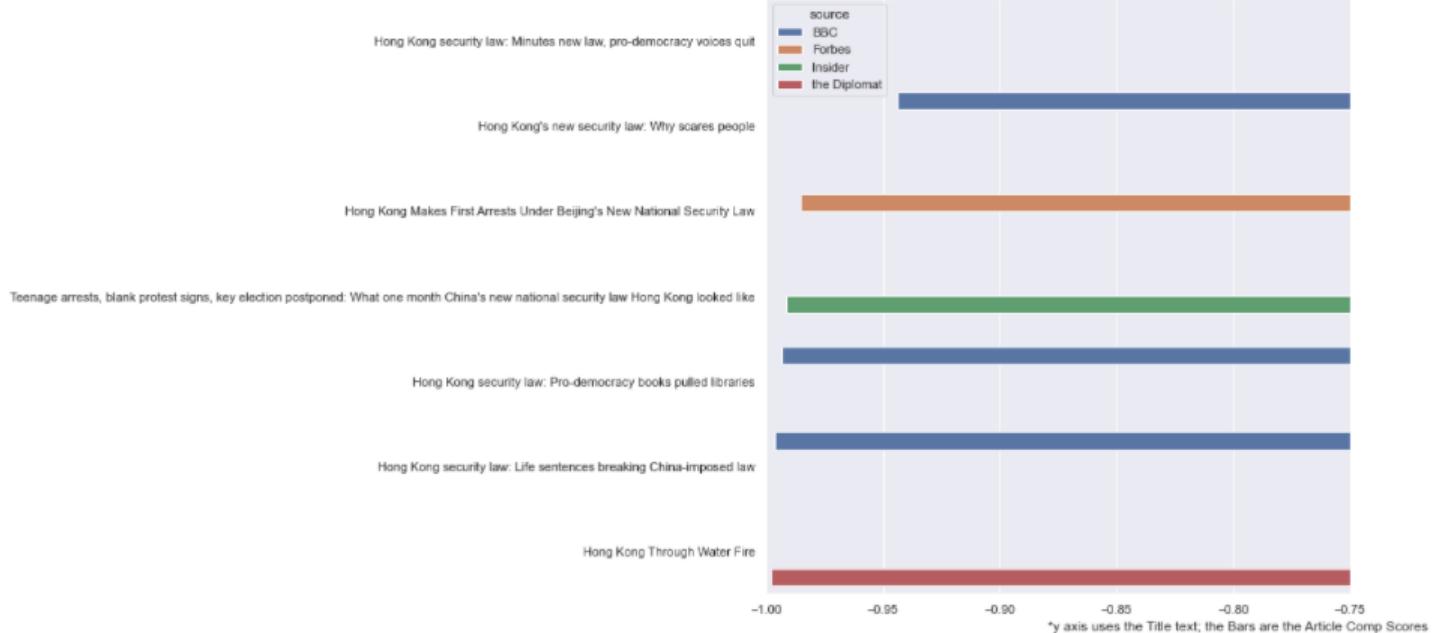
The breakdown between groups +50%:  
China | 60.0 percent  
Intl | 40.0 percent

The breakdown between groups -50%:  
China | 30.0 percent  
Intl | 70.0 percent

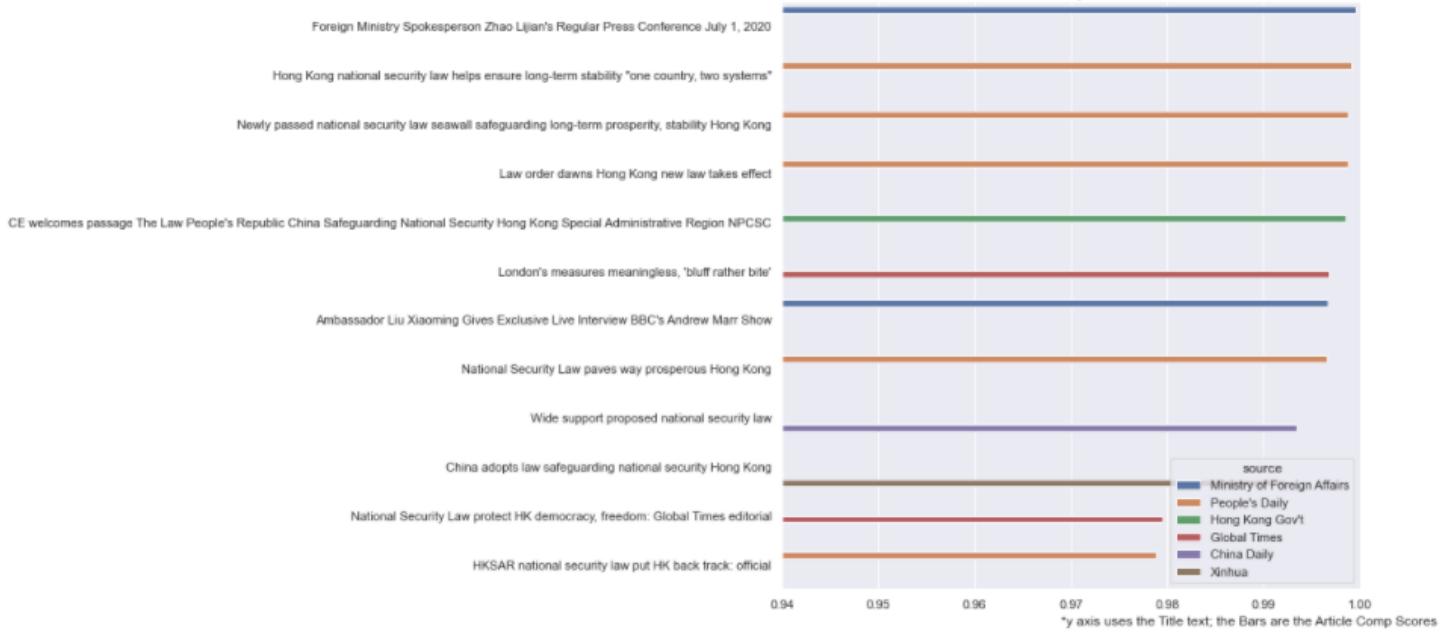
([back](#))

## [ Sentiment Intensity Analyzer Bar Charts, Articles ]

Negative Articles by Int'l Sources



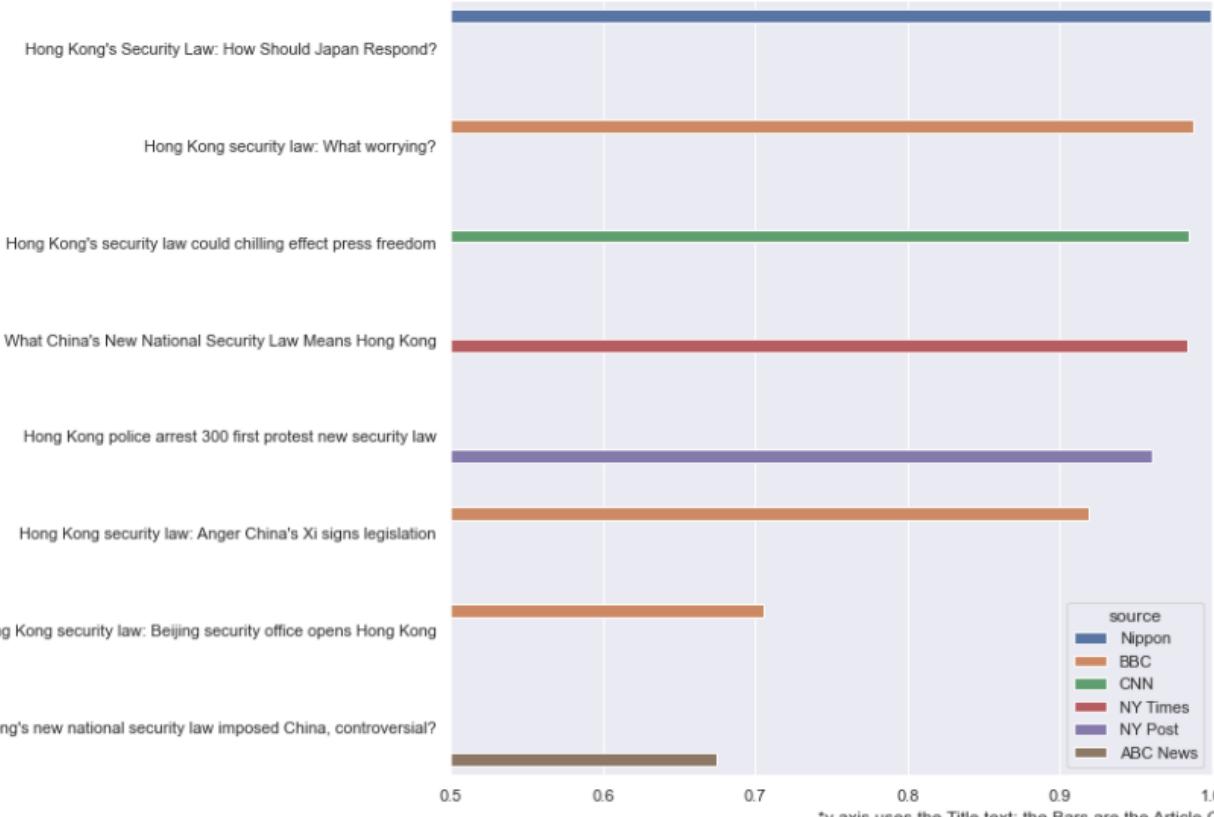
Positive Articles by China Sources



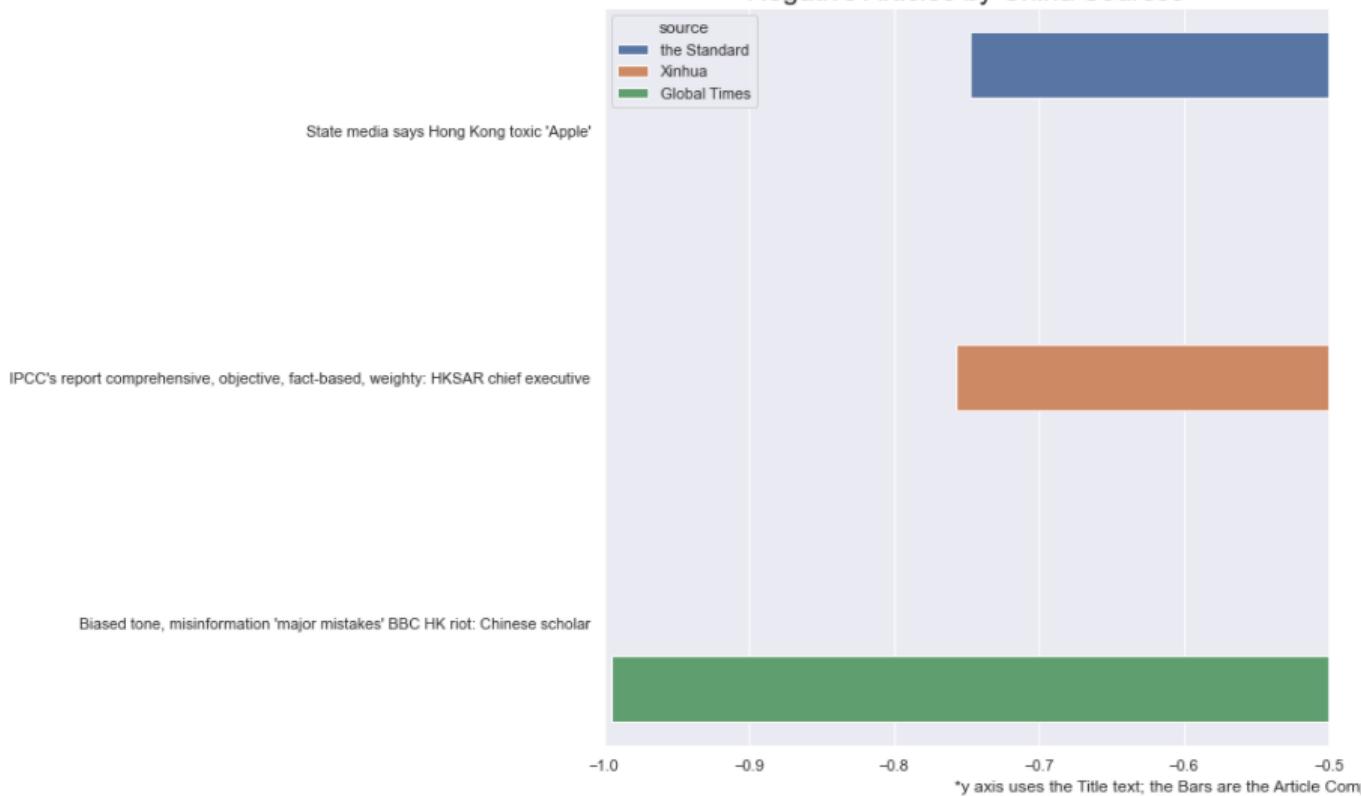
([back](#))

## [ Sentiment Intensity Analyzer Bar Charts, Articles ]

### Positive Articles by Int'l Sources



### Negative Articles by China Sources



[\(back\)](#)

## [ Word Clouds, Articles ]

### Nouns in the Articles



### Adjectives & Verbs in the Articles



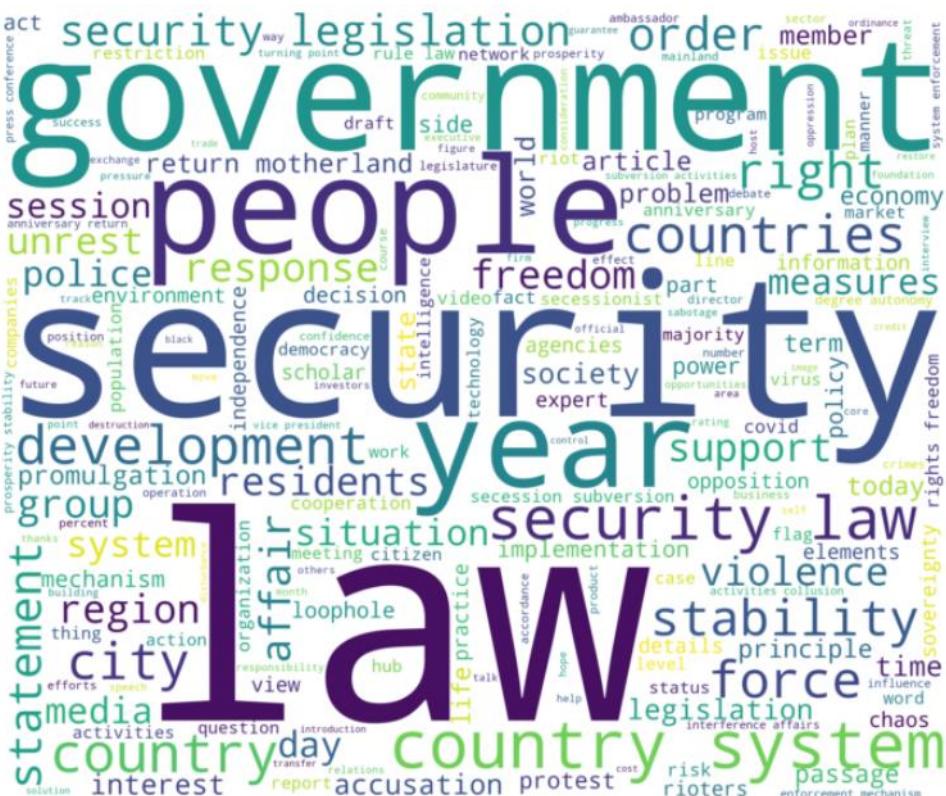
[\(back\)](#)

[ Word Clouds, China Articles ]

Adjectives & Verbs in the China-based Articles



Nouns in the China-based Articles



([back](#))

NLP Topic & Sentiment Analysis on the HK National Security Law, by Rand Sobczak Jr.

[ Word Clouds, International Articles ]

Adjectives & Verbs in the International Articles



Nouns in the International Articles

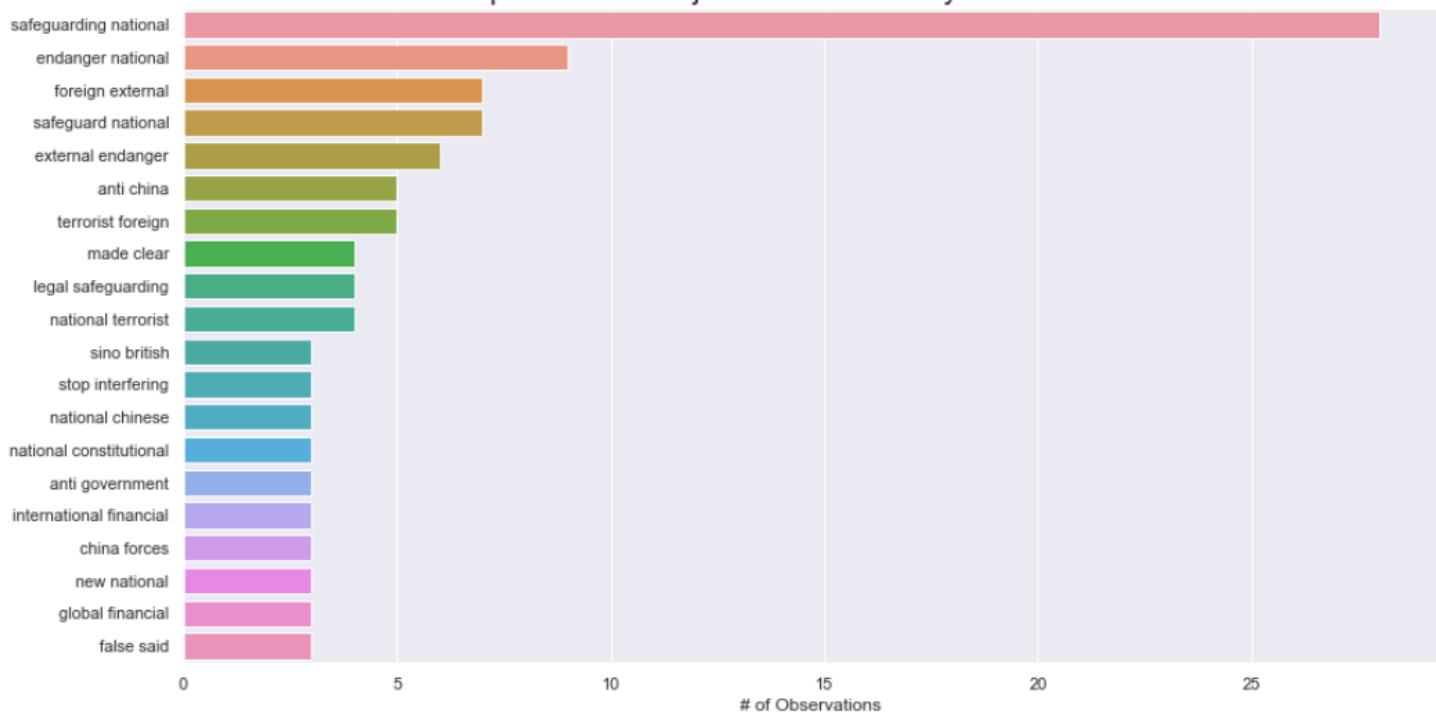


(back)

NLP Topic & Sentiment Analysis on the HK National Security Law, by Rand Sobczak Jr.

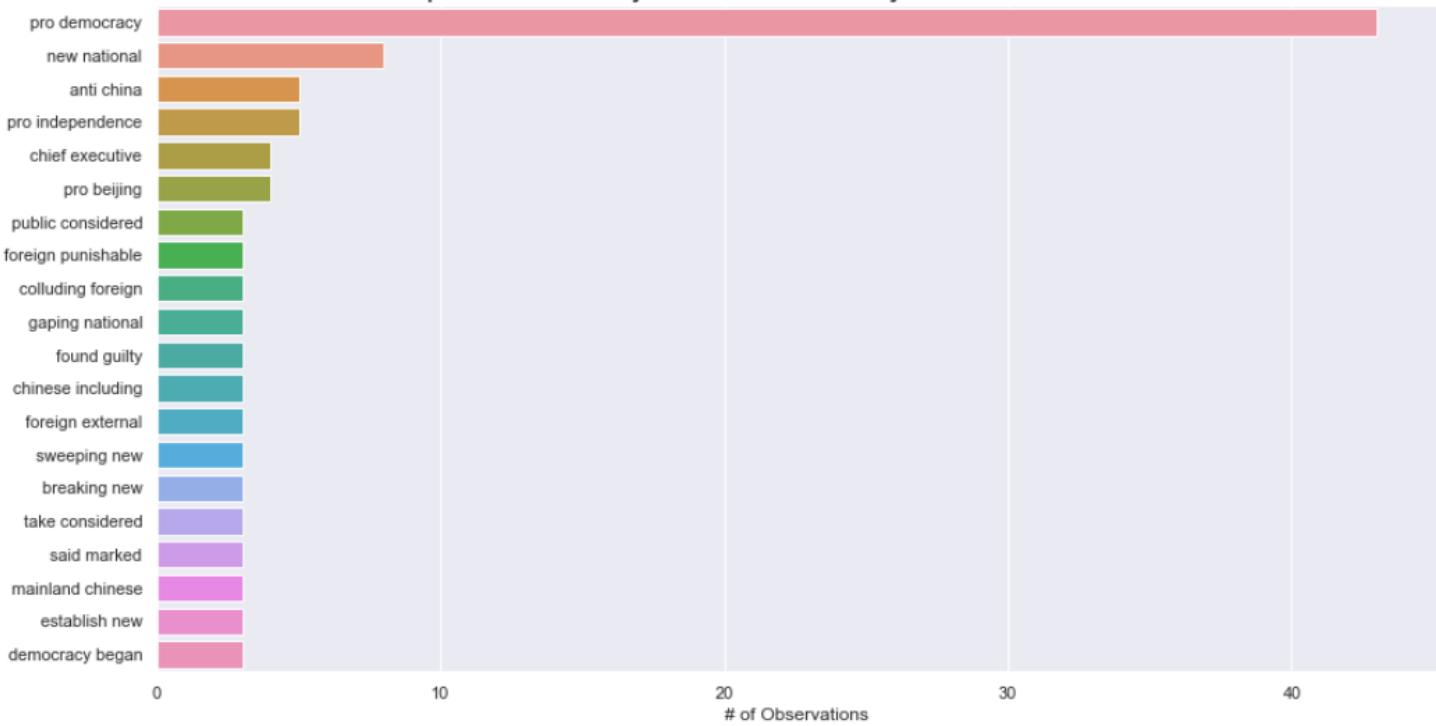
### [ Choice Words, Articles ]

#### Top 20 Choice Adjectives & Verbs\* by China Sources



\*These are words arrived through a Bi-gram

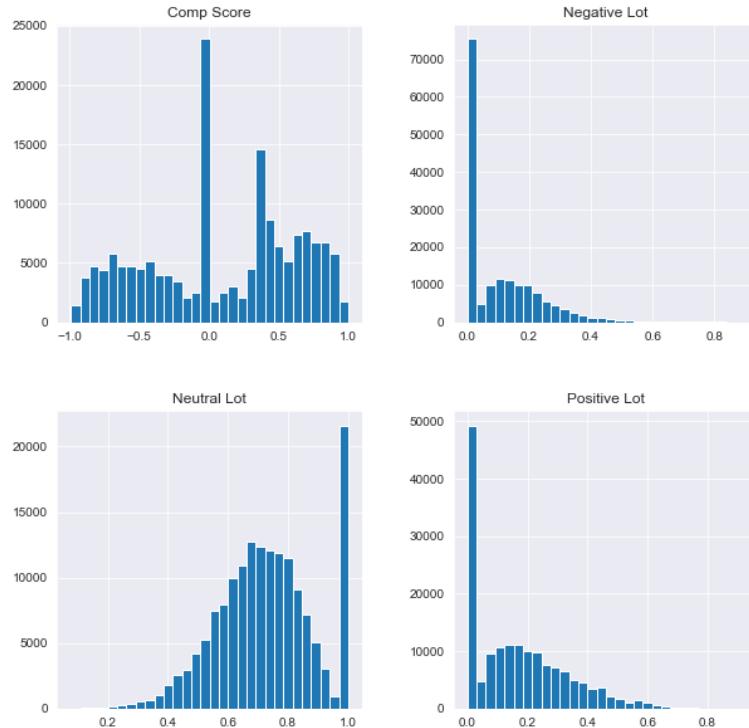
#### Top 20 Choice Adjectives & Verbs\* by International Sources



\*These are words arrived through a Bi-gram

## Appendix VII

### [ Sentiment Intensity Analyzer Histogram, Tweets ]

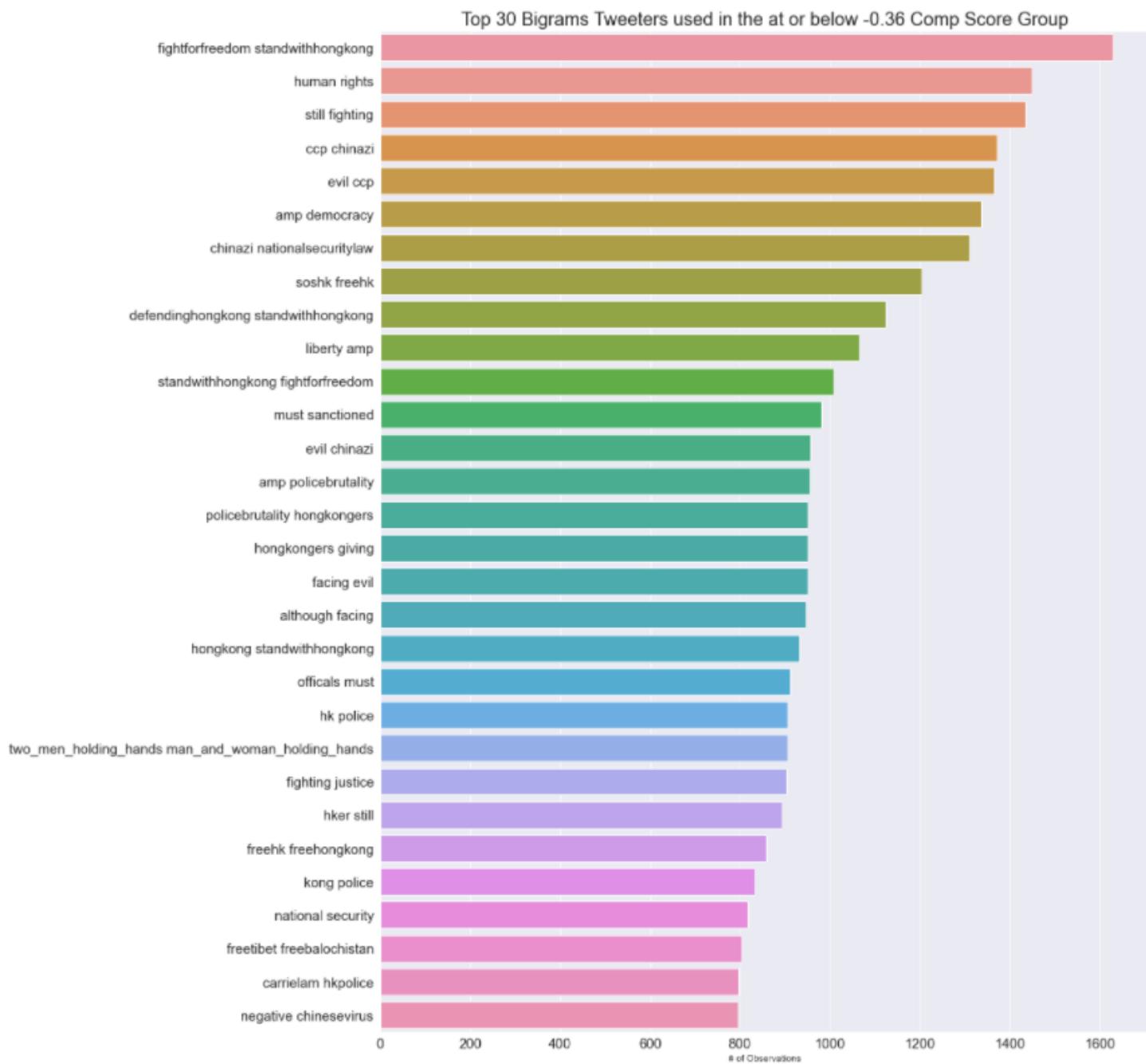


### [ Word Clouds, Tweets at or Below a -0.36 Comp Score ]



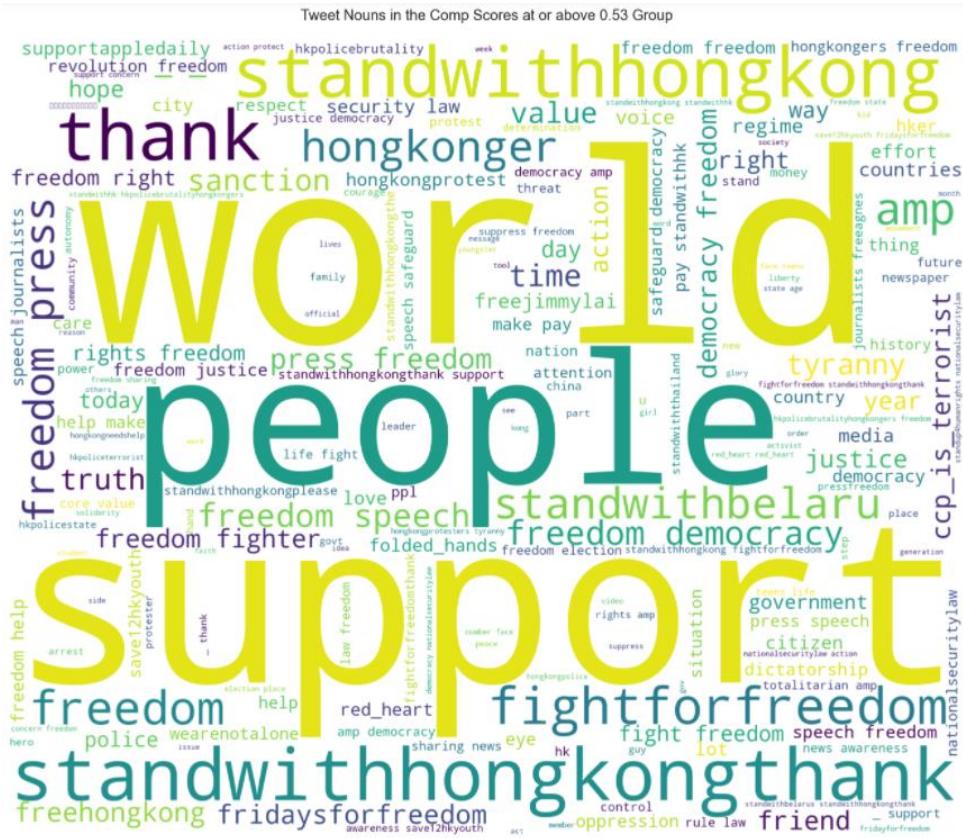
[\(back\)](#)

### [ Choice Words, Tweets at or Below a -0.36 Comp Score ]



([back](#))

[ Choice Words, Tweets at or Above a 0.53 Comp Score ]



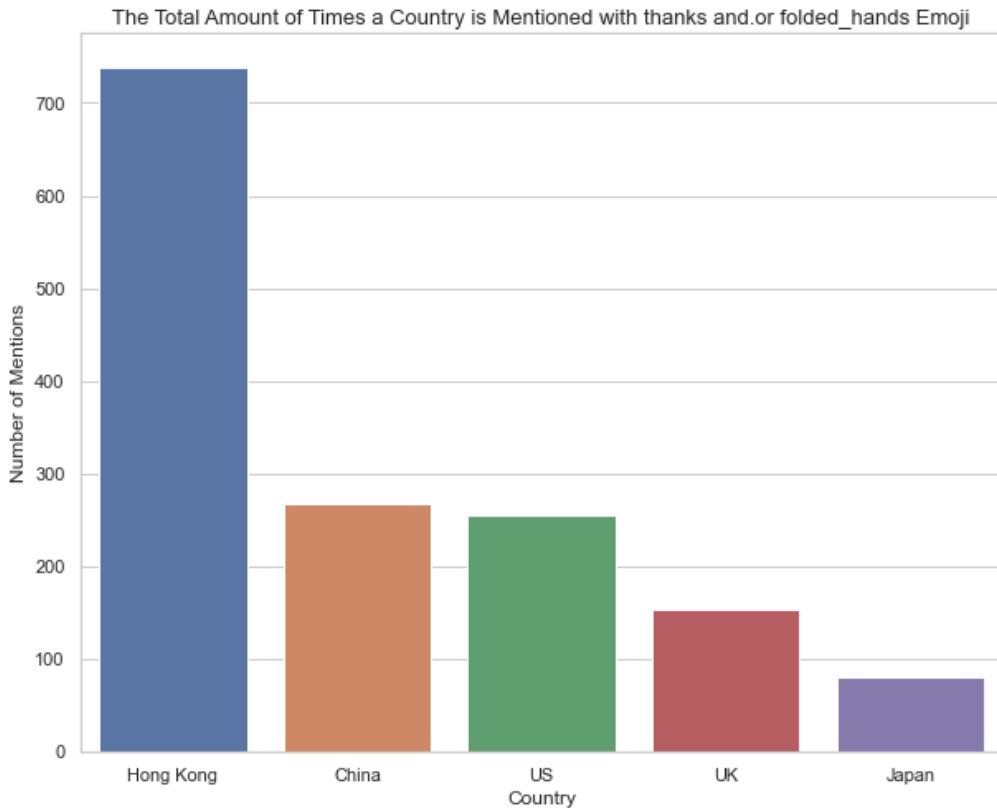
([back](#))

### [ Choice Words, Tweets at or Above a 0.53 Comp Score ]



([back](#))

### [ Bar Chart, Tweets with use of the folded\_hands Emoji ]



### [ Tweets with use of the folded\_hands Emoji & Mentions of China or USA ]

Examples of observations with both thanks & folded\_hands for China are below.

Thank #StandWithHongKong !!! We sincerely hope free world stand firm Communist China's human rights abuses Hong Kong. Many thanks folded\_hands light\_skin\_tone

Deeply touched heartfelt thanks done #HK folded\_hands light\_skin\_tone folded\_hands light\_skin\_tone folded\_hands light\_skin\_tone Always grateful international support fighting evil #CCPChina regime. #StandWithHongKong

Thank much #StandWithHongKong . folded\_hands Although don't know actually plans, still thanks UK helps. #newwestberlin #5De mandsNot1Less #BoycottChina

Examples of observations with both thanks & folded\_hands for the US are below.

Heartfelt thanks standing solidarity Hongkongers folded\_hands Pls consider implementation #MagnitskyAct sanction CCP & HK officials violate "1country2systems" & "high autonomy" promises Joint Declaration fire Let JUSTICE done UK #StandwithHongKong

\_nozomi Thank much US UK,thanks always supportive us red\_heart folded\_hands light\_skin\_tone #StandWithHongKong

As hongkonger, cannot save home without help US. We believe democracy freedom important everywhere world. HKers appreciate please see help ~ please save US HK folded\_hands light\_skin\_tone #freehongkong #thanksTrump #thankUS

([back](#))

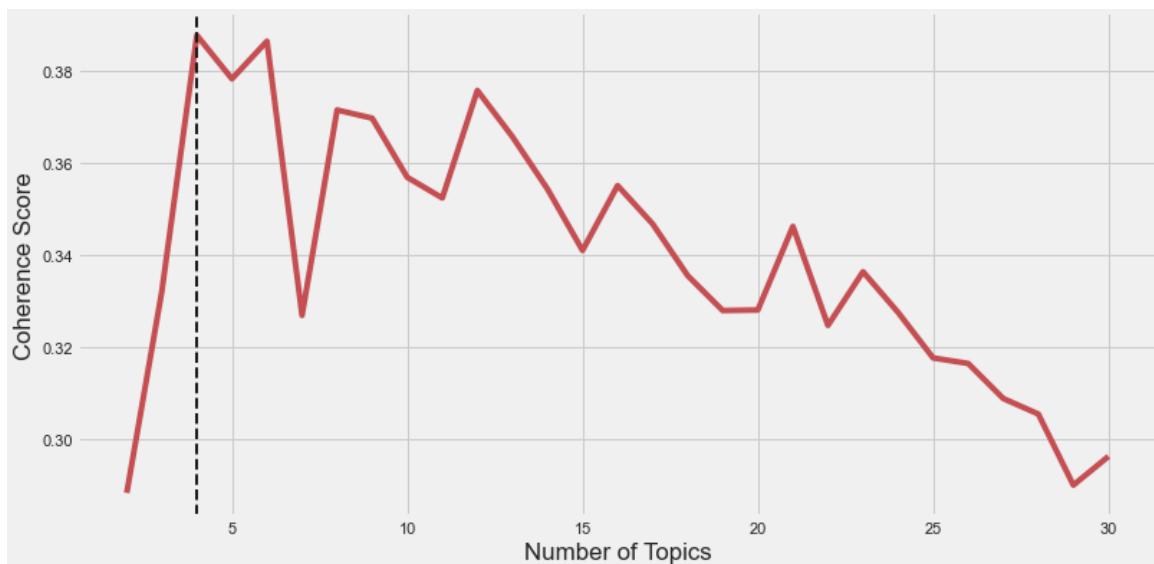
**Appendix VIII [ ~~ THIS IS RESERVED FOR THE NEWS ~~ ]**

*(back)*

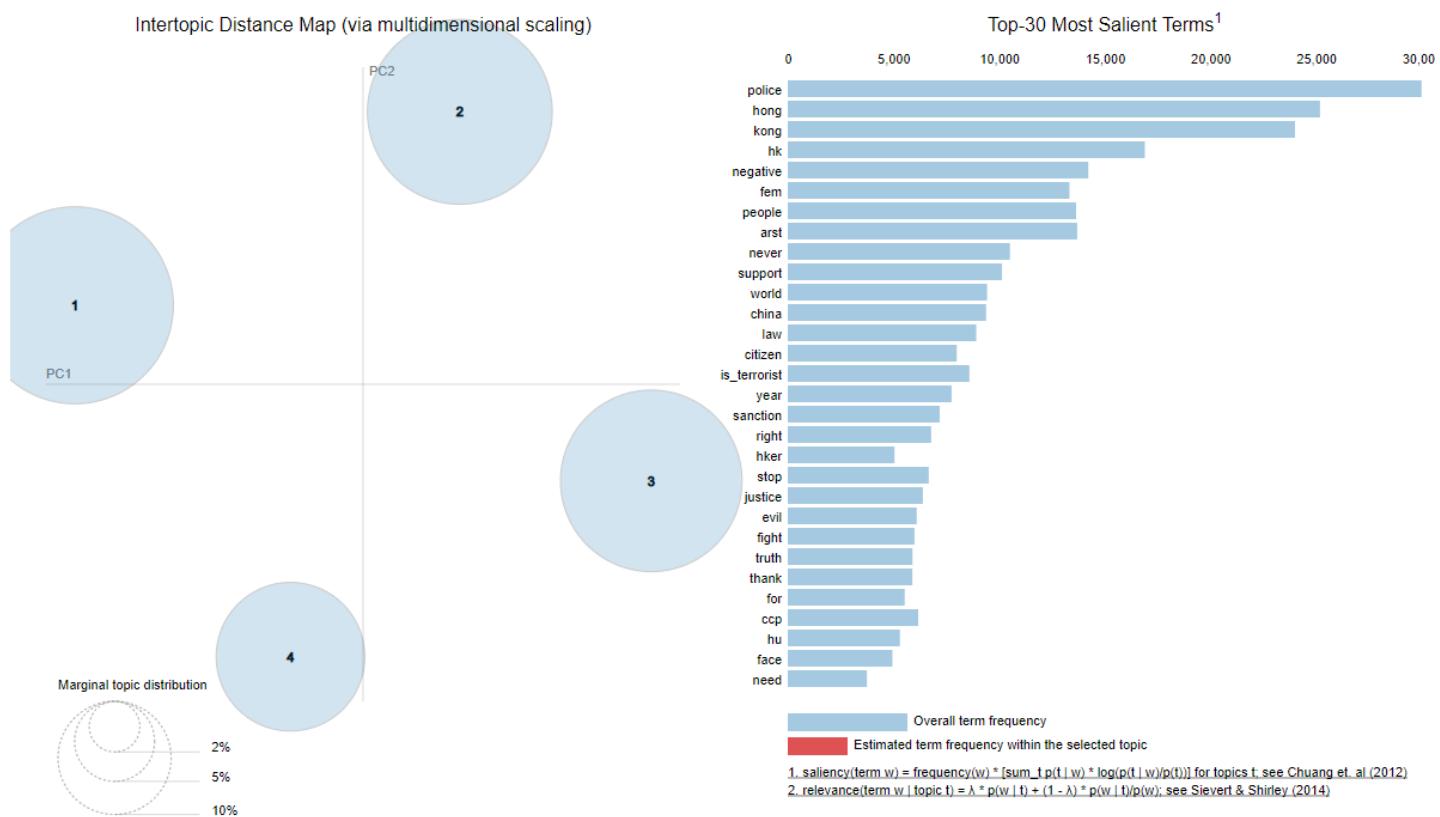
*(back)*

## Appendix IX

### [ Coherence Score, Tweets ]



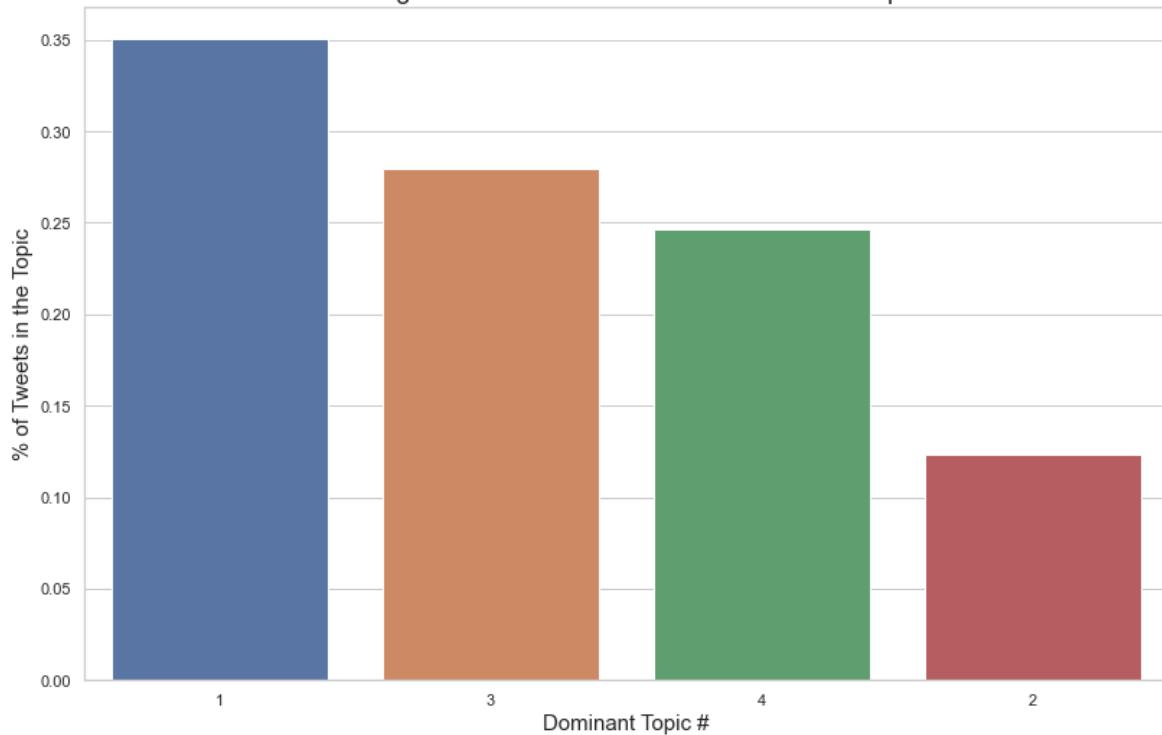
### [ pyLDAvis, Tweets ]



([back](#))

### [ Coherence Score Contribution, Tweets ]

Looking at the % Contribution of Tweets within Topics



(back)

## [ Topic 1, Tweets ]

4520 thank sir supporting hongkongers defending freedom humanrights hong kong defending core value free world please  
5078 thank continuous great effort thank ben joining speak hong kong we stand together freedom  
5116 thank ipac support support democracy freedom we really appreciate continued strong support continue fight now we see strength unity enormous  
5176 thank standing hong kong because continuous support antichina alliance stronger ever lets keep working freedom democracy hong kong together  
8871 thanks persistent effort fight freedom brave young man thank sir standing hong kong too grazie signore  
12654 thank your support means lot us hongkongers suffering limited freedom we sincerely hope free world work together stop chinas oppression many thanks  
13981 \_aderholt \_ thank sir supporting cause defending freedom defending core value free world please continue hope would  
15149 thanks liberatehongkong follow you telling world injustice anywhere threat justice everywhere free hong kong together support global vs amp  
15836 \_with\_hk \_hk \_wonghki thank folded\_hands everyone realize freedom free we stand together fight autonomy freedom  
18382 \_ballester thank support i hope move let people barcelona know situation hq lets fight freedom human rights together flexed\_biceps light\_skin\_tone please stay  
safe  
20829 \_news please keep supporting freedom hong kong thank you freedom free we need fight safeguard  
24284 \_aderholt thank sir supporting hongkongers defending freedom humanrights hong kong defending core value free world please continue hope would  
25179 thank supporting hk fire hongkonger give up we keep fighting totalitarianism please keep spreading hk situation japanese we need support  
25279 thank president tsai speaking hope stand united fight freedom democracy folded\_hands thumbs\_up red\_heart  
30752 thanks ttdog follow you telling world hong kongerinjustice anywhere threat justice everywherehuman rights advocatean eggdog lover together support global vs  
amp  
30934 thanks sec pompeo the solidarity among freedomloving nations support hong kong heartening lets stay strong keep fighting triumph democracy comes  
31176 thanks ka hing follow you telling world fight freedom stand hong kong together support global vs amp  
32829 thank american we touched under launched little now cautious now sincerely appreciate great brave move done us support human rights issues  
40814 \_with\_hk \_wonghki never give fight freedom flexed\_biceps thanks support flexed\_biceps in hk like human voice support hope world  
42882 thanks dobbi elf face\_with\_medical\_mask follow you telling world dobbi free elf thought\_balloon one dreamers we fight democracy amp human rights speaking\_he  
d yellow\_heart stand hong kong oncoming\_fist light\_skin\_tone fight freedom together support global vs  
52645 ive urgent appeal we need support continue standing hong kong campaign donate support great cause \_watch  
60080 we stand fight freedom democracy we never surrender we never defeated we we front free world we hongkongers  
60973 oh its touching heartfelt heartfelt thanks red\_heart yellow\_heart green\_heart blue\_heart purple\_heart ribbon we never give fight freedom flexed\_biceps fle  
xed\_biceps flexed\_biceps  
60974 oh its touching heartfelt heartfelt thanks red\_heart yellow\_heart green\_heart blue\_heart purple\_heart ribbon we never give fight freedom flexed\_biceps fle  
xed\_biceps flexed\_biceps  
60975 oh its touching heartfelt thanks french fellows red\_heart yellow\_heart green\_heart blue\_heart purple\_heart ribbon we never give fight freedom flexed\_bicep  
s flexed\_biceps flexed\_biceps  
61463 \_with\_hk \_angioi \_hk as hongkongers touching see ppl around world voicing support hong kong fight against tyranny thanks  
63823 thank speaking hong kong senator kitching injustice anywhere threat justice everywhere free world lock arm speak human rights hong kong  
63860 thank australia defending freedom human rights hong kong defending core value free world hongkongers urge free world  
70293 thanks standing hong kongit would tough battle hongkongers never surrender yet international support crucial us too  
71777 thank eu telling hongkongers alone fight defending freedom humanrights hong kong defending core value free world please  
72119 thank all free nations need stand together fight democray democray human right freedom suppressed ccp everywhere  
75485 thank senator housakos speaking hk we never surrender fight please continue the world needs united fight  
76840 thanks continuous support sir but hong kong loses freedom democracy us please take action support freedom fighters stop chinas aggression bastion free world  
81431 thank sen haley hong kong aspiring predatory fascistic s statement intent get away atrocities one stop anymore the socialimperialisticis waging war free world  
amp theus  
81436 thank rep riggleman hong kong aspiring predatory fascistic s statement intent get away atrocities one stop anymore the socialimperialisticis waging war free w  
ord amp theus  
82390 \_news \_watch world lets stand behind brave fighting tyranny freedom justice democracy stop fight freedom  
82673 thank rep gooden hong kong aspiring predatory fascistic s statement intent get away atrocities one stop anymore the socialimperialisticis waging war free wor  
d amp theus  
87768 \_watch thank senator we come back keep fight freedom plz let world connect together protect world future  
88165 \_watch \_with\_hk \_us from heart thanks support help we want live fear thanks speak us i love red\_heart red\_heart red\_heart red\_heart red\_heart red\_heart sunflow  
er sunflower sunflower rainbow rainbow rainbow folded\_hands folded\_hands folded\_hands folded\_hands folded\_hands \_is\_terrorist  
92488 red\_heart red\_heart red\_heart taiwan freedom human rights democracy the free world proud taiwan  
99093 \_with\_hk \_liesin\_ppl thank brother fight lot need support world we need fight believe values  
102183 its time tell free world embrace liberty amp democracy universal values injustice anywhere oppression freedom amp autonomy hong kong way  
tolerable  
104258 britian showing world fulfilling historical responsibility hong kong showing world fighting must safeguard democracy human rights  
104768 thank senators while hong kong people greatly appreciate us providing refuge place earth truly free long exists aspiring predatory nationalistic tyrannical cc  
p hk frontlinies to

(back)

## [ Topic 2, Tweets ]

3766 i pray everyday safe return i wish someone could help them folded\_hands folded\_hands crying\_face crying\_face crying\_face loudly\_crying\_face loudly\_crying\_face broken\_heart broken\_heart broken\_heart \_is\_terrorist  
6269 \_hker help kids pls no matter evil no matter brutal still fighting justice liberty amp democracy pouting\_face giving up  
7469 who serve evil who betray people hsbc bear results \_is\_terrorist  
8337 \_kwai i really love song loudly\_crying\_face folded\_hands hkers inexhaustible patriotism energy fire unite countries world globe\_showing\_americas one double\_exc  
lamation\_mark may negative bless hk hkers folded\_hands two\_men\_holding\_hands man\_and\_woman\_holding\_hands two\_men\_holding\_hands man\_and\_woman\_holding\_hands two\_men\_holding\_hands  
man\_and\_woman\_holding\_hands two\_women\_holding\_hands two\_men\_holding\_hands man\_and\_woman\_holding\_hands two\_men\_holding\_hands man\_and\_woman\_holding\_hands  
8840 \_hker\_alfieri no matter evil no matter brutal still fighting justice liberty amp democracy flexed\_biceps dark\_skin\_tone angry\_face giving up  
10685 \_hker\_hk no matter evil no matter brutal still fighting justice liberty amp democracy flexed\_biceps dark\_skin\_tone angry\_face giving up  
11706 \_craymer\_hker no matter evil no matter brutal still fighting justice liberty amp democracy flexed\_biceps dark\_skin\_tone angry\_face giving up  
12561 \_hk\_craymer no matter evil no matter brutal still fighting justice liberty amp democracy flexed\_biceps dark\_skin\_tone angry\_face giving up  
13242 \_sit\_hker no matter evil no matter brutal still fighting justice liberty amp democracy flexed\_biceps dark\_skin\_tone angry\_face giving up  
33272 \_news\_good news need more usa british need sanction people evil  
35663 \_news\_hker no matter evil no matter brutal still fighting justice liberty amp democracy flexed\_biceps dark\_skin\_tone angry\_face giving up  
37261 pope blind chi na hk negative show justice evil acts  
37626 \_hk\_hker no matter evil are brutal is wont give up still fighting justice liberty amp democracy help pls  
52494 what happened night never forgotten people behind never forgiven  
55540 \_hker\_china\_talk no matter evil are no matter brutal is giving up still fighting justice liberty amp democracy join us  
58624 \_adonis\_hker no matter evil no matter brutal still fighting justice liberty amp democracy flexed\_biceps dark\_skin\_tone angry\_face giving up  
61627 \_hker\_polumbo no matter evil are no matter brutal is giving up still fighting justice liberty amp democracy join us  
61697 \_hk\_hker no matter evil are no matter brutal is giving up still fighting justice liberty amp democracy join us  
64655 musical\_notes sing hallelujah lord sing hallelujah lord musical\_notes  
65804 \_spalding\_hker no matter evil no matter brutal is giving up still fighting justice liberty amp democracy pls help  
79586 \_hk\_hk no matter evil no matter brutal is giving up still fighting justice liberty amp democracy pls help  
80839 help help help help help \_is\_terrorist  
93297 never forget never forgetnever forget never forget never forget never forget never forget never forget never forget  
93592 govh also trustworthy opportunity na sana like could benefit potential exodus hkg  
95378 good song sharing thumbs\_up light\_skin\_tone facing enemy cooperate other  
113932 candle  
candle  
123431 all freedomlovers around world would abominate cruelty barbaric acts ccp  
131400 folded\_hands light\_skin\_tone please true folded\_hands light\_skin\_tone please true folded\_hands light\_skin\_tone please true folded\_hands light\_skin\_tone please  
true  
132369 lmaoo joke i remember happy nba back nba association clause contracts forcing everyone stand anthem not a single nba player took a knee during the anthem  
132426 what nobodys free everybody free  
132427 sad watermelon noises and sad hkers noises too \_\_\_\_\_  
132433 threat democracy world globe\_showing\_europeafrica \_is\_terrorist  
142748 hong kong people street national security law clearly ccp broken joint declaration  
145736 press freedom but condition police good friends thugs allow  
146455 hong kong great threat now please stand us strike freedom  
146811 this brave hearts taught pla lesson never forget  
149150 leung ling kit the first man suicided hong kong antiextradition movement conveying strong will lives light eyes tyrants heavy nation lets death pain remain hea  
rt carry on  
157937 already enough rapists murderers  
Name: Tweets, dtype: object

(back)

### [ Topic 3, Tweets ]

328 us condemns prodemocratic media tycoon jimmy lai asks china restore hong kongs rule law indiablooms first portal digital news management jimmy lai  
2137 hong kong police definitely rubbish totally control hold weapon hong kong longer safe city hong kong country seeking freedom  
12062 hk gov shows china style rule law is law weapon suppress prodemocracy camp law protect citizens equality law joke no justice hong kong anymore  
21203 it end it obviously political prosecution ccp adopting strong strategies silence opposition voices undermine hong kongs democracy amp rule law tool ccp arrest  
prodemocracy people please  
25833 \_hk ccp uses suppress press freedom hong kong hk police puppet political tool ccp only dictatorial regime would afraid press freedom it shows clearly ccp regime\_is\_terrorist  
25910 it ccps show order put white terror hongkongers hk police puppet political tool ccp only dictatorial regime would afraid press freedom it shows clearly ccp regime\_is\_terrorist  
26419 prodemocrats hong kong deeply concerned press freedom afte jimmy lai arrested monday colluding foreign powers offense new controversial chinese security law via  
26508 the enactment tool ccp supress freedom press hong kong it shows ccp afraid power press put white terror hk hk police puppet ccp \_is\_terrorist  
27746 \_hk after effective unlawful evil national security law hong kong freedom press freedom speech freedom election terminated hong kong longer suitable place live  
please  
29611 how police broke media office morning they destroying reputation freedom speech makes public lives white terror more people deserve sanctioned  
35711 hong kongs rule law nonexistent the city totally seized authoritarian rule now govt everything silence democracy activists hk always asset globe us  
41785 from arrest student activists yesterday dq prodemocracy candidates today hong kong suffering fullon political suppression prosecution china hong kong govt the  
world must implement sanctions authoritarians now  
42632 from arrest student activists yesterday dq prodemocracy candidates today hong kong suffering fullon political suppression prosecution china hong kong govt the  
world must implement sanctions authoritarians now  
42638 from arrest student activists yesterday dq prodemocracy candidates today hong kong suffering fullon political suppression prosecution china hong kong govt the  
world must implement sanctions authoritarians now  
42819 china care rule law autonomy hk they shows us clearly accept dissent so china turned hong kong dictatorship pls continue we urge world hold ccp accountable  
44095 from arrest student activists yesterday dq prodemocracy candidates today hong kong suffering fullon political suppression prosecution china hong kong govt the  
world must implement sanctions authoritarians now  
44099 from arrest student activists yesterday dq prodemocracy candidates today hong kong suffering fullon political suppression prosecution china hong kong govt the  
world must implement sanctions authoritarians now  
44221 from arrest student activists yesterday dq prodemocracy candidates today hong kong suffering fullon political suppression prosecution china hong kong govt the  
world must implement sanctions authoritarians now  
44637 the national security law totally brings threats freedom speech hong kong hong kong longer hong kong anymore she like province china  
45288 i believe law regulation revised clearly like people violated chinas national security law extradited as long everyone understand ns law means china deprive others freedom  
63186 the free world recognize ambition this authoritarian regime respect international order amp honor commitment ccp incapable become democratic regime what happens  
5 hong kong waken world  
63188 the free world recognize ambition this authoritarian regime respect international order amp honor commitment ccp incapable become democratic regime what happens  
5 hong kong waken world  
64677 the situation hk worrying sweeping powers granted citys police threatening personal safety local residents there hope us would continue enforce sanctioning regime  
ime china  
78669 china accept hong kong hong kong want us china city lack open mind narrow ego this act covidvirus make everyone world hate china  
81269 china passes controversial hong kong national security law a national security law imposed hong kong process nobody hong kong control over  
81961 today pass law without affected knowing even contents inside article next day pass decision void international agreements treaties they simply treat lives matte  
r but  
83887 ccp compelling national security law forcefully turns international business centre ant random china city unwise uniqueness hk gone democracy  
84403 the new laws take direct aim political dissident rather national security threats restricting freedom speech political expression hong kong

Name: Tweets, dtype: object

([back](#))

## [ Topic 4, Tweets ]

180 victim criminal on august arrested related yuen long attack protest tuen mun july including prodemocracy legislators journalist  
566 roger fisher wrote getting yes amp said separate people problem focus interests positions cant build online forums that we start brainstorming likely shared op  
posing interest conflict  
567 \_thaler roger fisher wrote getting yes amp said separate people problem focus interests positions cant build online forums that we start brainstorming likely s  
hared opposing interest conflict  
591 trying hard avoid officially locking city the idea carrie law refuses give style twisted ideas words just like responded antielab  
822 protester burned flag regime prodemocracy protests becomes second person see sentence upped jail time  
862 a lady sentenced jail months assault police laser pen no laser pen found arrested popo submit another laser pen exhibit pro ccp negative judge still accept ev  
idence \_is\_terrorist  
1870 yrsold pakistani plead guilty assaulting occasioning actual bodily harm assaulting probeijing man dec mong kok he detained days sentenced months probation or  
der conditions  
2444 great news stop chinese propaganda and im coming see top fans emotionally vulnerable spamming nmsl comment below lol  
4296 i sure probeijing sent create facts calvin in worst case may create evidence riot him this familiar with  
6945 similar ccp lies tiananmen massacre hongkong government try fraudulent history incident police disappeared hours let gangsters attack protesters  
7539 they lie us know theyre lying know theyre lying keep lying us keep pretending believe them elena gorokhova a mountain crumps  
9612 when think worst thing ever happen beginning as matter fact plan hk since \_is\_terrorist sources  
17903 when cant get real record make up like case woman convicted hurting police w laser pointer laser pointer ever found police showed court similar laser pointer h  
e brought  
25232 \_hk media first target activists second one ongoing now used working also arrested w whatever made reason  
31979 aged passing crossroad charged excise wearing mask wht obviously aged mask why charged even insulted street sanction hkpolic  
39167 index\_pointing\_up index\_pointing\_up index\_pointing\_up she children too when asked family r uk passport holders lan refused answer to protect family will uk p  
rotect dictators family will uk allow mugabes family hold uk passports  
42111 months ago riot police beaten passengers prince edward station arbitrarily it suspected passengers beaten death related cctv disclosed till now  
44860 the chinese secret police arrest everyone want our safety threatened we need care worse situation \_is\_terrorist  
45534 illegal thing say also group students arrested charge attempting split yhe nation posting facebook post  
53562 time refresh memory we still remember collude triad clearlythey totally sided tyranny suppress thusthey innocent thoughts  
58536 many government secretaries families living uk aus magnisky act must need put sanction let families enjoy live ccp governance  
61089 thats keep recording important proccp thugs pretend anything blame they familiarity abuse trust kind people  
63390 a yrsold clerk plead guilty assaulting police officer for kicking officer back nov causeway bay sentenced hrs community service source media hk  
64009 the first time ever primary election held weekend amp have banned kinds protest amp voting protest were left now please view hk  
66150 dead no evidence witness judge said community service orders may applicable here still judge chose imprisonment no guilty proven  
67288 the difference media hk west hk call peaceful protestors rioters west call rioters peaceful protestors  
67433 this lady carried us flag shopping mall n dragged n arrested by many riot po\_ice here obviously nation flags r  
69397 the party people one sounds like religious testimony ccp tries replace religion itself a concept always failed babylon moscow  
69576 on contrary hkgov made books famous now more people going find ways read them like mainlders used come hk read banned books  
69829 vetting necessary kick ccp supporters hk gov officials hk police away british they deserve live british  
71786 an yrsold plead guilty possession offensive weapon a flick knife nov chai wan sentenced months probation order source youarenotalonehk\_live  
72368 oh liberate hk revolution time liberate hk revolution time liberate hk revolution time liberate hk revolution time liberate hk revolu  
tion time  
77551 a man arrested becoz hv flags bag insane next time foreigner walk street police may arrest himher foreigner hv different nationality hidden head  
85428 a uni student possession pipe month sentence a girl related possession kg ketamine walk away with absolutely nothing  
86700 a yrsold gaffer plead guilty possession offensive weapon public places intent commit arson oct tseung kwan o sentenced yrs imprisonment  
86767 an yrsold plead guilty assaulting police officer nov tseung kwan o sentenced months probation order probation officers guidance source youarenotalonehk\_liv  
e  
90270 arrested youth lynched police can charged assaulting police not guilty magistrate said polices unreliable witness  
92830 we also claim say along ladakh we needs understand publishing providing fake news info media create confusion succeed  
99033 have seen police could rush station indiscriminately attack people shown difference terrorist attack what ridiculous faced consequences  
99376 the people not asking violence asking justice these protests one ways use voices make change made difference before everything power make sure happens again  
99710 when dare mention they know could likely cause death have stopped yet no have faced consequences that no thats todays  
100588 judge i would like ask u seriously injured n medical report prove popo get hurt u treat turth \_is\_terrorists  
102137 judge said cause testimonies polices found reliable inconsistent the judge sentenced yearold teacher assaulting police

(back)

*(back)*

