Social Media Analysis on Panic Buying in the COVID-19 Pandemic

Alex Yinan Guo

Department of Economics, Keio University

Fukuhara Research Seminar

Supervisor: Masahiro Fukuhara, Department of Economics, Keio University

Co-supervisor: Peter Romero, Graduate School of Economics, Keio University

July 31, 2021

Abstract

This paper aimed to explore people's collective response to the COVID-19 outbreak and the behavior of panic buying with the impact of social media. To achieve this goal, sales data of facemasks in Tokyo region from the end of 2019 to mid-2020 was analyzed to explore people's changes of purchasing patterns and consumption behaviors during the pandemic. In addition, this paper examined social media data from 7,435 Japanese Twitter users and 23,501,189 Tweets from 2009 to mid-2020 to inspect impacts of social media factors, i.e. extracts topics, keywords, and sentiment, on panic buying.

Keywords: COVID-19; social media analysis; topic analysis; sentiment analysis; panic buying

Social Media Analysis on Panic Buying in the COVID-19 Pandemic

1 Introduction

The global COVID-19 pandemic became the most severe public health crisis for humankind in the 21st century and presented challenges to our society in numerous aspects. In December 2019, cases of pneumonia caused by a newly identified β-coronavirus occurred in Wuhan, China. It has soon spread and escalated into a global pandemic. The World Health Organization officially named the new virus Coronavirus Disease 2019 (COVID-19) on February 11, 2020 (Guo et al., 2020).

Survival psychology states that individuals may experience behavioral changes due to major events such as natural disasters and disease outbreaks, which have the potential to disrupt normal lives or pose a health risk (Leach, 1994). Such behavior was witnessed in past major disasters or health crises (Ding, 2014; Forbes, 2016). Similarly, the COVID-19 pandemic and the corresponding lockdown and social distancing had caused changes in consumer behaviors, and one is panic buying (Sheth, 2020; Islam et al., 2021). Panic buying and the corresponding spikes in demand for consumer goods provide challenges for ordering, replenishment, and distribution, leading to disruptions of supply chains (Zheng et al., 2021). In addition to adversities in retail, individuals and more vulnerable groups in greater need of the products could not obtain them due to panic buying situations, creating negative externalities in societies (Wesseler, 2019).

The rapid and vast adoption of SNS platforms enables online social interactions. It has become the platform for individuals to acquire information, create connections, and build communities. SNS platforms could play essential roles in understanding public attitudes and behaviors during a crisis (Shah et al., 2019; Sinnenberg et al., 2017; Steffens et al., 2019). The analysis of social media data could help us better understand how users react and communicate during the pandemic (Jordan et al., 2018).

Thus, this paper would like to offer a viewpoint to explore this abrupt change of consumption behavior from the perspective of social media analysis. By studying censurers' collective response on social media platforms, we hoped to provide an insight to the supply chain management for better preparation under similar circumstances. Furthermore, we would like to present a possibility for the field of behavioral economics in terms of social media analysis. To explore if the dynamics of social media users would reflect on consumption behaviors during the pandemic, we extracted and examined keywords, sentiment information, and topics discussed by Japanese users on Twitter. In addition, we demonstrated panic buying behavior via analysis of sales data of face masks in the Tokyo region. Moreover, the relationship between social media factors and consumption behaviors were explored.

2 Literature Review and Hypotheses

2.1 Social Networking Services (SNS)

Because of the origin of COVID-19 and major inconveniences it has brought to people's lives, COVID-related posts online tend to display negative emotions in many regions around the world (Nemes & Kiss, 2021; Wang et al., 2020). The outbreak of COVID-19 could be interpreted as a sudden change of environments, and people would respond to such change by entering an "emergency mode" by changing their perception, cognition, and behavior accordingly (Romero et al., 2021). Past research showed that alternations in behaviors and communication modes after personal tragedies (Mechanic, 1986) and natural disasters (Finau et al., 2018; Savage, 2019). Recent laboratory studies further demonstrated people's increasing desire for communication when facing a common simulated enemy (De Jaegher, 2020). With the outbreak of COVID-19, such corresponding change in communication would affect not only social-distancing individuals who become more active and spend more time online during the pandemic but also the overall information field and collective response on SNS (Abd-Alrazaq et al., 2020; Islam et al., 2021). Therefore, this paper hypothesized:

H1 More COVID-19-related contents leads to overall increasing negativity on SNS.

2.2 Panic Buying

Panic buying is a socially unwelcomed herd behavior (Steven, 2015, p.12) in which excessive quantities of everyday essentials and medical supplies are purchased from markets, resulting in stockouts. One possible explanation for panic buying suggested by Yuen, Wang, Ma, and Li (2020) was fear of the unknown, which would amplify the perception of threat, leading to panic buying as a means of self-protection during the pandemic. In addition, the sudden changes and inconveniences brought by COVID-19 led to more negative emotions, which would trigger panic buying as a coping mechanism to regulate negativity. Panic buying under such circumstances could also be interpreted as displacement activity (Bacon & Corr, 2020). Instead of facing uncertainties and accumulating negative feelings, people would instead distract themselves and take actions such as hoarding to regain a sense of control over the situation. Thus, we propose:

H2 During COVID-19 pandemic, increasing negative sentiment on SNS causes increasing purchases of COIVD-related products.

Informational conformity and observational learning were two other possible contributors to panic buying (Yuen et al., 2020). Individuals are members of society, and social environments inevitably influence their behaviors. Informational influence is the type of conformity (Deutsch & Gerard, 1955), referring to the inclination to accept messages from others as guidance when confronting complicated information and simplifying the decision-making process (Aronson & McGlone, 2009). Conformity behavior maximizes the possibility of effective action while minimizing the cost of one's cognitive resources during confusing situations (Chartrand & Bargh, 1999). Previous studies demonstrated that informational conformity could be a contributor to panic buying under life-threatening crises. Murray and Schaller (2012) stated that when consumers face the threat of infectious disease, their consumer behaviors tended to be irrational and guided by the group's information (Murray & Schaller, 2011). Song et al. (2020) indicated that information shared during a health-related crisis driven by self-protection as a social motivation established the basis for consensus among groups and influenced the behavior of group members during health-threatening crises. More consumers' choices were based on the information from others to help them mitigate the threat of death (Song et al., 2020). Observational learning refers to a phenomenon where people observe the behavior of others and then make the same decision sequentially, leading to a herd behavior (Shettleworth, 2010). As one can only observe, one has no access to the outside information probably owned by others. This lack of knowledge could cause some individuals to believe that the majority has a better assessment of the situation, think that the behavior of the majority is the optimum choice, override their own opinions, and in turn copy the group behavior, i.e., panic buying.

During the COVID-19 pandemic, SNS platforms has enabled individuals to form online social groups and communities beyond physical limitations in which individuals could build connections and share information. People's decision-making, including their purchasing decisions, would be consequently affected by attitudes, opinions, and perspectives from SNS platforms.

Therefore, we assumed that what people talk about on SNS could significantly influence consumers' decisions making, and thus the third hypothesis was:

H3 Occurrence and higher frequency of COVID-related keywords and topics encourage panic buying behaviors.

3 Methodology

3.1 Data

For the analysis, the following data was collected, and features were created accordingly using Excel, Python, and R:

1. Daily Sales of facemasks in the Tokyo region (Appendix A);
2. Keywords, categories, and sentiment features from Twitter Data (Appendix B & Appendix C);
3. COVID-related events (Appendix D);
4. Additional data that might affect sales of facemasks (Appendix E).

3.1.1 Sales Data

We generated features from sales data to further examine possible contributors to alternations in consumption behaviors. For instance, this paper engineered daily sales features and daily share of sales features based on materials, sizes, design, package size, and manufacturers of facemasks to inspect whether types of facemasks would change people's purchasing decisions during the pandemic. Moreover, we generated daily percentage changes of each feature mentioned previously because consumers could be more sensitive to the evolving situations at stores purchasing utilitarian goods. Past research demonstrated that the perception of shortage was one possible reason for panic buying (Yuen et al., 2020). Perceived scarcity of goods would motivate individuals to engage in panic buying due to psychological reactance and anticipated regret. Thus, we assumed that a rather significant decrease or shortage of certain products could trigger changes in consumer behaviors.

3.1.2 Twitter Data

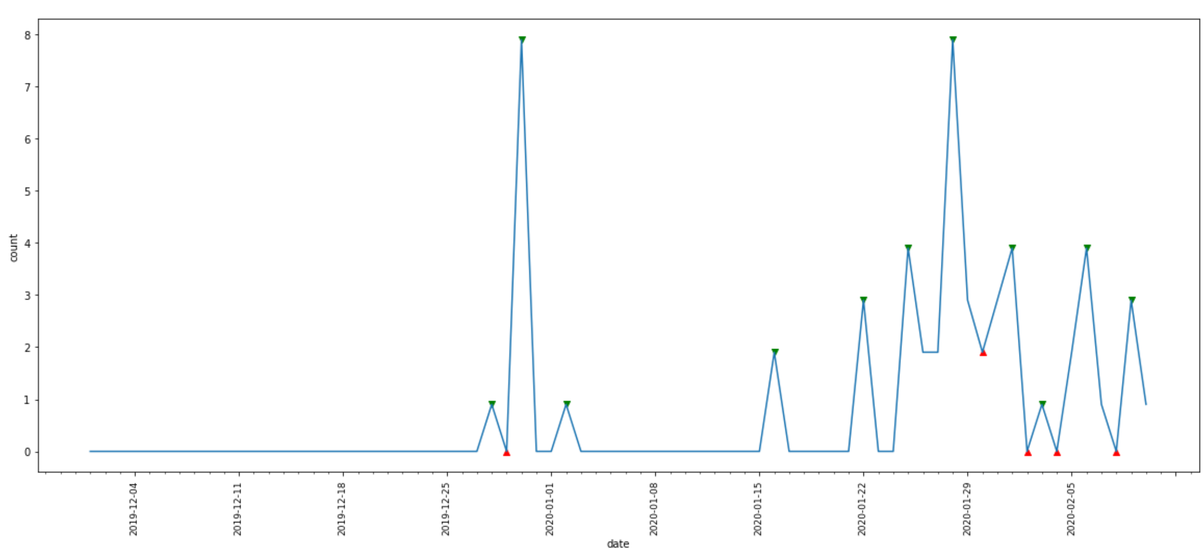
We collected the recent 3,200 Tweets from all Japanese Twitter users who tweeted between January and June 2020 about the following COVID-related keywords: "Corona," "COVID," "Pandemic," "Facemask," "Hand Sanitizer," "Hand Soap," "Toilet Paper," "Napkins," and "Tissues." This collection method allowed us to capture a broader range of the phenomena over and above purely COVID-19-related Tweets. In total, we had 7,435 Twitter users and 23,501,189 Tweets. Since we assumed that there was mutual influence among users in SNS platforms, we analyzed daily Twitter text data combined, i.e., the focus was on a macro collective response on social media instead of individual ones.

The choice of tools for Japanese language analysis was limited due to the infancy of psycholinguistics in Japan (Romero et al., 2021). The dictionary for Japanese Linguistic Inquiry and Word Count (LIWC) was still developing, and other approaches such as translating the Chinese version of LIWC were not ideal (Guntuku et al., 2019; Shibata et al., 2016). Therefore, this paper chose the commercial version of IBM Watson Natural Language Understanding (NLU), a natural language processing (NLP) service for text analysis, to extract keywords, categories, and sentiment features from daily total Twitter data (Appendix F).

3.1.2.1 Topics Analysis

Using IBM NLU, we had topics data extracted from Twitter text from the end of 2009 to mid-2020. In this dataset, every day, there were specific topics picked from a pre-existed list of categories with hierarchy, called 'taxonomy'. As this research was conducted on a collective base, there was no frequency of topics mentioned by individuals, making feature engineering, especially for numerical features, rather challenging. The solution proposed was to create daily coverage of topics based on their hierarchies to explore whether certain topics reach a broader range of populations during the pandemic. To be more specific, "health and fitness" was a level-1 category with 52 sub-categories. We calculated the daily sum of occurrence of topics to which belong "health and fitness," and this sum of occurrence was divided by 52 to generate the daily coverage of "health and fitness." Another two sets of features created for topic data were peaks of coverage and intervals between two peaks. The idea is illustrated in Figure 1 with a time series plot of the topics "health and fitness". A Python code was created to detect peaks of coverage of each topic, which are green arrows in Figure 1. These peaks were marked in binary variables. Additionally, lengths of intervals between each peak were calculated as numerical variables. Decreasing cadence, i.e., length of intervals, indicated that the topic was getting more popular. Topic-based features also contained binary variables representing the occurrence of each topic and daily percentage change of coverage of topics.

Figure 1
Illustration of the Feature "Peak" in Topics Analysis.



3.1.2.2 Keywords and Sentiment Analysis

Using IBM NLU, we had keywords and sentiment data of the same period from the Twitter text of Japanese users. This dataset included sentiment labels (positive, neutral, or negative), sentiment scores, and counts of keywords on a daily basis. The authors organized this dataset into the following features: daily overall sentiment score in average, the daily count of three sentiment labels respectively, binary variables for each keyword to detect positivity, binary variables for each keyword to detect negativity, the daily count of each keyword, and daily percentage change of keyword counts. This approach of feature organization allowed us to discover sentiment details in both macro and micro scopes since we had features representing overall daily sentimental changes as well as sentimental fluctuations of individual keywords.

3.1.3 COVID-19-Related events

To explore influences of big events to consumption behaviors, this paper used the list from Romero et al. (2021) to eleven features from major pandemic-relevant events extracted from Japan and foreign news (Appendix E). Even though the duration of news framing effects tends to last approximately one to two weeks depending on individuals (Lecheler & de Vreese, 2011), considering that COVID-relevant news is frequent during the pandemic, we assumed that the length of influence of each news would be three days. Accordingly, we labeled the day of the event and the following two days of the event as 1 in each variable.; the other days were marked as 0.

3.1.4 Additional Variables

In order to specify influences on the sales of facemasks and rule out other external factors, this paper included a daily amount of PM 2.5 and a binary variable to represent rainy days in the data to control impacts from weathers to sales of facemasks. Binary variables representing common paydays and national holidays in Japan were included as well since some customers tend to make purchases after receiving paychecks or during day-offs.

3.2 Time-based features

To address the time-series nature of our datasets, this paper included time components when using machine learning based models and GLM methods to conduct analysis. Thus, we created features such as lags, the difference between lag values and current values, rolling means, and the difference between rolling means and current values to capture time dynamics for social-media-based variables and sales-based variables. In addition, features representing days of the week were also included.

3.3 Data Visualization

3.3.1 Sales of Facemasks

Firstly, there was an obvious and sudden increase in sales of facemask illustrated in Figure 2, indicating the existence of panic buying during the pandemic. Furthermore, we discovered before the end of January 2020, shares of sold facemask units by materials (Figure 3), sizes (Figure 4), and design (Figure 5) of facemasks and shares of sold packages by package size (Figure 6) were all rather stationary, indicating that there was no change in preferences of consumers. However, in the end of January 2020, i.e. the same time panic buying occurred, daily shares of sales suddenly started fluctuating drastically. It seemed that people just purchased any facemask they could get hand on without considering types of facemasks or size of packages during panic buying. In addition, it appears that average price of facemasks per unit increased (Figure 7), possibly driven by higher demands and limited supplies.

Figure 2

Daily Sales of Facemasks in Unit

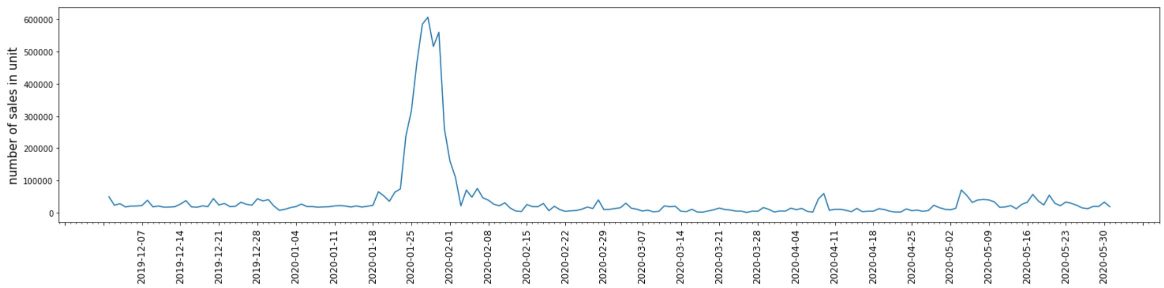


Figure 3

Shares of sales by materials of facemasks

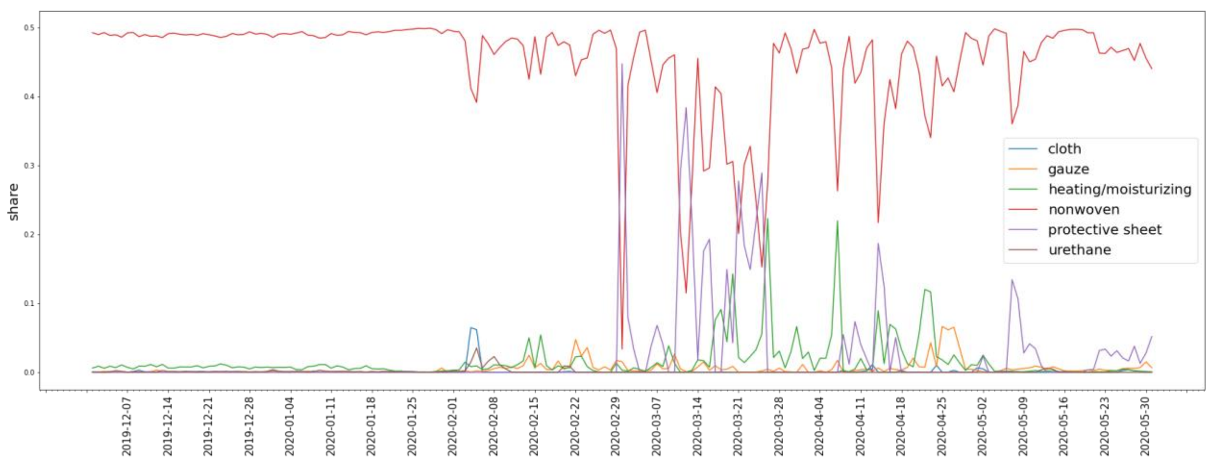


Figure 4

Shares of Sales by Sizes of Facemasks

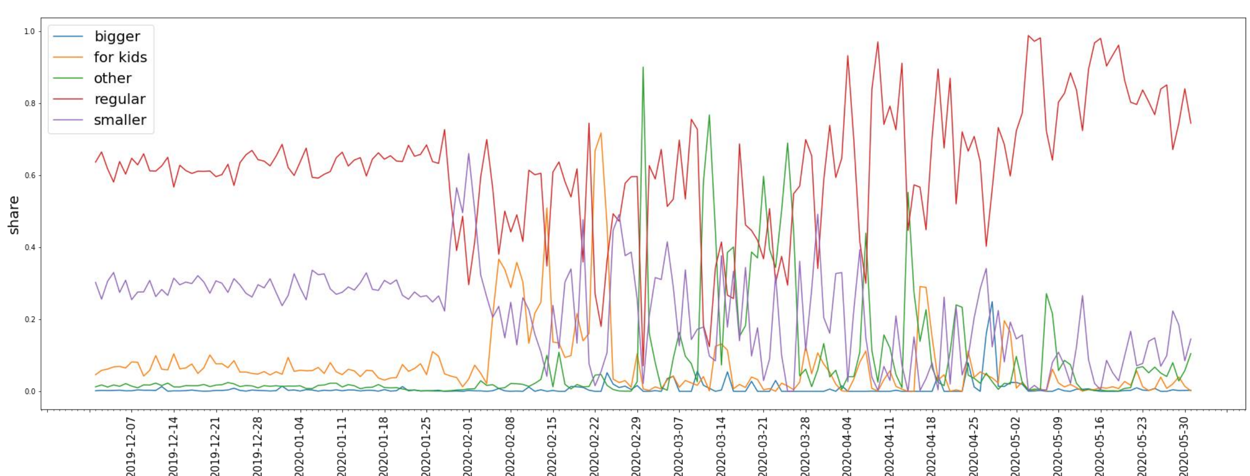


Figure 5

Shares of Sales by Design of Facemasks

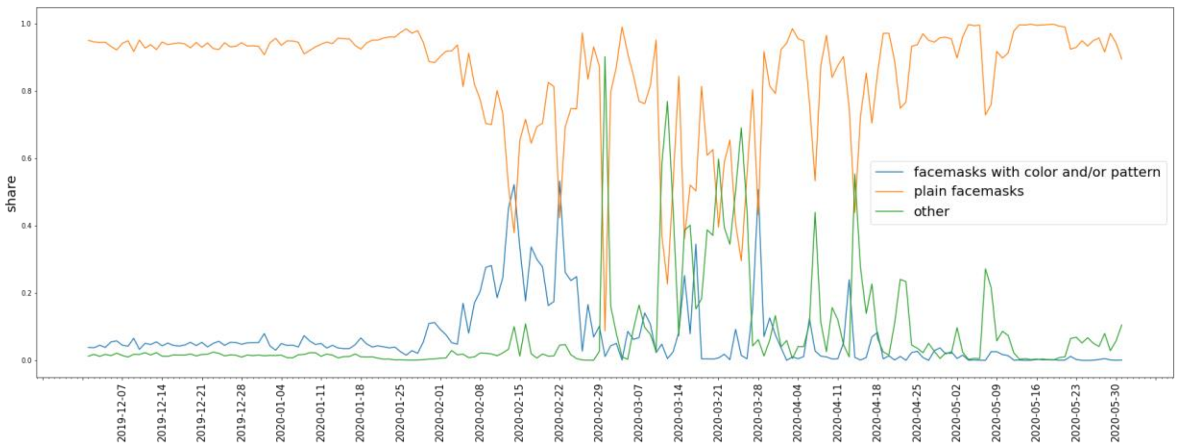


Figure 6

Shares of Sold Packages of Facemasks by Package Sizes

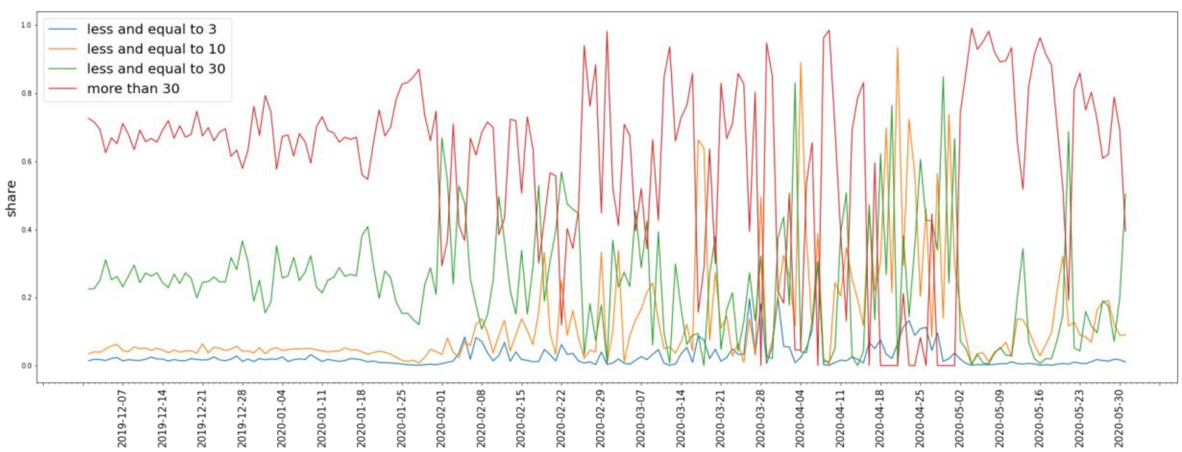
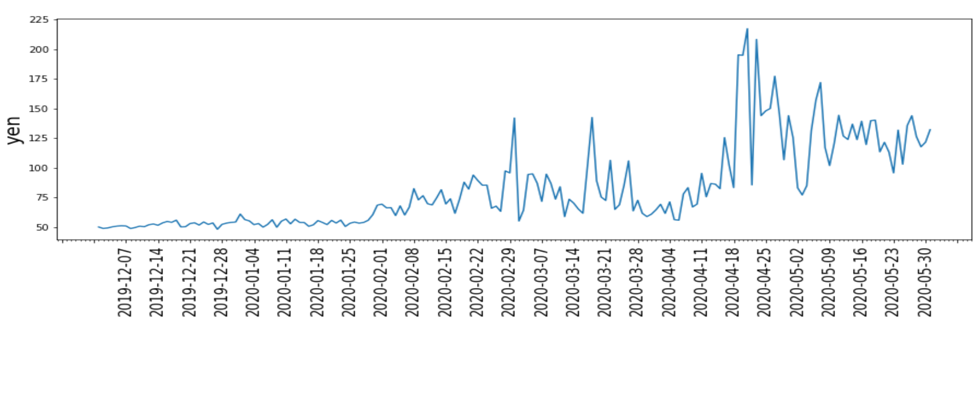


Figure 7

Daily Average Price of Facemasks per Unit

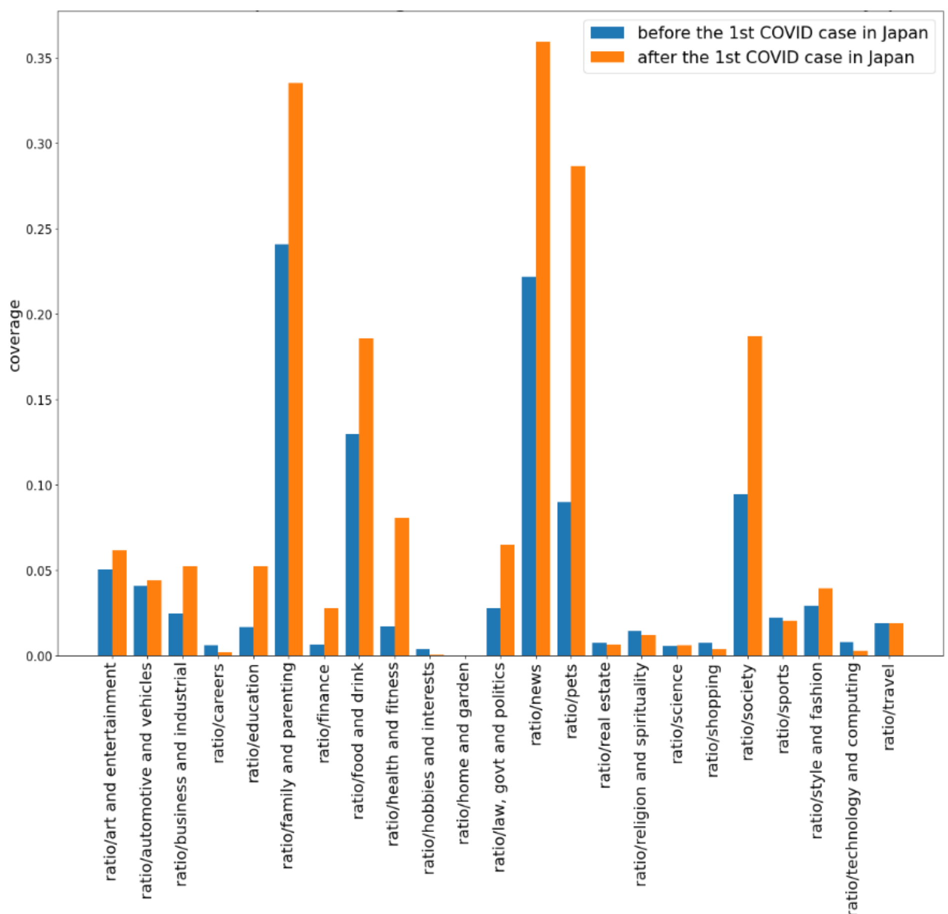


3.3.2 Topics

From the visualization of topics data, we could observe that there was obvious difference in topic coverages (see Figure 8 & Figure 9). Topics with increasing coverages were rather more relevant to the present, while topics with decreasing coverages tended to be more future-related. We could say that during the pandemic, people tend to focus more on the current topics and "live in the moment" during the pandemic.

Figure 8

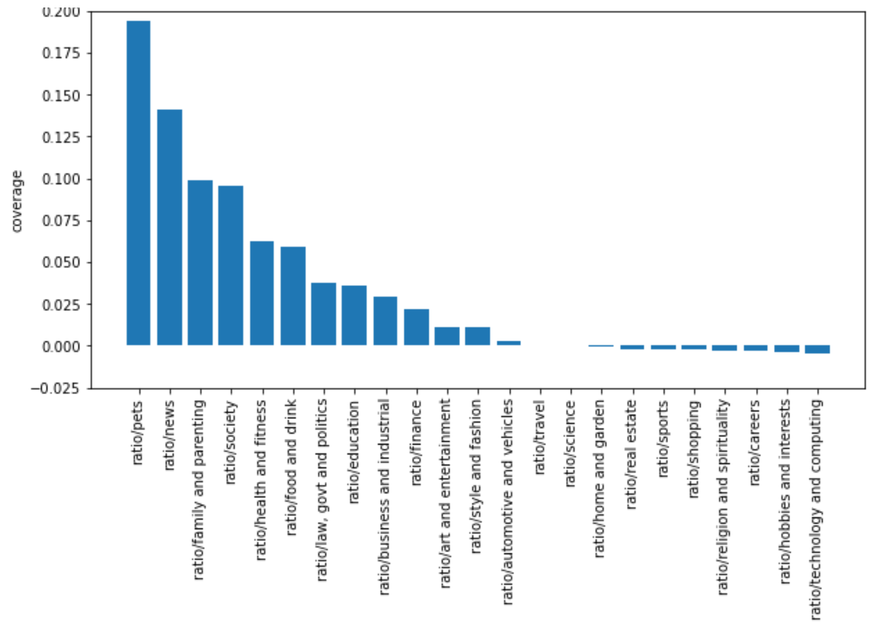
Difference in Topics Coverage: Before and After the 1st COVID Case in Japan



Note. Topics were extracted from Twitter text data from 2009 to 2020 using IBM Watson NLU.

Figure 9

Difference in Average Coverage Before and After the 1st COVID Case in Japan



Note. Topics were extracted from Twitter text data from 2009 to 2020 using IBM Watson NLU. Positive values indicated that the coverage increased after the 1st case.

3.3.3 Keywords and Sentiment

To testify H1, we conducted comparison and contrast between COVID-19-related keywords and the rest keywords. The time period of this dataset began from the pandemic to February 2020. Table 1 demonstrated that the first three keywords display higher negativity compared to the mean value of the rest keywords. Figure 10 also indicated that COVID-19-related keywords had rather large percentage of negative labels overall. Thus, we concluded that H1 is valid.

Table 1

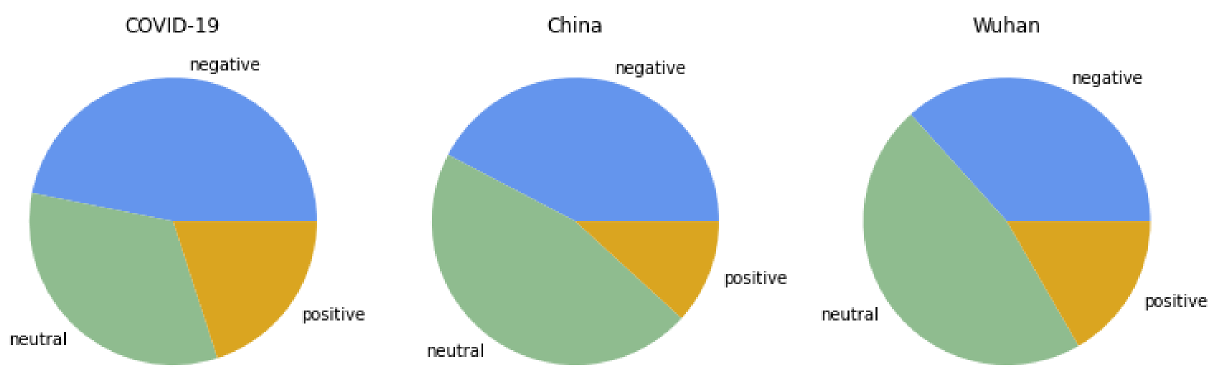
Sentiment Information of Covid-19-Related Keywords and the Average Value to the Rest

	新型コロナウイルス (COVID-19)	中国 (China)	武漢 (Wuhan)	Other Keywords
Mean Sentiment Score	-0.107186	-0.285291	-0.062097	0.081098
Ratio of Negative Label	0.615385	0.631579	0.6	0.285661

Note: Here we picked COVID-19, China, and Wuhan as three representative COVID-19-related keywords. Sentiment score was a measurement for the sentiment of keywords range from 0-1 to 1, in which -1 indicated complete negativity and 1 indicated complete positivity. Ratio of negative label was calculated as the number of negative labels divided by the total number of occurrence of keywords.

Figure 10

Total Number of Sentiment Labels of COVID-19 Related Keywords



Note. “COVID-19” percentage of “negative”, “neutral”, and “positive” was 0.47, 0.33, and 0.2, respectively. “China” percentage of “negative”, “neutral”, and “positive” was 0.42, 0.46, 0.12, respectively. “Wuhan” percentage of “negative”, “neutral”, and “positive” was 0.37, 0.47, 0.17, respectively.

In addition, we examined the collective sentiment changes since the pandemic. The average daily sentiment score had been decreasing (see Figure 11), and more negative keywords were mentioned as time went by (see Figure 12). A drop around the confirmation of the 1st COVID-19 case in Japan (January 16th, 2020) was observable in Figure 5 as well. Figure 13 and Figure 14 are word clouds before and after the 1st COVID-19 case in Japan. We could see that since Figure 12 was generated based on Twitter text around the end of 2019, keywords such as “クリスマス (Christmas)” , “新年 (New Year)” , and “プレゼント (present)” were frequently mentioned. “新型コロナウイルス (COVID-19)” were also in Figure 13, although the frequency was not high. In Figure 13, the frequency of “新型コロナウイルス (COVID-19)” increased comparatively.

Figure 11
Daily Sentiment Label Counts

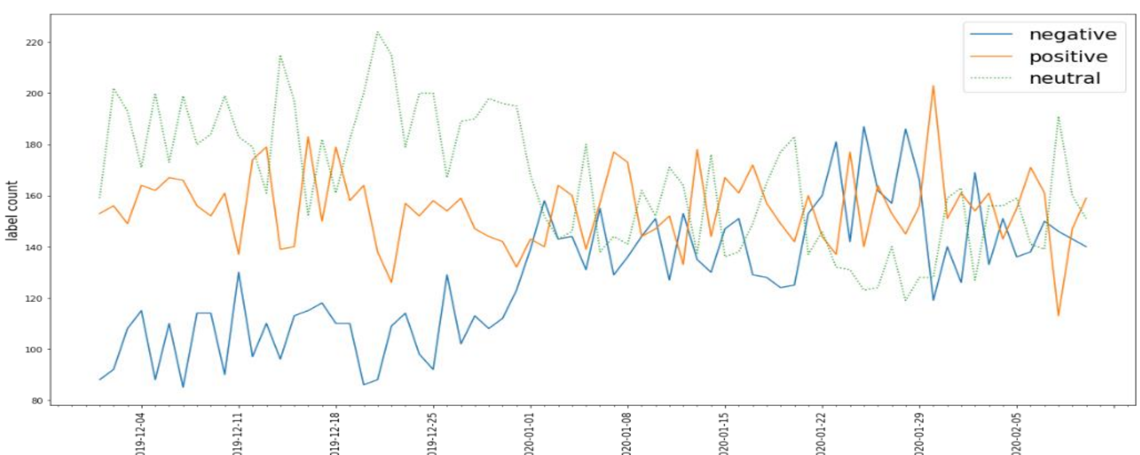


Figure 12
Daily Average Sentiment Score

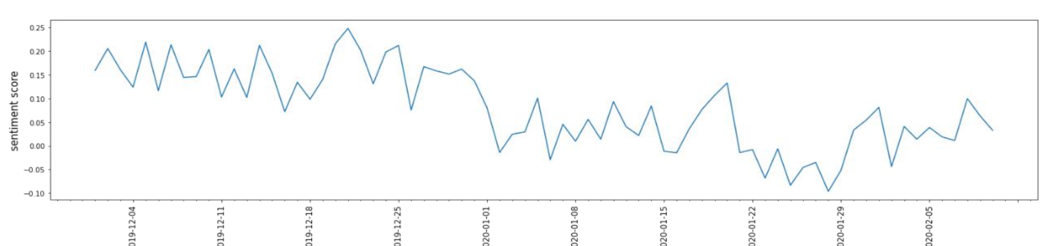


Figure 13

Word Cloud from the End of 2019 to the 1st COVID-19 Case in Japan

Note. The size of keywords in the figure were generated from frequencies of keywords from Twitter text data.

Figure 14

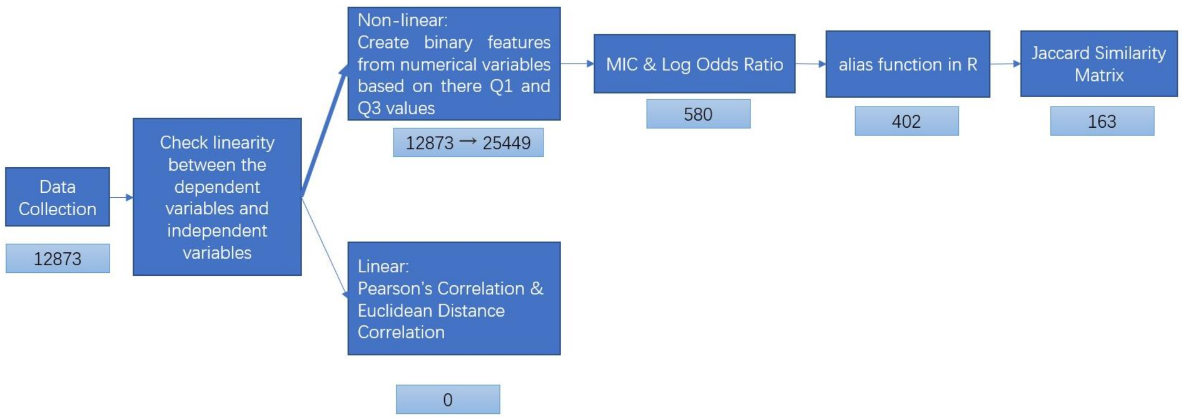
Word Cloud from the 1st COVID-19 Case in Japan to February 2020

Note. The size of keywords in the figure were generated from frequencies of keywords from Twitter text data.

3.4 Feature Engineering and Selection

Figure 15

Flow of Data Engineering and Selection



Firstly, we labeled days with sales of facemasks higher than 150% of the average daily sales as “spike” because our focus is the panic buying behavior. Thus, instead of using numerical variables such as daily sales of facemasks, we chose the binary “spike” as the dependent variable. After initial data cleaning and organization, we retrieved 11,454 variables from datasets mention before. The subsequent feature engineering was conducted in the following steps (see Figure 15):

1. Remove features with less than 5% observations
2. Remove features for which less than 0.02 percent of the observations differ from the mode value
3. Transfer numerical variables into three binary variables
   (a) Values above Q3 are marked as “high”
   (b) Values between Q1 and Q3 are marked as “medium”
   (c) Values below Q1 are marked as “low”

After feature engineering, we had 25,449 binary features. Next, following the flow (Figure 1), we check the linearity between the dependent variable “spike” and all other features. Since all were non-linear relationships, we moved on to calculate Log-Odds ratios and Maximal Information Coefficient (MIC) between “spike” and other features. Both of which allowed us to examine the strength of non-linear association between the dependent variable and independent variables. We removed features with MIC values less than 0.1 and with absolute values of Log-Odds ratio less than 0.1 and then compared two datasets to find the common 580 features.

The next filter is the alias function in R, which detects linearly dependent terms. We deleted features returning 1 or -1 to prevent multicollinearity. This left us with 402 features. The following step is Jaccard Similarity Matrix, a matrix of similarity coefficients for binary terms, in which

$$J(A,B) = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|}.$$

The Jaccard coefficient, *J(A, B)*, is between 0 and 1 by design. Here, since we compared sets of binary variables, $|A \cap B|$ equals cases when A=B, and $|A \cup B|$ is the length of the dataset. Based on Jaccard Similarity Matrix, we deleted features with Jaccard coefficient less than 0.5, and removed some features that are highly correlated with other explanatory variables to prevent multicollinearity in the GLM models later. This screening left us with 163 features.

3.5 Models

Using the filtered 163 features, we created ***M1*** and ***M2*** by step wise logistic regression and individual feature significance. ***M3*** is generated by Random Forest.

M1: spike = -6.929 + 2.456*count.武漢.high + 3.115*Lag1.count.新型コロナウイルス.high + 2.526*count.中国.high + 2.421*Lag1.health.and.fitness.disease

M2: spike = -6.204 + 2.481*negative.武漢 + 3.095*Lag1.positive_label_count.low + 3.095*Lag1.nonwoven_share.high

For ***M1*** and ***M2***, we first ran a two-direction stepwise logistic regression with 163 variables to detect potential variables with high significance. Then, we ran single logistic regressions with the 163 variables once at a time to examine their individual significance. We then ruled out features with high individual p-values and/or with 0 in confidence intervals. The previous two screening process left us with 92 features. We then divided them into batches by their themes and sentiment. Finally, we selected at most one feature from each batch (to decrease the possibility of multicollinearity) and found a combination of features that yielded highest significance by P-values of the F-test of each variable and the relatively low AIC value.

Table 2
M1 Results

```
glm(formula = spike ~ count.武漢.high + Lag1.count.新型コロナウイルス.high
+ count.中国.high + Lag1.health.and.fitness.disease, family = binomial, data = df)

Deviance Residuals:
Min        1Q     Median      3Q       Max
-1.64889  -0.15065  -0.04424  -0.04424   2.08479
Coefficients:
Estimate Std. Error z value Pr(>|z|)
(Intercept)                      -6.929      2.296  -3.018  0.00254 **
keywords_df1231.high              2.456      1.423   1.725  0.08444 .
keywords_df2995.high              3.115      1.588   1.962  0.04979 *
keywords_df1051.high              2.526      1.461   1.729  0.08375 .
Lag1..health.and.fitness.disease  2.421      1.455   1.664  0.09616 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
Null deviance: 53.435  on 68  degrees of freedom
Residual deviance: 17.520  on 64  degrees of freedom
AIC: 27.52

Number of Fisher Scoring iterations: 8
```

Table 3
M2 Results

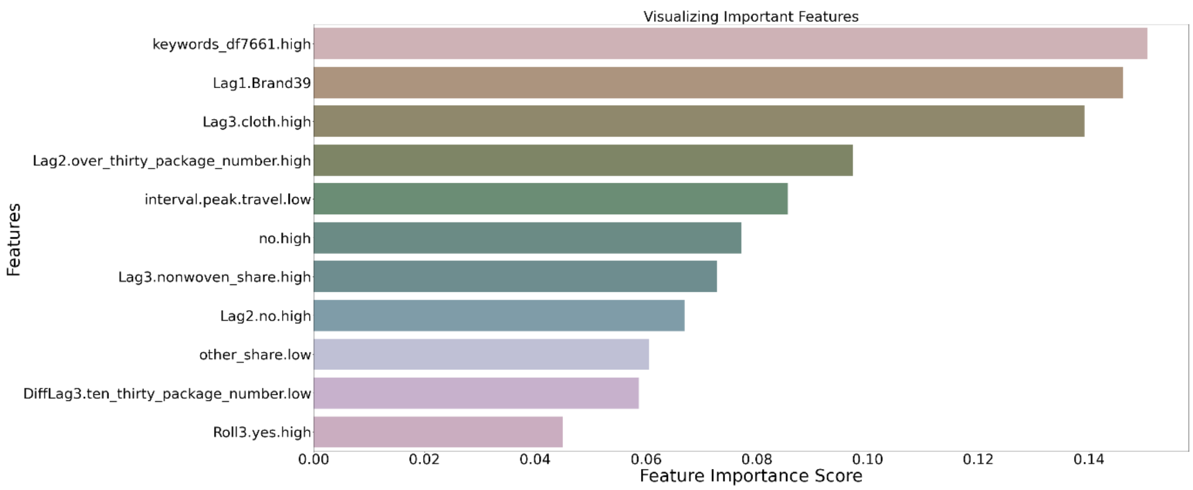
```
glm(formula = spike ~ keywords_df300 + keywords_df1829.low +
Lag1.nonwoven_share.high, family = binomial, data = df)

Deviance Residuals:
    Min        1Q    Median        3Q       Max
-1.17131  -0.06354  -0.06354  -0.06354   2.51118
Coefficients:
                         Estimate Std. Error z value Pr(>|z|)
(Intercept)                -6.204      1.975  -3.142  0.00168 **
keywords_df300              2.481      1.238   2.004  0.04506 *
keywords_df1829.low         3.095      1.401   2.209  0.02720 *
Lag1.nonwoven_share.high    3.095      1.401   2.209  0.02720 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
Null deviance: 53.435  on 68  degrees of freedom
Residual deviance: 19.683  on 65  degrees of freedom
AIC: 27.683
Number of Fisher Scoring iterations: 7
```

M3 is generated from function “RandomForestClassifier” from the package “sklearn” on Python. Firstly, we inputted all 163 features from the previous step into one random forest model. Then, according to feature importance based on the mean decrease in impurity (MID), or Gini Importance, we selected the first 12 features and ran another random forest model using these 12 features (Figure 16). Both individual importance of features and the in-sample prediction accuracy improved the second round.

Figure 16

Feature Importance using MDI



4 Results

4.1 Model Interpretation

In ***M1***, from the positive coefficients of "count.武漢.high", "Lag1.count.新型コロナウイルス.high", and "count.中国.high", we could say that when Twitter users talk about "Wuhan", "COVID-19", and "China" more often, it is more likely that panic buying occurs. Similarly, the positive coefficient of "Lag1.health.and.fitness.disease" demonstrates that when people tweet content relevant to disease, panic buying is more likely to occur the day after. According to the P-values, all variables are significant to bring an impact to the dependent variable. Therefore, we could confirm that H2 is valid. More CIVID-19-related topics and keywords on SNS encourage panic buying behavior.

In ***M2***, the positive coefficient of "negative.武漢" indicates that the negative emotions displayed when people post the keyword "Wuhan" online increase the possibility of panic buying. Exp(-6.204)-1 equals 10.66, indicating that the variable "negative. 武漢" as 1 has 10.66 more odds of “spike”, holding other variables constant. In other words, if the sentiment of “Wuhan” is negative, the odds of panic buying will increase 10.66 times. Moreover, the positive coefficient “ Lag1.positive_label_count.low” represents that when fewer people talk about positive keywords during the precious day, panic buying will have a higher possibility to occur the next day. The results from ***M2*** enable us to conclude that H2 holds as well. Increasing overall negativity on SNS could cause panic buying. In addition, in ***M2***, a sales-related feature demonstrates significance as well. The positive coefficient of “Lag1.nonwoven_share.high” demonstrates that higher shares of sales of facemasks made of nonwoven materials could stimulate consumers’ urge to hoard, which also confirmed our assumption about the impact of perceived scarcity on consumption behavior.

In terms of ***M3***, according to Figure 16, “keywords_df7661.high” (Roll3.count.コナウイルス) demonstrates the highest MDI importance in ***M3***, meaning that the previous three days’ mean of the count of the keyword “COVID-19” could have a relatively strong impact on the current day’s consumption behavior. More discussion of “COVID-19” on Twitter in the past three days could encourage consumers’ panic buying behavior today. We could notice that “Brand 39”, i.e. products from the manufacturer “ユニフリー”, displays relatively high importance as well. One possible reason is most facemasks of this manufacturer are with large package sizes and one of the lowest prices per unit in the market. We could also see that among the three models, “interval.peak.travel.low” displays importance as the only keyword/topic that is not directly related with COVID-19, such as “COVID-19”, “China”, “Wuhan”, “disease”, etc. This feature represents that during the pandemic, people talked less often about travel on Twitter, and it could be related to panic buying. The inability of recreational travel could build up anxiety and fear, leading to sudden changes of consumption behaviors. ***M3*** includes both significant sentiment features and topic/keyword features; hence H2 and H3 are further proved.

4.2 In-sample and Out-of-sample Validation

The total dataset was divided into training data and testing data to conduct validation of the three models. Confusion matrix, sensitivity, specificity, ROC, and AUC are calculated for three models, respectively. Detailed results are displayed in Table 4 and Table 5. Figure 17, Figure 18, and Figure 19 are the ROC of the three models. We observed overfitting in M3 from the rectangle-shaped ROC and AUC as 1. Overfitting also occurred in M1 out-of-sample validation.

Table 4

Summary of In-Sample Validation

	Sensitivity	Specificity	AUC
M1	0.75	0.96	0.9865
M2	0.625	1	0.973
M3	1	1	1

Table 5

Summary of Out-of-Sample Validation

	Sensitivity	Specificity	AUC
M1	1	1	1
M2	0	1	0.8889
M3	1	1	1

Figure 17

M1 ROC In-Sample Validation (Left) and Out-of-Sample (Right)

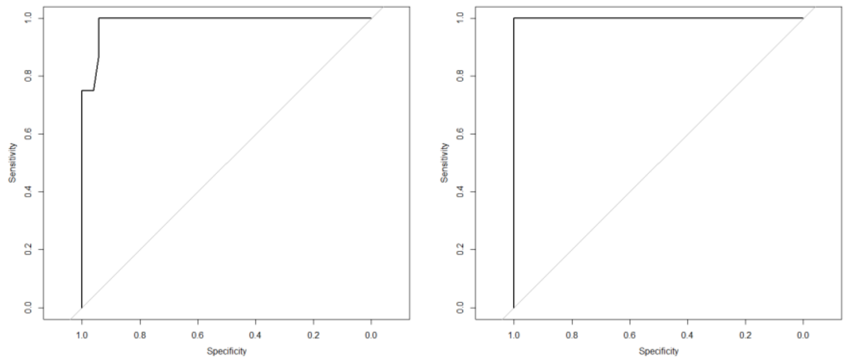


Figure 18

M2 ROC In-Sample Validation (Left) and Out-of-Sample (Right)

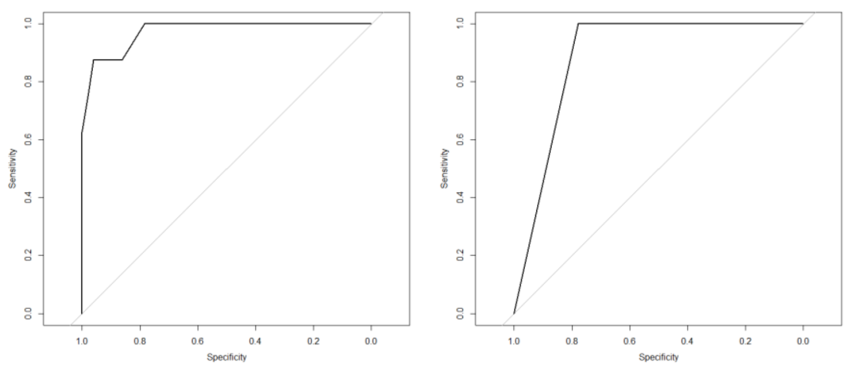
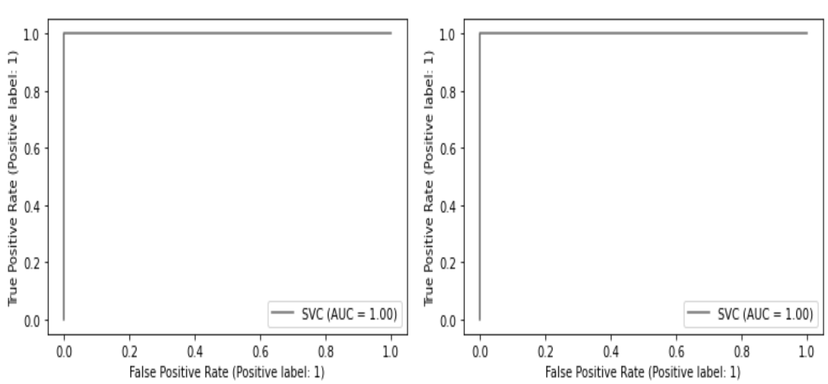


Figure 19

M3 ROC In-Sample Validation (Left) and Out-of-Sample (Right)



5 Discussion and Limitation

First of all, our initial plan for the final model is to combine previous models into one overall prediction model to explore SNS dynamics and consumption behaviors further. However, after the combination and feature selection from previous features, we could not find a good combination of features that satisfied sound comprehensiveness, high significance, and solid prediction results altogether. One contributing factor to this could be multicollinearity. Since we have created numerous lagged and rolling windows features, they could be mutually related, leading to multicollinearity in the models and corresponding poor results. In addition, as a post facto research, it is possible to leak information from the future into variables during feature engineering and model construction. Nonetheless, decent results yielded from M1 and M2 indicate that sub-components along display good performance. This paper could in turn valid H2 and H3 using these smaller models. On the other hand, we are aware that the validation is preliminary, and further research needs to be done to reach more convincing findings. This paper demonstrates that tree-based machine learning approaches are more practice-oriented, while regression-based perditions are more theoretic in nature. The limitation of our models indicates the need to use more dynamic approaches to combine the two such as Bayesian time space model to explore SNS data.

Secondly, the focus on the macro-level of SNS response toward the pandemic of this paper weakened the ability to analyze from observations on an individual level. This paper could not control the contextual conditions under which Twitter data was created nor sort individuals by personal situations. In addition, due to the lack of storage data, we were unsure whether specific changes in facemasks sales were results from SNS dynamics or inventory-related reasons. Future research could consider using inventory data as control variables to examine the relationship between social media and sales.

Finally, during the data collection, we collected 3,200 Tweets from each Twitter user who posted COVID-related keywords. However, because users had different posting frequencies, these 3,200 tweets could represent all contents for certain users while only recent several years or months of content for other users. This limitation is due to the restricted number of Tweets allowed to collect by the official Twitter API. Moreover, there was no demographic information either because of the restrictions from Twitter API. In this paper, we assumed that considering the high coverage and importance of Twitter in the Japanese social media market, our sample could be representative. In future research, additional demographic information would make the findings more accurate.

6 Conclusion

This paper utilized SNS data and sales data of facemasks to explore possible contributors to panic buying. To sum up, all three hypotheses proposed by this paper were examined to hold using two GLMs and one Random Forest model. The discussion around the pandemic has increased, leading the overall negativity on SNS to rise. Growing frequency and coverage of COVID-19-releted contents and the expanding negative sentiment on social media platforms could be possible drivers of panic buying.

References

Abd-Alrazaq, A., Alhuwail, D., Househ, M., Hamdi, M., & Shah, Z. (2020). Top concerns of tweeters during the covid-19 pandemic: Infoveillance study (preprint). https://doi.org/10.2196/preprints.19016

Aronson, J., & McGlone, M. S. (2009). Stereotype and social identity threat. In T. D. Nelson (Ed.), Handbook of prejudice, stereotyping, and discrimination (pp. 153–178). Psychology Press.

Bacon, A. M., & Corr, P. J. (2020). Coronavirus (COVID-19) in the United Kingdom: A Personality-based perspective on concerns and intention To self-isolate. *British Journal of Health Psychology*, *25*(4), 839–848. https://doi.org/10.1111/bjhp.12423

Chandra Guntuku, S., Li, M., Tay, L., & Ungar, L. H. (2019). Studying Cultural Differences in Emoji Usage across the East and the West. *Proceedings of the International AAAI Conference on Web and Social Media*, *13*(01), 226-235. Retrieved from https://ojs.aaai.org/index.php/ICWSM/article/view/3224

Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception–behavior link and social interaction. *Journal of Personality and Social Psychology*, *76*(6), 893–910. https://doi.org/10.1037/0022-3514.76.6.893

De Jaegher, K. (2020). Common-enemy effects: Multidisciplinary antecedents and economic perspectives. *Journal of Economic Surveys*, *35*(1), 3–33. https://doi.org/10.1111/joes.12394

Deutsch, M., & Gerard, H. B. (1955). A study of normative and informational social influences upon individual judgment. *The Journal of Abnormal and Social Psychology*, *51*(3), 629–636. https://doi.org/10.1037/h0046408

Ding, H. (2014). *Rhetoric of a global epidemic: Transcultural communication about sars*. Southern Illinois University Press.

Finau, G., Cox, J., Tarai, J., Kant, R., Varea, R., & Titifanue, J. (2018). Social media and Disaster communication: A case study of CYCLONE WINSTON. *Pacific Journalism Review : Te Koakoa*, *24*(1), 123–137. https://doi.org/10.24135/pjr.v24i1.400

Forbes, S. L. (2016). Post-disaster consumption: Analysis from the 2011 Christchurch earthquake. *The International Review of Retail, Distribution and Consumer Research*, *27*(1), 28–42. https://doi.org/10.1080/09593969.2016.1247010

Guo, Y.-R., Cao, Q.-D., Hong, Z.-S., Tan, Y.-Y., Chen, S.-D., Jin, H.-J., Tan, K.-S., Wang, D.-Y., & Yan, Y. (2020). The origin, transmission and CLINICAL therapies ON

CORONAVIRUS disease 2019 (COVID-19) OUTBREAK – an update on the status. *Military Medical Research*, *7*(1). https://doi.org/10.1186/s40779-020-00240-0

Islam, T., Pitafi, A. H., Arya, V., Wang, Y., Akhtar, N., Mubarik, S., & Xiaobei, L. (2021). Panic buying in the Covid-19 pandemic: A MULTI-COUNTRY EXAMINATION. *Journal of Retailing and Consumer Services*, *59*, 102357. https://doi.org/10.1016/j.jretconser.2020.102357

Jordan, S., Hovet, S., Fung, I., Liang, H., Fu, K.-W., & Tse, Z. (2018). Using twitter for public health surveillance from monitoring and prediction to public response. *Data*, *4*(1), 6. https://doi.org/10.3390/data4010006

Leach, J. (1994). *Survival psychology*. Macmillan.

Lecheler, S., & de Vreese, C. H. (2011). Getting real: The duration of framing effects. *Journal of Communication*, *61*(5), 959–983. https://doi.org/10.1111/j.1460-2466.2011.01580.x

Mechanic, D. (1986). The concept of illness behaviour: Culture, situation and personal predisposition. *Psychological Medicine*, *16*(1), 1–7. https://doi.org/10.1017/s0033291700002476

Murray, D. R., & Schaller, M. (2011). Threat(s) and conformity deconstructed: Perceived threat of infectious disease and its implications for conformist attitudes and behavior. *European Journal of Social Psychology*, *42*(2), 180–188. https://doi.org/10.1002/ejsp.863

Nemes, L., & Kiss, A. (2021). Social media sentiment analysis based on COVID-19. *Journal of Information and Telecommunication*, *5*(1), 1-15.

Romero, P., Mikiya, Y., Nakatsuma, T., Fitz, S., & Koch, T. (2021). Modelling personality change during extreme exogenous conditions. https://doi.org/10.31234/osf.io/rtmjw

Savage, D. A. (2019). Towards a complex model of disaster behaviour. *Disasters*, *43*(4), 771–798. https://doi.org/10.1111/disa.12408

Shah, Z., Surian, D., Dyda, A., Coiera, E., Mandl, K. D., & Dunn, A. G. (2019). Automatically appraising the credibility of Vaccine-related web Pages shared on social media: A Twitter SURVEILLANCE STUDY. *Journal of Medical Internet Research*, *21*(11). https://doi.org/10.2196/14007

Sheth, J. (2020). Impact of covid-19 on consumer behavior: Will the old habits return or die? *Journal of Business Research*, *117*, 280–283. https://doi.org/10.1016/j.jbusres.2020.05.059

Shettleworth, S. J. (2010). *Cognition, evolution, and behavior*. Oxford University Press.

Shibata, D., Wakamiya, S., Kinoshita, A., & Aramaki, E. (2016, December). Detecting Japanese patients with Alzheimer's disease based on word category frequencies. In *Proceedings of the Clinical Natural Language Processing Workshop (ClinicalNLP)* (pp. 78-85).

Sinnenberg, L., Buttenheim, A. M., Padrez, K., Mancheno, C., Ungar, L., & Merchant, R. M. (2017). Twitter as a tool for Health research: A systematic review. *American Journal of Public Health*, *107*(1), 143–143. https://doi.org/10.2105/ajph.2016.303512a

Song, W., Jin, X., Gao, J., & Zhao, T. (2020). Will buying follow Others ease their threat of Death? An analysis of consumer data during the period of COVID-19 in China. *International Journal of Environmental Research and Public Health*, *17*(9), 3215. https://doi.org/10.3390/ijerph17093215

Steffens, M. S., Dunn, A. G., Wiley, K. E., & Leask, J. (2019). How organisations promoting vaccination respond to misinformation on social media: A qualitative investigation. *BMC Public Health*, *19*(1). https://doi.org/10.1186/s12889-019-7659-3

Steven, D. (2015). *The new politics of strategic resources: Energy and food security challenges in the 21st century*. Brookings Institution Press.

Wang, T., Lu, K., Chow, K. P., & Zhu, Q. (2020). COVID-19 sensing: Negative sentiment analysis on social media in China via BERT Model. *IEEE Access*, *8*, 138162–138169. https://doi.org/10.1109/access.2020.3012595

Wesseler, J. (2019). Storage policies: Stockpiling versus immediate release. *Journal of Agricultural & Food Industrial Organization*, *18*(1). https://doi.org/10.1515/jafio-2019-0055

Yuen, K. F., Wang, X., Ma, F., & Li, K. X. (2020). The psychological causes of panic buying following a health crisis. *International Journal of Environmental Research and Public Health*, *17*(10), 3513. https://doi.org/10.3390/ijerph17103513

Yuen, K. F., Wang, X., Ma, F., & Li, K. X. (2020). The psychological causes of panic buying following a health crisis. *International Journal of Environmental Research and Public Health*, *17*(10), 3513. https://doi.org/10.3390/ijerph17103513

Zheng, R., Shou, B., & Yang, J. (2021). Supply disruption management under consumer panic buying and social learning effects. *Omega*, *101*, 102238. https://doi.org/10.1016/j.omega.2020.102238