

Observing the Recent Frictions Between China and United States of America

A Transnational Perspective Based on Google Trend Data

Junhan Tang & Ziheng Zhang

2019-12-13

Contents

Introduction	1
Data	1
Results	4
Discussion	12
References	13

Introduction

Since March 2018, the escalation of tensions between China and the United States has obtained ubiquitous attention all over the world. This massive attention is probably caused by the tangible impact of the dispute on the global economy. According to a survey conducted by Fox Business, more than two-thirds respondents reported that their business have been influenced by the recent frictions (Fordham,2019). However, the dispute not only influences economy, but it also extends to other aspects. For example, three bills has passed unanimously in the House of Representative on Oct 16 this year (2,3,4), all aiming at Hong Kong. Moreover, there was also another bill passed on the same day about Meng Wanzhou, the CFO of Huawei (5,7). The four acts were passed while there was an ongoing negotiation about trade war between China and America with low chance of agreement (Zengerle,2019). Under this circumstance, the unanimously passing of the four acts does not seem like a separate event, rather we can consider it as related to the trade war. We are interested in the global reaction on these three topics. More specifically, in this project, we want to explore the amount of attention people have paid to the trade war, Huawei and Hong Kong protest based on the assumption that when people intend to know something, they will google it.

Data

It is viable to utilize Google trend data to construct a database for analysis. Google trend data is not being selected heuristically, rather according to the data from statcounter(8), it has 92.95 % market share across all platforms, which enables us to get the data from all over the world. Firstly, by using web scraping, we obtain a list of countries and areas from the International Organization for Standardization (ISO), and the list includes geocode for each of the country and area. Next, we check whether there are searching records of the interested keywords for each of the country or area according to geocodes (e.g. US, CN), by using gtrends function.

Then, for the first topic (trade war between China and the United States), we start our data gathering by using a more general keyword—‘trade war’. By writing a new function (geocode_1_exist) and inserting it into apply function, we can filter out the countries and areas where there are searching records of “trade war” based on geocodes from the ISO list and collect data of transnational searching frequencies. Combined with the time period we are interested in, we can find the searching frequencies of ‘trade war’ of the filtered countries and areas from March 1st, 2018 to November 1st, 2019. Meanwhile, since Google trend shows the

related queries of 'trade war', we can also choose the related keywords and gather their frequencies to obtain a more comprehensive database for analysis. The keywords we choose are 'us china trade war' and 'china trade war', since the searching frequencies of the two keywords of the filtered countries and areas are highest and the missing values about the searching frequencies of 'us china trade war' and 'china trade war' are fewer than the missing values of other related keywords about 'trade war'. In addition, we also collect searching frequencies of another keyword 'US trade war' to amplify the database. Finally, we combine the searching frequencies of the four keywords about the trade war between China and the United States and create a new database by data cleaning.

```
#asp_1 selecting countries#
geocode_1<-as.matrix(geocode_country[,1])
geocode_1_exist<-function(x){
  Country<-gtrends("trade war",time="2018-03-01 2019-11-01",geo=x)
  overtime<-Country$interest_over_time
  if (is.null(overtime)==FALSE) {return (x)}
  if (is.null(overtime)==TRUE) {return (NA)}
}
geocode_1_list<-apply(geocode_1,1,geocode_1_exist)
geocode_1<-as.data.frame(geocode_1_list)
geocode_1<-geocode_1 %>% drop_na(.)
colnames(geocode_1)<-"geo_code"
#finding hits of 'trade war' #
over_time_1<-function(x){
  Country<-gtrends("trade war",time="2018-03-01 2019-11-01",geo=x)
  return (Country)
}
data_lists_1<-apply(geocode_1,1,over_time_1)
database_asp_1<-data.frame()
for (i in 1:length(data_lists_1)) {
  overtime<-data_lists_1[[i]][[1]]
  database_asp_1<-rbind(database_asp_1,overtime)
}

database_asp_1 %<>%
  mutate(date_geo=paste0(date,"",geo)) %>%
  select(date_geo,geo,keyword,hits)

#finding related topics#
related_query_database_1<-data.frame()
for (i in 1:length(data_lists_1)) {
  country<-data_lists_1[[i]]
  related_query<-country[[7]]
  related_query_database_1<-rbind(related_query_database_1,related_query)
}
query_freq_1<-related_query_database_1 %>%
  drop_na(.) %>%
  select(value,geo) %>%
  group_by(value) %>%
  summarize(freq=n()) %>%
  arrange(desc(freq)) %>%
  filter(freq>=15) %>%
  filter(value %in% c("us china trade war","china trade war"))

#finding hits for related queries#
```

```

finding_hits_1<-list()
for (i in 1:length(query_freq_1$value)) {
  cn<-function(x) {
    Country<-gtrends(query_freq_1$value[i],time="2018-03-01 2019-11-01",geo=x)
    return (Country)
  }
  data_lists_1_1<-apply(geocode_1,1,cn)
  finding_hits_1<-c(finding_hits_1,data_lists_1_1)
}
#finding other keywords#
over_time_1_2<-function(x){
  Country<-gtrends("US trade war",time="2018-03-01 2019-11-01",geo=x)
  return (Country)
}
data_lists_1_2<-apply(geocode_1,1,over_time_1_2)
finding_hits_1<-c(finding_hits_1,data_lists_1_2)
data_related_1<-data.frame()
for (i in 1:length(finding_hits_1)) {
  country<-finding_hits_1[[i]]
  overtime<-country[[1]]
  if (is.null(overtime)==FALSE) {data_related_1<-rbind(data_related_1,overtime)}
}
data_related_trade_war<-data_related_1 %>%
  mutate(date_geo=paste0(date,",",geo)) %>%
  select(keyword,hits,date_geo)
data_related_kw_1<-database_asp_1
keywords<-c(query_freq_1$value,"US trade war")
for (i in 1:length(keywords)) {
  kw<-data_related_trade_war %>% filter(keyword==keywords[i])
  data_related_kw_1<-data_related_kw_1 %>% left_join(kw,by="date_geo")
  i<-i+1
}
#combining#
database_asp_1<-data_related_kw_1 %>%
  rename(trade_war=hits.x,us_china_trade_war=hits.y,china_trade_war=hits.x.x,US_trade_war=hits.y.y) %>%
  select(date_geo,geo,trade_war,us_china_trade_war,china_trade_war,US_trade_war)

```

The procedure of collecting searching frequencies of keywords about Huawei (a Chinese high-tech giant in the center of the frictions between China and the United States) is similar with the data collection process of trade war. For this topic, we choose two different clue keywords, ‘Meng Wanzhou’ and ‘huawei ban’ at the beginning. Then we further filter related keywords of the two clue keywords and obtain the searching frequencies of the related keywords. After combining the searching frequencies of both the clue keywords and the related keywords about Huawei and cleaning data, we obtain the second database.

For the third topic, Hong Kong protest, the only meaningful keyword is ‘Hong Kong protests’ and it does not possess any useful related keywords. Thus, the third database only contains the keyword ‘Hong Kong protests’.

After establishing the database for each topic, we combine the databases of all topics (the trade war between China and the United States, Huawei, Hong Kong protest) and create a comprehensive database which we also add continent names for each filtered country or area. Consequently, the combined database has 7569 observations from 87 candidate countries with 15 variables.

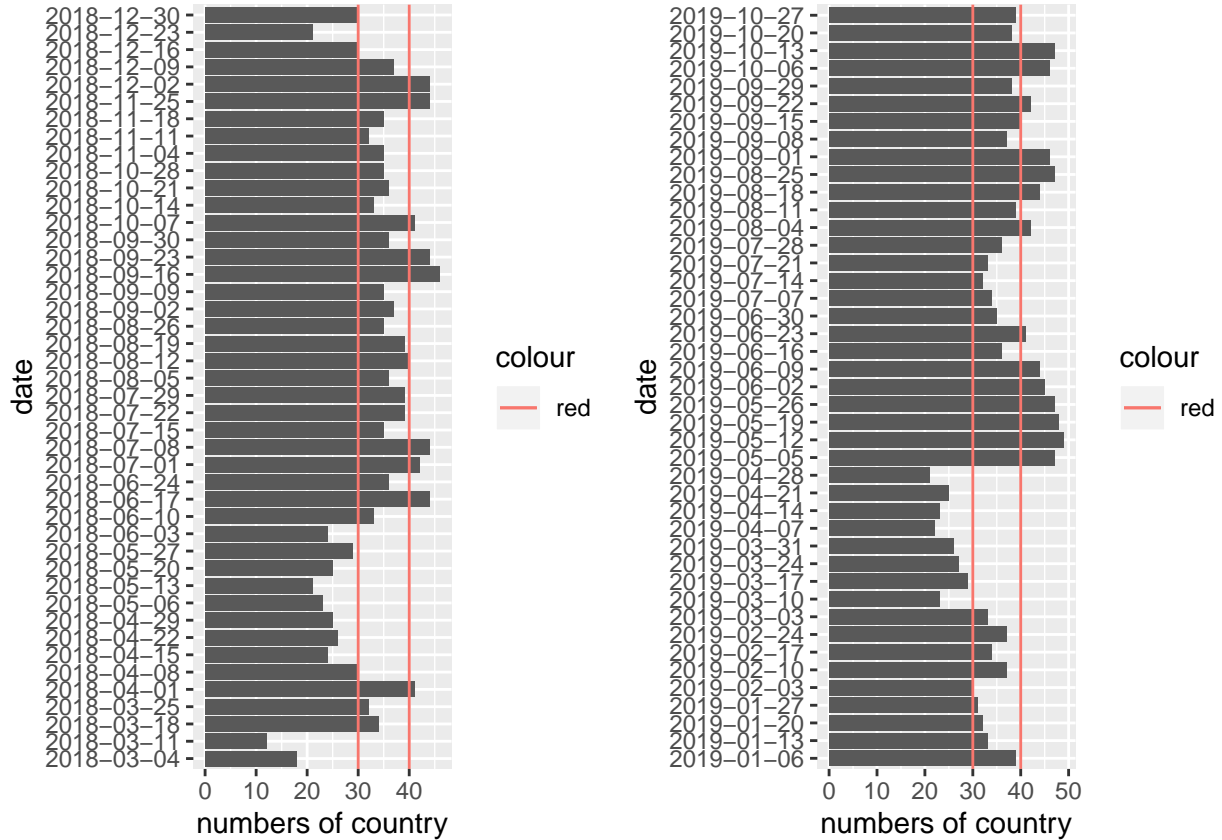
Results

Analysis and Data Exploration

US-China trade war

First, we analyze the searching frequencies of the keywords about the trade war between China and the United States. Comparing with ‘trade war’, we find that ‘us china trade war’ is a better measurement of the attention being paid by the people of the candidate countries to the trade war between China and United States. Then, we filter the searching records of ‘us china trade war’ based on the hierarchical clustering with keywords ‘US trade war’ and ‘china trade war’. As a result, we can classify searching records of ‘us china trade war’ into four clusters. For each cluster, we first test whether the searching frequency of ‘us china trade war’ is correlated with the searching frequency of ‘US trade war’, since ‘US trade war’ is weakly related to ‘us china trade war’. Next, we check if there are correlations between the searching frequencies of ‘us china trade war’ and the interaction of the searching frequencies of ‘US trade war’ and ‘china trade war’. Since people paying attention to the trade war between China and US may choose different keywords about this topic, the searching frequency of ‘us china trade war’ may reflect more universal attention of people if it is contributed by both the searching frequency of ‘china trade war’ and the searching frequency of ‘US trade war’ collectively.

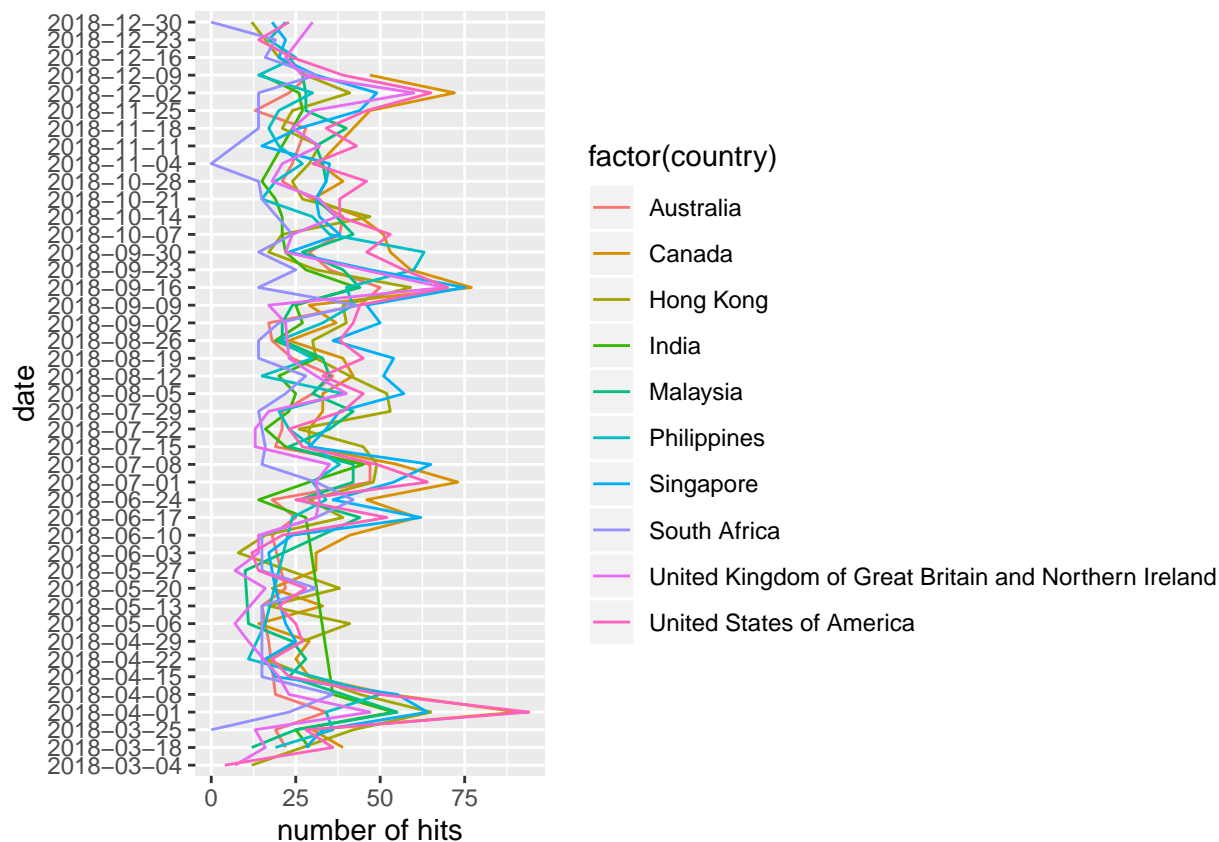
After clustering and filtering, we extract two clusters of searching records of ‘us china trade war’ for 53 countries and areas from the database. By visualization, we can find that along the timeline from March 1st, 2018 to November 1st 2019, in some weeks which is signified by one day of the weeks, the numbers of countries with searching records of ‘us china trade war’ are over 40. In contrast, the numbers of countries are between 30 and 40 or lower than 30 in other weeks.



Thus, we argue that the countries where people keep longest duration of attention to the trade war between China and US should have searching records of ‘us china trade war’ in our entire interested time period. Meanwhile, the countries where people pay shortest duration of attention should have searching records only at the weeks when the numbers of countries with searching records of ‘us china trade war’ are over 40.

The reason might be that people in these countries pay attention to US-China trade war when something noticeable happens. Thus, according to the duration of attention people paying to ‘us china trade war’ in each country, we adopt k-means cluster and classify the 53 countries and areas into 3 groups based on the result of clustering. In the first group of countries and areas, the searching records of ‘us china trade war’ exist in the entire interested period, which demonstrates that the people in the countries and areas pay the most attention to the US-China trade war. Accordingly, the length of people’s attention in the second group is shorter and the third group hold shortest attention.

For instance, visualizing the search frequencies of the first group, it can be found that searching records of ‘us china trade war’ in these countries and areas exist in almost all of weeks of interest in 2018, and the variations of searching frequencies among this group members are similar. The result of the weeks in 2019 also follows this pattern. This means that this group of countries and areas continuously pays attention to the US-China trade war, and the peaks of searching frequencies of ‘us china trade war’ show that when there are important news about US-China trade, people from all countries and areas start to pay attention.



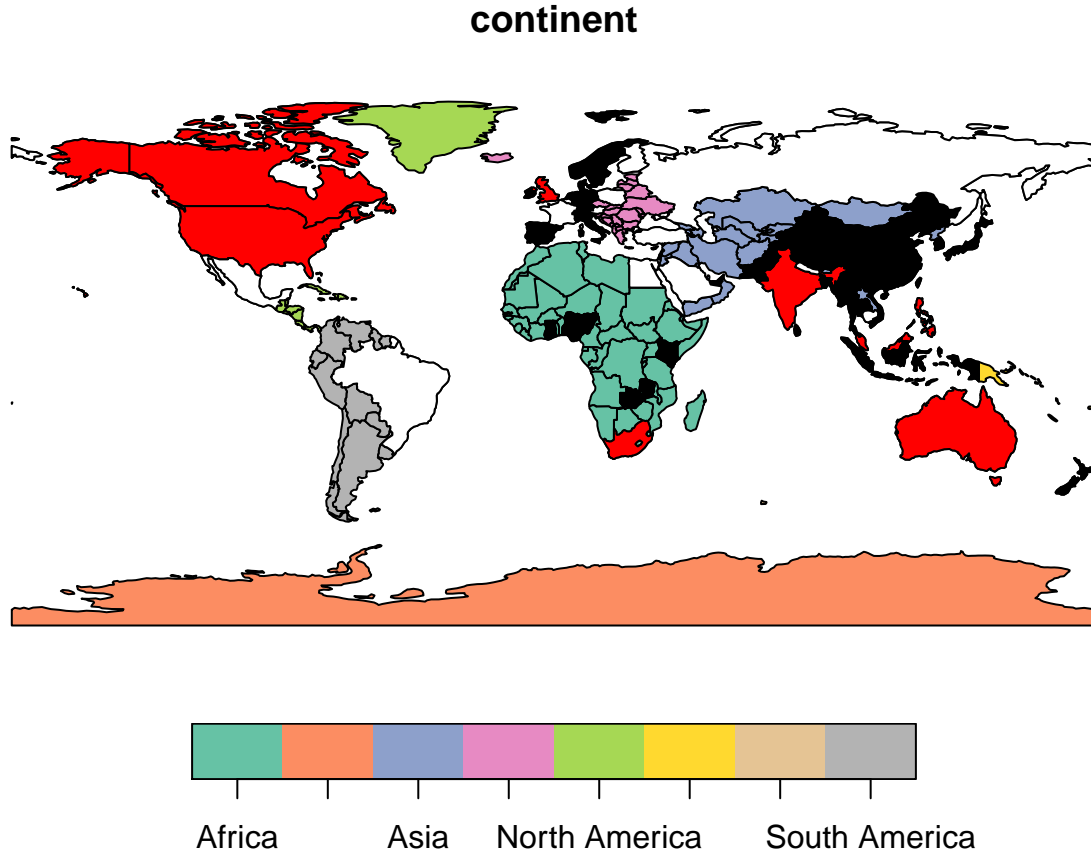
Then, we fill the three categories of countries and areas into a world map, identifying the geographical distribution of these countries. Since Singapore, Hong Kong and Saint Helena, Ascension and Tristan da Cunha do not have corresponding geocodes in the world map database, they will not emerge in the filled map. In this map, the countries marked in red represent the countries where people pay the most attention to the US-China trade war; the countries marked in black represent the countries where people pay less attention; the countries marked in white represent the countries pay the least attention.

```
#Using map to show the countries#
world_geo<-as.data.frame(world$iso_a2)
colnames(world_geo)<-"geo"
world_geo$geo<-as.character(world_geo$geo)
world_map<-rbind(most_paid,next_most_paid,least_paid)
maps_data<-world_map %>%
```

```

inner_join(geocode_country,by="country") %>%
inner_join(world_geo,by="geo")
plot(world["continent"],reset=FALSE)
for (i in 1:nrow(maps_data)) {
  cn = world[world$iso_a2 == maps_data$geo[i], ]
  cn_1=st_union(cn)
  if (maps_data$kind[i]==1) {plot(cn_1,add=TRUE,col="red")}
  if (maps_data$kind[i]==2) {plot(cn_1,add=TRUE,col="black")}
  if (maps_data$kind[i]==3) {plot(cn_1,add=TRUE,col="white")}
}

```

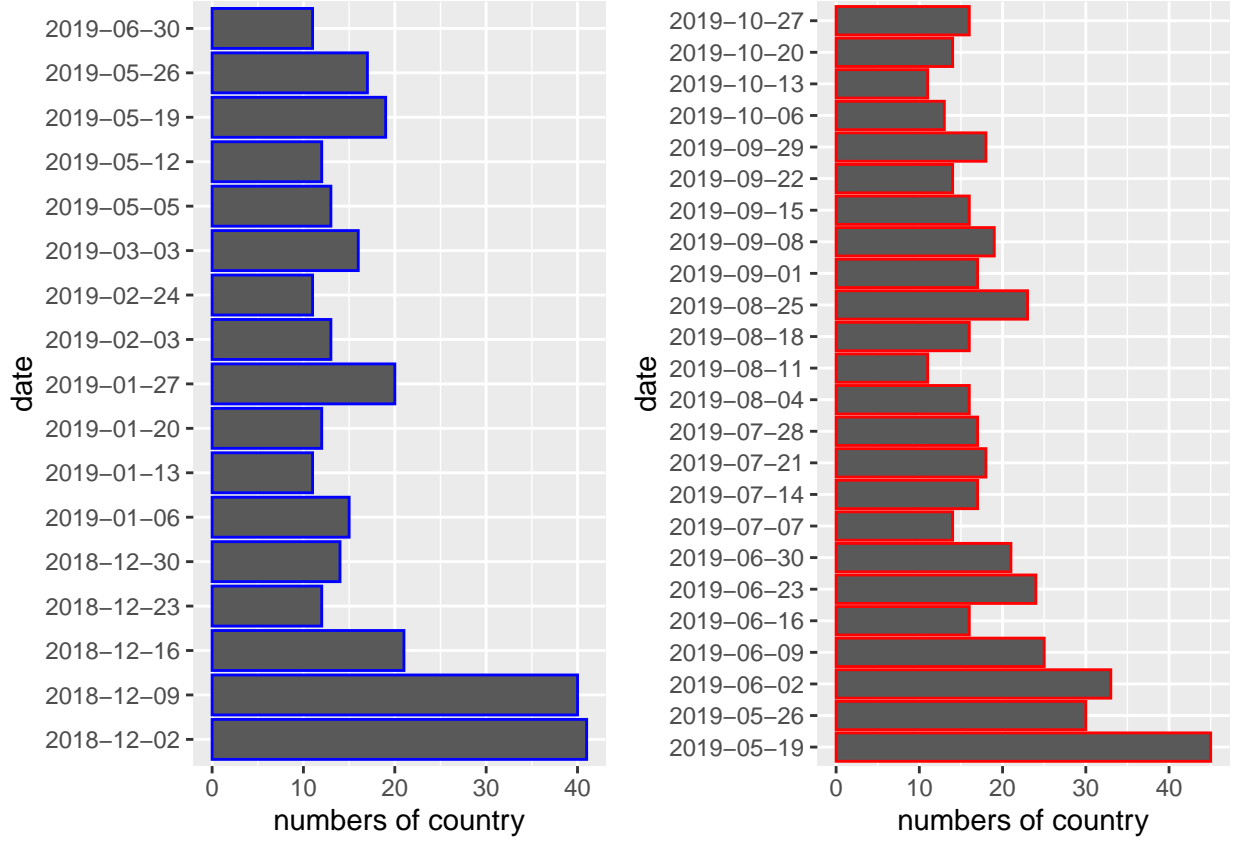


From the above map, it can be found that people in the countries with English as the native language pay the longest duration of attention to the US-China trade war by searching the keyword ‘us china trade war’. Then it can be found that countries in East Asia have similar duration of attention as countries in the West part of Europe and Africa. The countries with shortest duration of attention are mainly distributed in middle east, Eastern Europe and Latin America.

Huawei

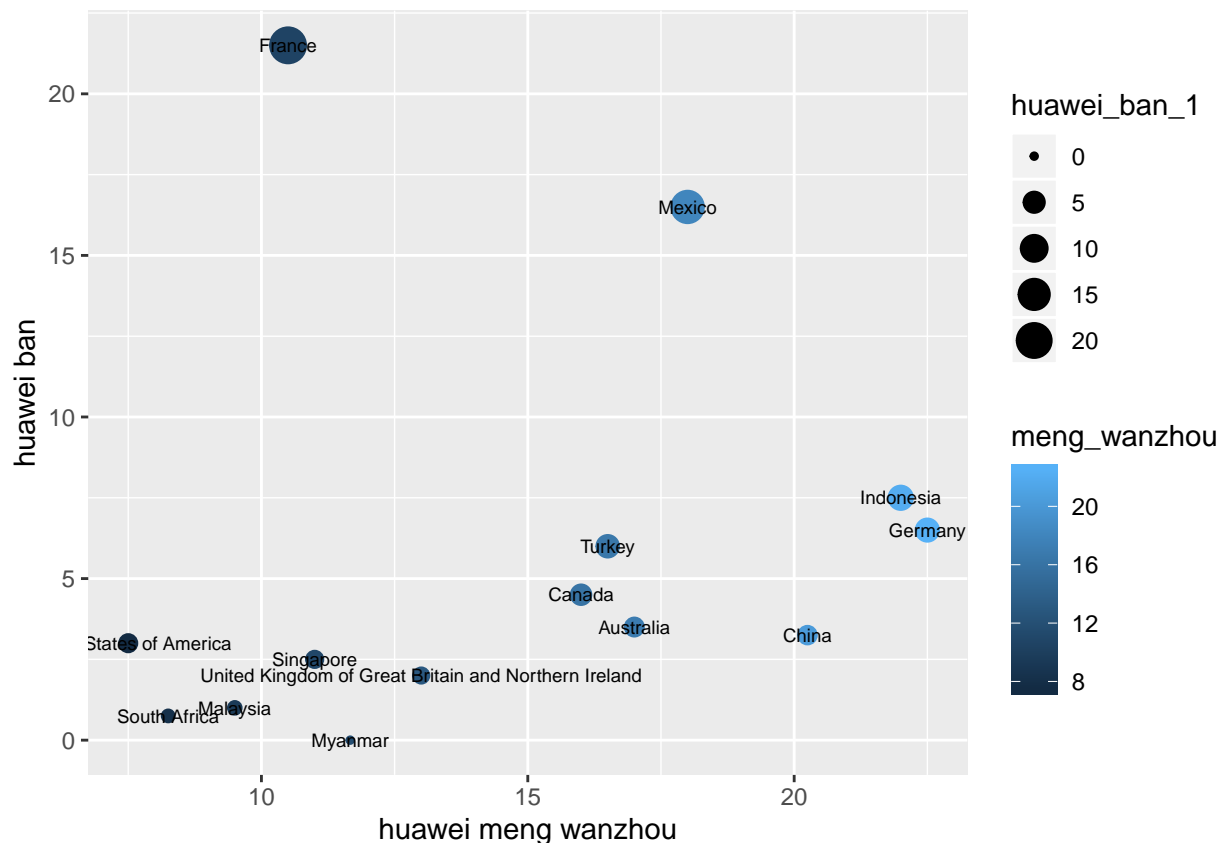
Now we start the analysis of the keywords about Huawei. During the data collection process, we find two seemingly separate key words ‘meng wanzhou huawei’ and ‘huawei ban’, so we would like to explore whether the attention to Meng Wanzhou (CFO of huawei, being confined in Canada) is correlated with the attention to bans about Huawei. Meanwhile, although another keyword ‘huawei google ban’ is a related keyword of ‘huawei ban’, we also would like to know whether there is a relationship between the attention to bans about Huawei and the attention to the ban initiated by Google to Huawei. In other words, we would like to explore whether the ban from Google contributes to searching frequencies of ‘huawei ban’.

Firstly, we respectively filtered the weeks when more than 10 countries show searching records of ‘meng wanzhou huawei’ and ‘huawei google ban’, since important news of the two keywords occur in different time periods, we can concentrate on the periods when the keywords acquire sufficiently wide attention among the countries.



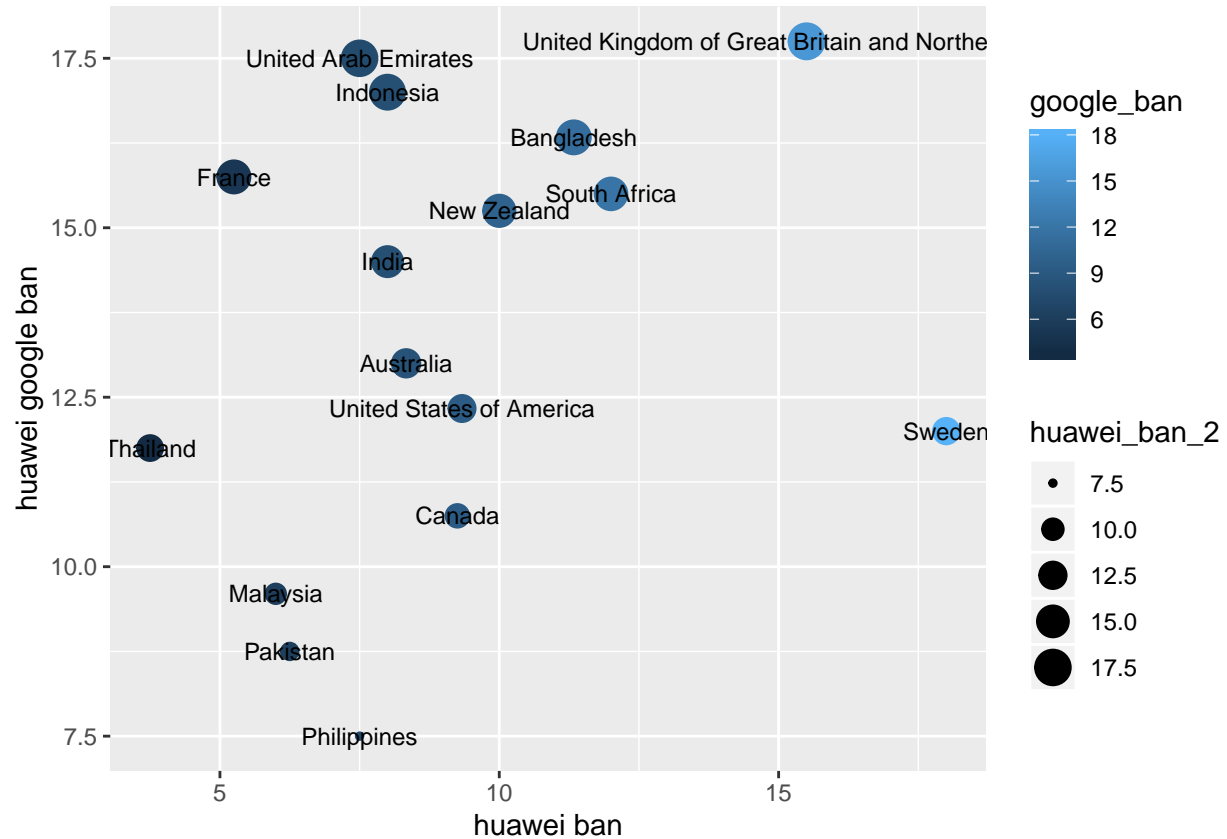
It can be found that relatively universal searching records of ‘meng wanzhou huawei’ start from the first week of December 2018 and end at June 2019 (left side bars), while relatively universal searching records of ‘huawei google ban’ are from May 2019 to October 2019 (right side bars). Then, by visualizing searching records of ‘huawei ban’ along the whole timeline, we find that the number of countries with searching records of ‘huawei ban’ is higher in the periods when ‘meng wanzhou huawei’ or ‘huawei google ban’ obtain attention from more countries. Thus, we filter the searching frequencies of the three keywords by removing the dates before December 2018 and carry out the k-means clustering of the three keywords.

Based on the k-means clustering, we choose one cluster where there is a linear relation between the searching frequency of ‘huawei ban’ and those of ‘meng wanzhou huawei’ from December 2018 to June 2019, and there is also a linear relationship between the searching frequency of ‘huawei ban’ and ‘huawei google ban’ from May 2019 to October 2019. Within this cluster, we would like further subset the period to one or two months. For ‘huawei ban’ and ‘men wanzhou huawei’, we concentrate on searching frequencies of the two keywords in December 2018, when the arrest of Meng wanzhou was reported (7). Through visualization, we demonstrate the relationship between the searching frequency of ‘huawei ban’ and searching frequency of ‘meng wanzhou huawei’ among the available countries.



From the above plot, we can find that the relationship between the frequencies of ‘huawei ban’ and ‘meng wanzhou huawei’ exist in 14 available countries. The difference of color of the points in this plot demonstrates the searching frequencies of ‘meng wanzhou huawei’, and the size for each point in the plot demonstrates the search frequencies of ‘huawei ban’. Thus, there are diverse modes of relationship of the two keywords among some available countries. For example, although France and Mexico commonly have high searching frequencies of ‘huawei ban’, the searching frequencies of ‘meng wanzhou huawei’ in France is obviously lower than the searching frequencies in Mexico. This difference may reflect that in Mexico, people’s attention to the event of Meng Wanzhou extend their interest to bans about Huawei, while in France, the event of Meng Wanzhou should not be a factor contributing to the high attention of people to bans of Huawei. On the other hand, while the 12 other available countries demonstrate diverse searching frequencies of the two keywords, there seems to be a linear relationship of searching frequencies of the two keywords. Specifically, points with lighter color tend to be larger, meaning that in these countries such as Indonesia and Germany, more attention being paid to Meng Wanzhou contribute to more attention to bans about Huawei. In contrast, points with darker color tend to be smaller, meaning that in these countries such as Singapore and South Africa, less attention to the event of Meng Wanzhou contribute less to attention of bans about Huawei.

Similarly, through visualization, we find the relationship of searching frequencies of ‘huawei ban’ and ‘huawei google ban’ among the available countries, from late May, 2019 to late June, 2019.



From the above plot, we can find that there are 15 available countries showing relationship between frequency of 'huawei ban' and 'huawei google ban'. The difference of color for each point in this plot demonstrates the different searching frequencies of 'huawei google ban', and the size of each point in the plot demonstrates the search frequencies of 'huawei ban'. In this plot, we can find that points with darker colors tend to be smaller, while points with lighter colors tend to be larger. This means that for countries such as Malaysia and Philippines, low searching frequency of 'huawei ban' are accompanied with low searching frequency of 'huawei google ban'. In contrast, for the countries such as the United Kingdom, high searching frequency of 'huawei ban' are accompanied with high searching frequency of 'huawei google ban'. In sum, in the period from late May 2019 to late June 2019, there should be a linear relationship of searching frequencies of 'huawei ban' and 'huawei google ban' which means that the attention to the bans initiated by Google to Huawei partly contributes to the attention paid to bans about Huawei.

Hong Kong protest

Now we come to the searching frequencies of the keyword Hong Kong protest.

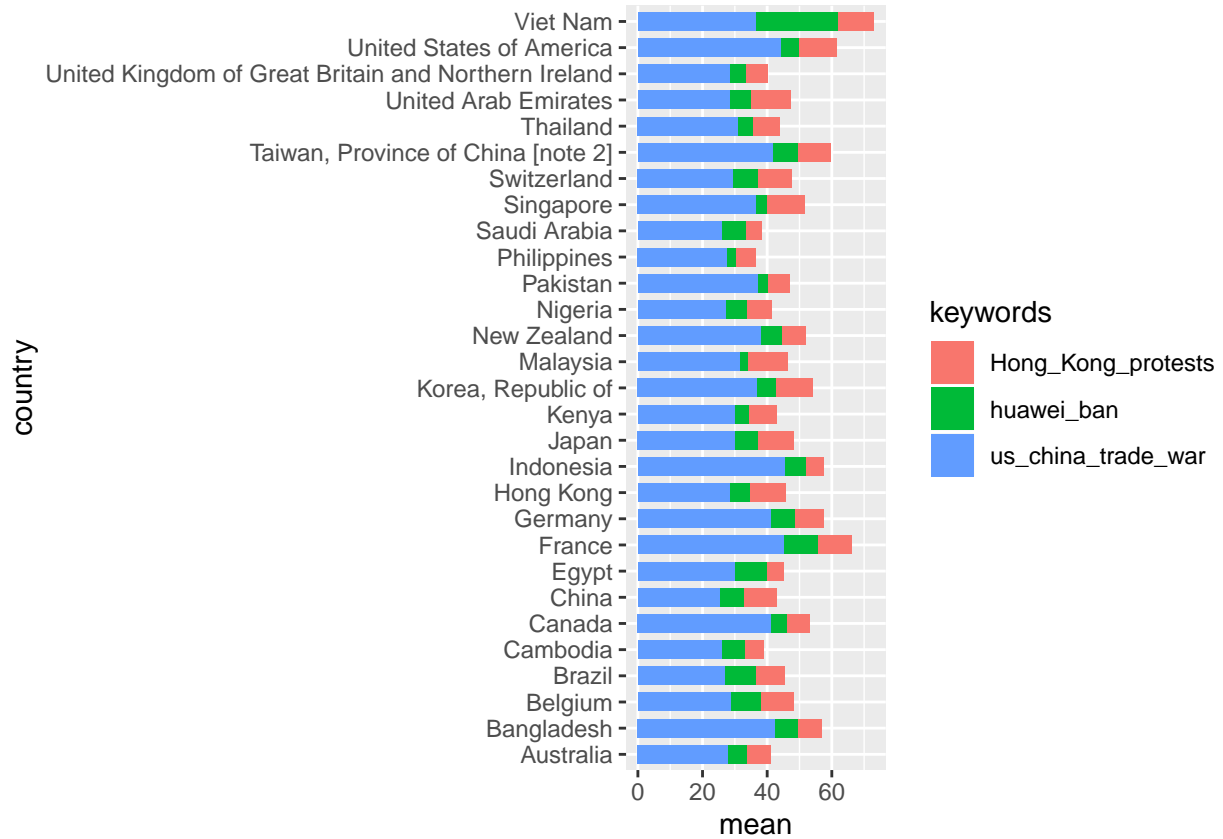
We use interactive graph to visualize the searching frequency of the keyword 'Hong Kong protests', which can demonstrate the different combinations of countries about the searching frequency of 'Hong Kong protests'.

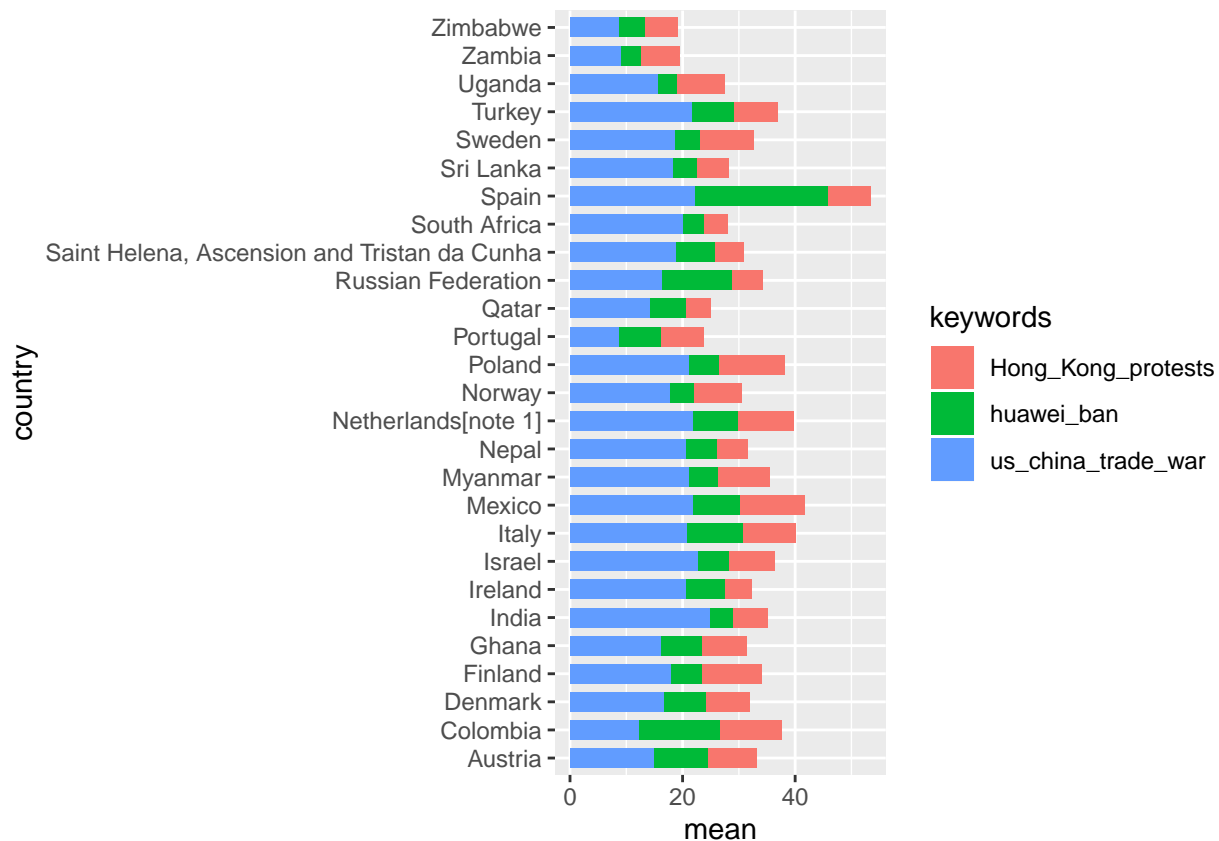
Combining topics

In the final part of the analysis, we combine the searching frequencies of the core keywords for the three topics: 'us china trade war', 'huawei ban' and 'Hong Kong protests', and we compute the mean of searching frequency for each keyword in each available country, which can be used for comparing the degree of people's attention to the three topics (US-China trade war, Huawei and Hong Kong protests).

Then, we visualize the mean searching frequencies of the three keywords 'us china trade war', 'huawei ban' and 'Hong Kong protests' for each available country, which is beneficial to further comparison among

the available countries. In the process of visualization, we find that approximately half of the available countries have the mean searching frequencies over 25 of 'us china trade war' so we extract these countries into one group and other available countries into the other group, which is beneficial to create graphs.





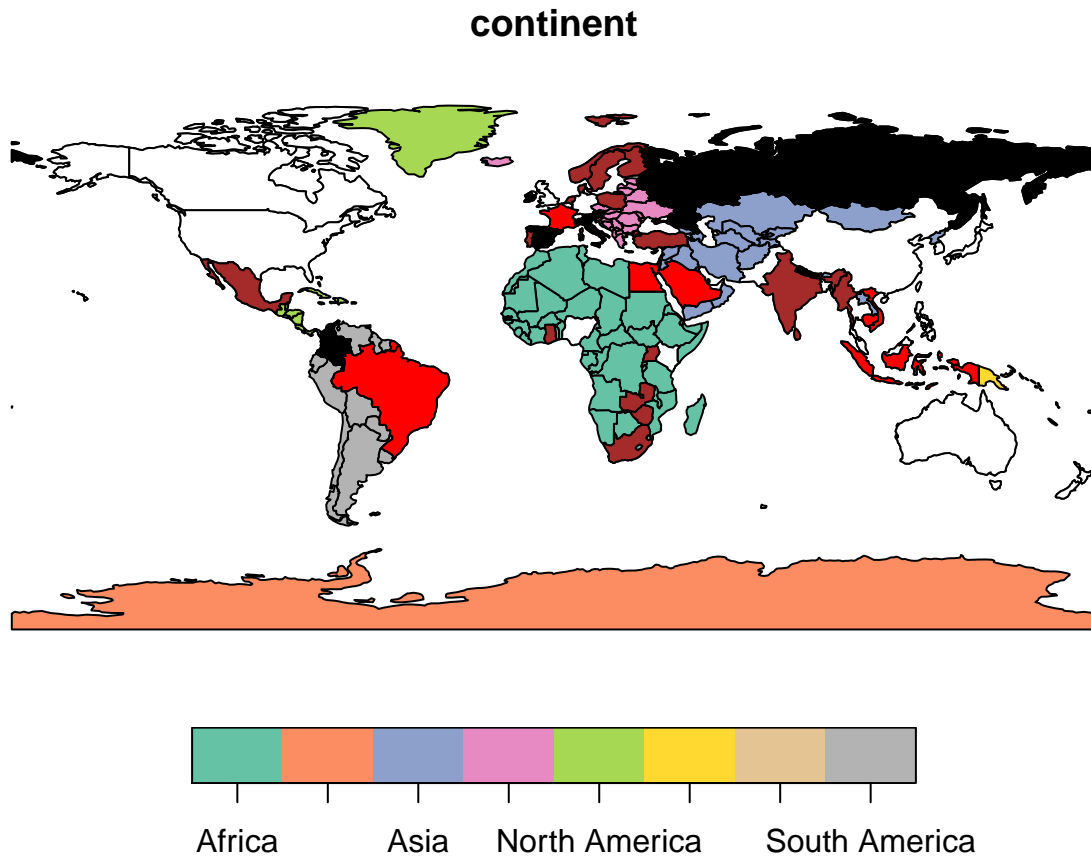
Based on the two graphs, we can find that compared with the mean searching frequencies of ‘huawei ban’ and ‘Hong Kong protests’, the mean of the searching frequency of ‘us china trade war’ is higher in all of the 56 available countries. This difference demonstrates that from March 2018 to October 2019, US-China trade war generally receives more attention than other two topics about the frictions between China and the United States. In addition, in some countries, the mean searching frequencies of ‘huawei ban’ is higher than that of cHong Kong protests⁷, while in other countries, the result is different. Thus, although the duration of attention to Hong Kong protests is shorter than that of attention to Huawei, people in the countries with mean searching frequency of ‘Hong Kong protests’ higher than mean searching frequency of ‘huawei ban’ may reflect that in the eruption of Hong Kong protest, these people pay much more attention to Hong Kong protest than Huawei.

Thus, we combine the standard of whether the mean searching frequency of ‘huawei ban’ is higher than the mean searching frequencies of ‘Hong Kong protests’ with the previous standard of whether the mean searching frequency of ‘us china trade war’ is higher than 25 and separate the 53 available countries (excluding Singapore, Hong Kong and Saint Helena, Ascension and Tristan da Cunha) into four groups.

Finally, we fill the four groups of available countries into another world map.

```
#Using map to show the countries#
asp3_1$group<-0
for (i in 1:nrow(asp3_1)) {
  if (asp3_1$us_china_trade_war[i]>25 & asp3_1$huawei_ban[i]>asp3_1$Hong_Kong_protests[i])
  {asp3_1$group[i]<-1}
  if (asp3_1$us_china_trade_war[i]<=25 & asp3_1$huawei_ban[i]>asp3_1$Hong_Kong_protests[i])
  {asp3_1$group[i]<-2}
  if (asp3_1$us_china_trade_war[i]>25 & asp3_1$huawei_ban[i]<asp3_1$Hong_Kong_protests[i])
  {asp3_1$group[i]<-3}
  if (asp3_1$us_china_trade_war[i]<=25 & asp3_1$huawei_ban[i]<asp3_1$Hong_Kong_protests[i])
  {asp3_1$group[i]<-4}
```

```
{asp3_1$group[i]<-4}
}
```



In this map, countries marked in red represent the countries where the mean searching frequency of ‘us china trade war’ is over 25, and simultaneously the mean searching frequency of ‘huawei ban’ is higher than the mean searching frequency of ‘Hong Kong protests’. The countries marked in black represent the countries where the mean searching frequency of ‘us china trade war’ is lower than or equal to 25, and simultaneously the mean searching frequency of ‘huawei ban’ is higher than the mean searching frequency of ‘Hong Kong protests’. The countries marked in white represent the countries where the mean searching frequency of ‘us china trade war’ is over 25, and simultaneously the mean searching frequency of ‘huawei ban’ is lower than the mean searching frequency of ‘Hong Kong protests’. The countries marked in brown represent the countries where the mean searching frequency of ‘us china trade war’ is lower than or equal to 25, and simultaneously the mean searching frequency of ‘huawei ban’ is lower than mean searching frequencies of ‘Hong Kong protests’.

Discussion

In this part, we would further discuss the main results from the previous analysis.

Results about trade war between China and the United States

From the first filled map, it can be found that from March 2018 to October 2019, there are diverse modes about the duration of attention to the US-China trade war among Western European countries. The reasons behind the difference may be about with the frequency of using English. The United Kingdom, where people’s main language is English, shows the longest duration of recorded searching frequencies about ‘us china trade war’, and therefore is classified as the country where people successively pay attention to the trade war between China and the United States. For Germany, Scandinavian countries and Spain, although people’s

first language is not English, they are also accustomed to adopting English search engines like Google, so they also get information about the trade war through English keywords searching. Thus, from searching records of Google trend, these countries still show relatively long period of attention to US-China trade war. For France, where people are so proud of their own language and reluctant to use English in some degree, searching records of English records are lower than that of other western countries, so France is classified as the category of countries where people pay least attention to US-China trade war. In sum, transnational comparison of duration of people's attention to the US-China trade war should sufficiently consider people's habits of using language among different countries, and English keywords from the searching engine can only partially explain the different length of attention for different countries.

Results of cross-topic comparison

In the second filled map, comparing the countries marked in red and countries marked in white, both of which belong to the group of available countries with more attention to US-China trade war than other available countries. In the countries marked in red, for instance France, Brazil and Egypt, people pay more attention to issues about Huawei than Hong Kong protest from March 2018 to October 2019, while in the countries marked in white including Germany, America, Japan and China, people pay more attention to Hong Kong protest than to news about Huawei. Therefore, it seems interesting to compare economically weaker countries (countries marked in red), where people pay more attention to economic frictions between China and the United States, to most of the countries marked in white which have better economic conditions and correspondingly pay more attention to political friction between China and United States. In other words, the economic conditions of a country may affect the amount of attention people paid to economic frictions and political frictions. Then, we pay specific attention to those countries in white, which includes two subgroup countries: English-speaking countries and Eastern Asian countries. While most Eastern Asian countries are not English-speaking countries, since the protest erupts in Hong Kong, the English keywords searching about Hong Kong protest in these countries also rise, which reflects the huge attention people paid in these Eastern Asian countries. Finally, compared to the countries marked in red and countries marked in white, people in both the countries marked in black and brown pay less attention to the trade war of China and the United States. We can find that countries in black and brown are either slightly affected by US-China trade war or even beneficiary from the trade war. In other words, since these countries are almost immune from negative economic effects of US-China trade war, people in these countries relatively pay less attention to the trade issue than that of the countries in red and white.

Advantages and Limits of the study.

This study has attempted to compare people's attention in different countries about topics of frictions between China and the United State in 2018 and 2019. By collecting searching frequencies from Google trend data, this study has demonstrated the similarities and differences of available countries about people's attention to various topics. However, there are some limits within this study. First, because of the occurrence of the missing value for the interested keywords, the number of countries which can be used for comparison is limited. In particular, missing values causes significant drop-out of countries, for example, only about 15 countries have available data for analyzing the relationship of searching frequencies about the Huawei keywords. There might be potential patterns for the countries without searching records of Huawei keywords. Furthermore, since keyword searching can only partially explain people's attention to certain public issues, it is necessary to collect data from other media, which may contribute to more comprehensive understanding about people's attention. Finally, because of limited access to internet or official bans, people in many countries have no access to Google search, which may cause potential biases of the searching record data.

References

- 1.Many firms have no contingency plans should US-China trade war worsen: Survey, FOXBusiness
- 2.Hong Kong Human Rights and Democracy Act of 2019, Pub. L. No. 116-76
- 3.Placing restrictions on teargas exports and crowd control technology to Hong Kong Act, H.R.4270, 116th Cong. §1 (2019)

4. Stand with Hong Kong Resolution, H.R. 543, 116th Cong (2019)
5. Commending the Government of Canada for upholding the rule of law and expressing concern over actions by the Government of the People's Republic of China in response to a request from the United States Government to the Government of Canada for the extradition of a Huawei Technologies Co., Ltd., executive, H.R. 521, 116th Cong (2019)
6. U.S. House takes hard line on China over Hong Kong, Huawei, Reuters.
7. Huawei CFO Meng Wanzhou arrested in Canada, faces extradition to United States, CNN
8. Search Engine Market Share Worldwide (Nov 2018- Nov 2019), Statcounter, <https://gs.statcounter.com/search-engine-market-share>