

THE “MILK TEA WAR”: A CASE STUDY ON USING SOCIAL MEDIA AS A SOURCE OF BUSINESS INTELLIGENCE

Michael Van B. Supranes

University of the Philippines School of Statistics

ABSTRACT:

Billions of Filipinos leave footprints of their lives in social media sites. In this study, more than 23,000 posts from Twitter and Instagram were analyzed to describe the competition among six popular milk tea brands, representing both “old” and “new” generations. The researcher sampled posts generated within the Philippines, mostly in Metro Manila, from January 2014 to June 2019. Google Trends was also used to supplement discussion. Exploratory Data Analysis was used to understand the data, and ARIMA models were used to evaluate the social media presence of brands in 2019. Data show evidence of changing social media landscape, as more people use Instagram to share their milk tea experience. Data show that curiosity may not always lead to “social-media-worthy experience”, as some brands peak in terms of google searches but not in terms of social media mentions. ARIMA models hinted which brands are gaining momentum in 2019. Interestingly, momentum gains may be attributed to introduction of new products, and/or the brand’s strategic campaigns.

KEYWORDS: Twitter, Business Intelligence, Milk Tea, ARIMA, Market Research, Social Media Analytics

I. INTRODUCTION

Billions of Filipinos leave footprints of their lives online. Based on We Are Social and Hootsuite Digital Report 2019, the Philippines ranks 1st worldwide in various aspects of internet usage. On the average, Filipinos uses the internet for 10 hours and 2 minutes per day, the highest in the world. In addition, 99% of internet users in the Philippines are on at least one social media platform, also the highest in terms of social media penetration. (We Are Social and Hootsuite Digital Report, 2019)

A social media platform allows online social interaction through highly accessible and scalable publishing techniques (e.g. status updates, video/photo sharing, blogs etc.), and it is also allows various modes of participation (e.g. commenting, clicking reactions, etc.) (Cromity, 2012). Several posts, likes, and comments are generated every minute in social media platforms. Each post or action may contain or reflect a person’s thought or experience. Social media platforms contain tons of data that may be useful in understanding human behavior.

In this study, more than 23,000 posts from Twitter and Instagram were analyzed to describe the competition among six popular milk tea brands, representing both “old” and “new” generations. The number of brand-related posts were viewed as a measure of brand awareness and/or interestingness. The trend was analyzed to identify, which brands are slowing down, or gaining momentum in terms of awareness and/or interestingness.

II. THE MILK TEA CRAZE IN THE PHILIPPINES

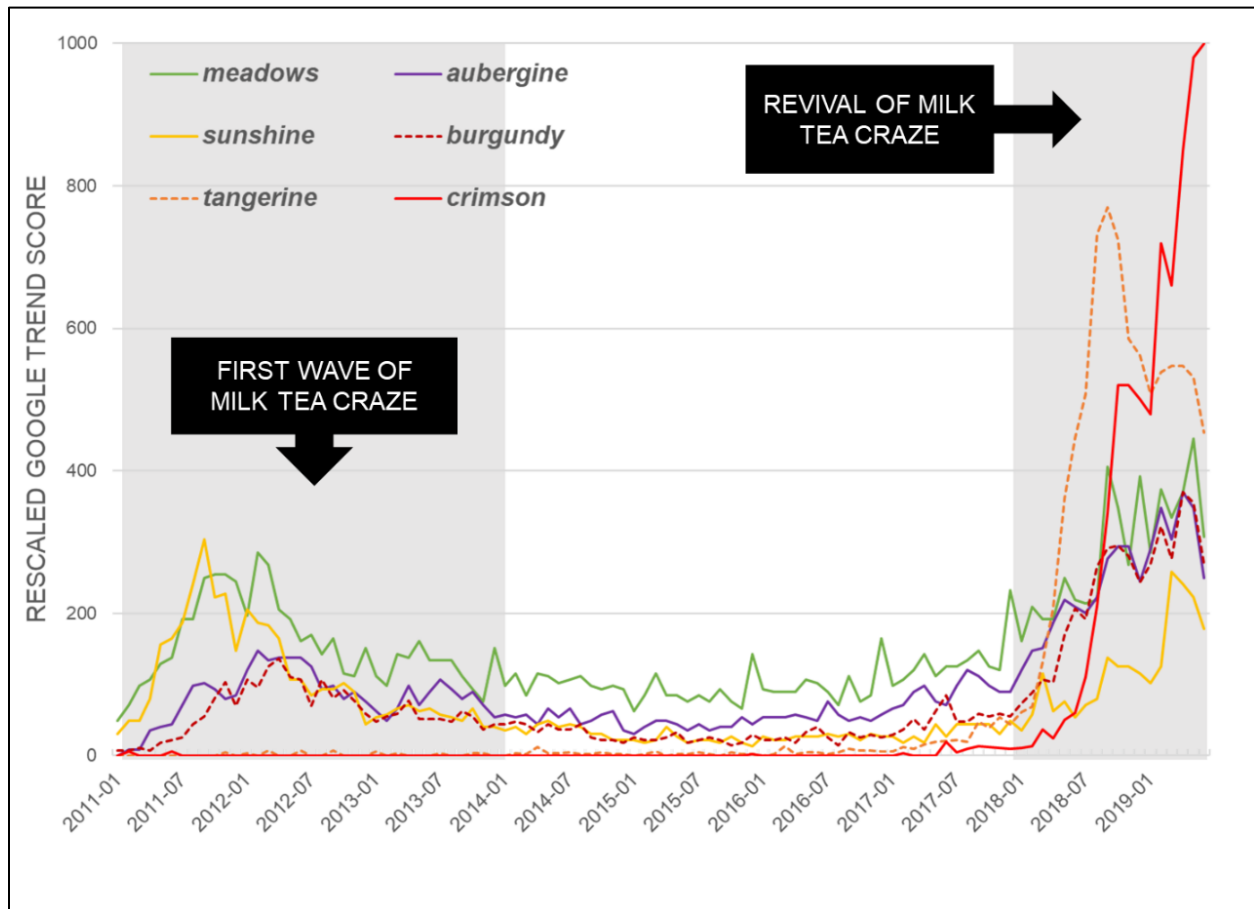
The research focuses on 6 milk tea brands. They represent “old” and “new” generations of milk tea brands. They also represent popular brands in the past and recent milk tea craze in the Philippines. This research is not sponsored by, and was conducted independent of any milk tea brand. In the interest of fairness, brand names are masked in the entire research paper.

Below are the aliases and descriptions of each brand included in the study.

GENERATION	BRAND ALIAS	DESCRIPTION
OLD: Present during the peak of the first milk tea craze (circa 2010 to 2012)	Meadows	Dec 2008 – inspired by their stay in Taiwan, a couple established the first milk tea chain, <i>Meadows</i> , in the Philippines.
	Sunshine	In 2010, a young Filipino TV/sports celebrity established the Philippine franchise of <i>Sunshine</i> , a milk tea brand from Taiwan. When it started, the brand well-known for its cream cheese infused drink.
	Burgundy	After the success of <i>Sunshine</i> , other international big brands entered the Philippines. In 2011, <i>Burgundy</i> (a Korean brand) and <i>Aubergine</i> (a Taiwanese brand) joined the competition.
	Aubergine	
NEW: Entered the Philippines after the first milk tea craze (2013 or later)	Tangerine	<i>Tangerine</i> , also a Taiwanese brand, entered the Philippines a bit late. It established its first store in 2013, and it expanded when milk tea popularity is already declining. A combination of innovation and effective marketing led to its popularity in the revived milk tea craze. It's popular for their milk tea that combines two types of <i>boba</i> – white pearl and black pearl (often called <i>panda</i>).
	Crimson	It's the newest among the 6 selected brands. <i>Crimson</i> is a Chinese brand, and it entered the Philippines in 2017. It is best known for colorful or “ <i>instagrammable</i> ” drinks, and cream cheese infused drinks.

Figure 1 (see next page) shows that the number of google searches from 2011 until 2019. There are two evident “bumps” in the time series, which signifies two periods of *milk tea craze*. Based on google trends, four of the selected brands were key market players during the first milk tea craze. By 2013, the search trend is already declining, and it's quiet flat and relatively low until 2017. By 2018, the search trend went up again. Based on google trends, it can be argued that new brands are leading the revived milk tea craze.

FIGURE 1: GOOGLE TREND SCORES OF 6 MILK TEA BRANDS IN METRO MANILA (JAN 2011 TO JUN 2019)



Footnote: The brand names above are aliases, and the score was rescaled to [0, 1000].

In an article published in Yummy.ph (2019), it was mentioned that the first milk tea craze peaked in 2013. It's the year when most shops stabilize and several brands became almost household names. The peak may not be evident in search trend, because people's awareness and knowledge of milk tea brands must have stabilized in 2013. Hence, the need to google search the store, menu, or brand is already declining by 2013. In the same article, it was mentioned that the unprecedented popularity of milk tea in 2018 may be attributed on various product innovations. (online article: Yummy.ph, 2019)

III. DATA AND METHODOLOGY

A big chunk of social media data is text data. In this study, we focus on analyzing text in tweets and caption in Instagram posts.

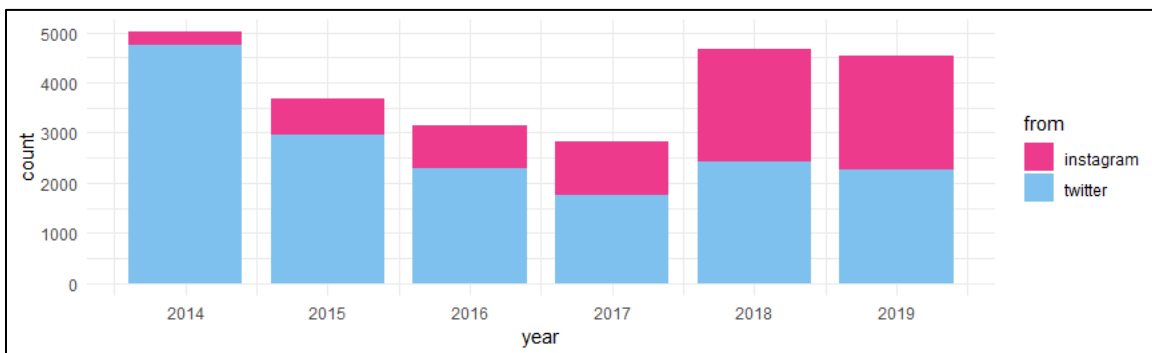
DATA COLLECTION

Twitter posts were collected programmatically using twitter search API (Application Programming Interface). All public posts that mentioned at least one of the 6 brand names were included in the data.

The researcher was not able to scrape data from the entire Philippines, because of time constraints and the upper limit set on twitter search. Instead, the researcher focused on gathering tweets generated from metro manila, where most branches of milk tea stores are located. A different approach was applied to Instagram, since it has no search API similar to twitter. For Instagram, the strategy starts with searching posts containing certain hashtags. Then, a web scraping software was used to collect data from search results.

To ensure that the post was generated in the Philippines, the researchers targeted hashtags containing both the brand name and “PH” (e.g. #milkteaph, #milkteaphilippines). In this study, the data set covers posts published within January 2014 and June 2019.

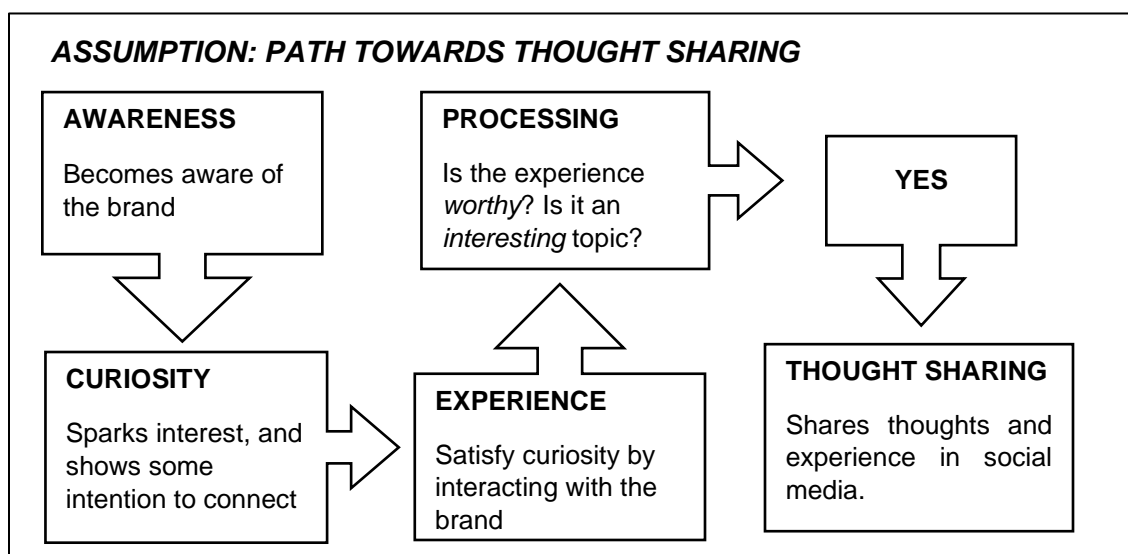
FIGURE 2: Number of Sampled Public Posts by Year and Source



Footnote: For 2019, the sample covers from January until June.

Figure 2 shows that total number of posts gathers by year and source. Majority of the sample came from twitter, since it is much easier to collect data from twitter. A total of 23,845 posts were included in the sample.

NUMBER OF POSTS AS MEASURE OF INTERESTINGNESS

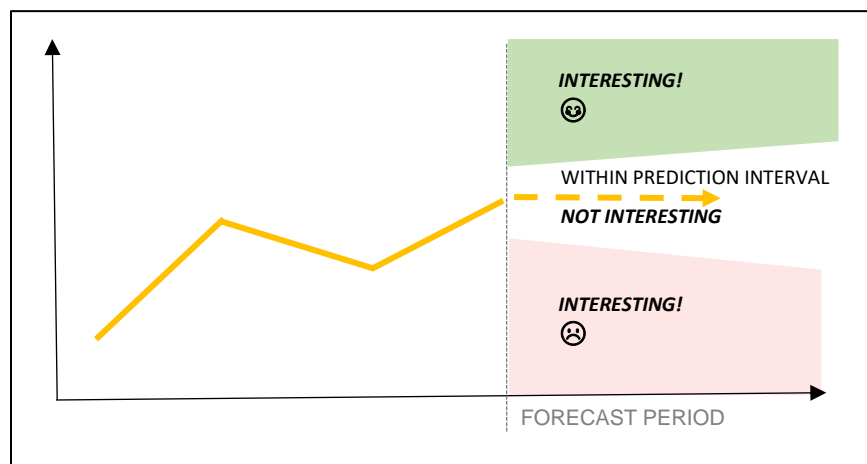


In general, people share posts/photos of *worthy* experiences. It is safe to hypothesize that people would only share influential ideas or interesting experiences, which may have positive or negative undertones. In addition, it is safe to assume that people can only talk about a brand, only if they are aware of the brand. Hence, each post about a brand is an indication of awareness and interestingness. For this study, we focus on posts with positive conversations about the brand. In a prior text analysis (not included in this paper), the sample of posts are found to be generally positive or neutral. Hence, one may associate a higher count with better brand awareness and more interesting experiences.

ARIMA MODELS

For each brand, the posts mentioning the brand name was counted per month. Then, a predictive model, specifically an Autoregressive Integrated Moving Average (ARIMA) model, for the monthly counts was fitted for each brand. The monthly data from January 2014 until December 2018 were used for training the model. The fitted model was used to forecast the number of posts in the next 6 months (Jan 2019 to June 2019). For each month from January 2019 until June 2019, a 90% bootstrap-based prediction interval was computed.

To evaluate whether the brand's interestingness is decelerating or accelerating, the observed counts in 2019 were compared with the respective prediction intervals. If the observed counts lie within the prediction interval, then the level of interestingness is within expectation. There's no sufficient evidence to say that the brand is improving nor worsening faster relative to the past 5 years.

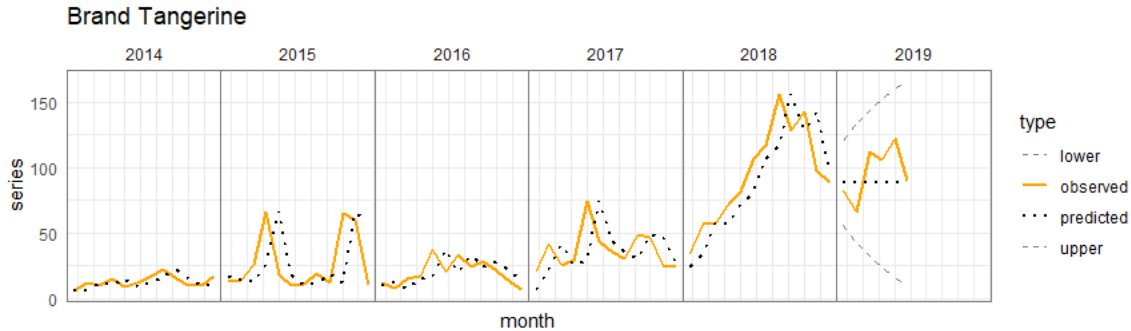


On the other hand, it is more interesting if the observed counts lie outside the prediction interval. It may be viewed as an evidence of deviation from historical pattern. If the observed counts lie above the prediction interval, then there is some evidence of improving interestingness, i.e. there are more conversations about the brand than expected. It might also be a sign of effective intervention. In the case when the counts are above the forecast, it is of interest to identify or at least speculate which event in 2018 led to a better performance in 2019. The interpretation will be reversed for the case when the observed counts are below the prediction interval.

IV. RESULTS AND DISCUSSION

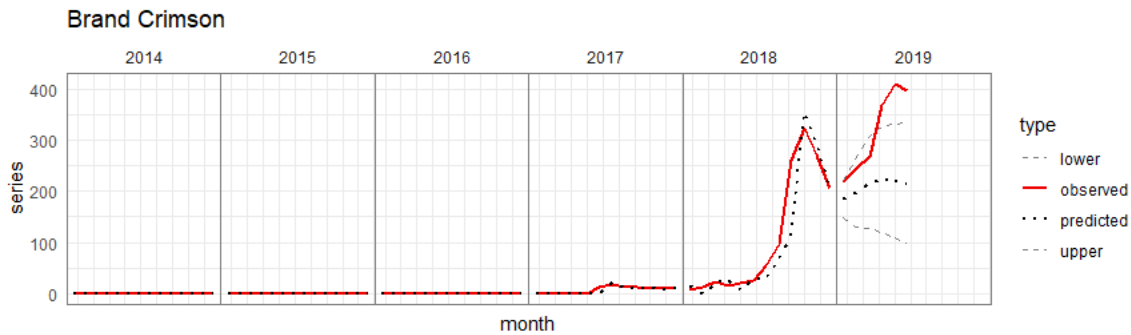
HOT & NEW: TANGERINE AND CRIMSON

FIGURE 3: Number of Posts about Tangerine by Month (2014 to 2019)



Tangerine is holding its ground. The brand has an obvious increase in mentions during the first half of 2018. At its peak, 100 to 150 posts are observed per month, which is the highest during the first half of 2018. However, the hype seemed to calm down and stabilize by the last quarter of 2018. Nevertheless, the mentions is not declining, as observed counts in 2019 safely wiggles within the prediction interval. Tangerine is maintaining its status.

FIGURE 4: Number of Posts about Crimson by Month (2014 to 2019)

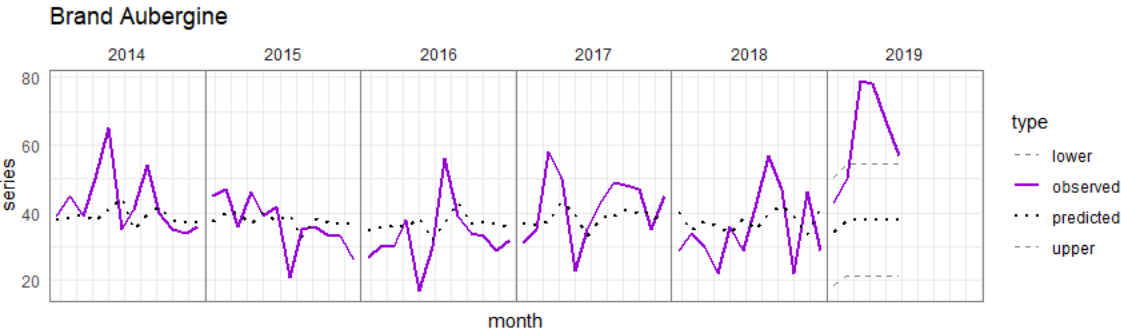


Crimson is getting hotter. As a new comer, it is important to offer something new in the market. One may argue that Crimson was successful in that aspect. The interestingness of brand significantly increased in 2019 – i.e. beyond what is expected. Moreover, it was able to break through the popularity of Tangerine. One may speculate that Crimson’s visually pleasing beverages attracted so much attention, most esp. the attention of Instagram users. Even though Crimson came much later than the other brands, the total hashtag mentions of “Crimson-PH” in is twice as much that of other brands. For instance, the total mentions of “#CrimsonPH” in Instagram is ~6,000, while the corresponding hashtags for Tangerine has no more than 3000 mentions.

STAYING FLAT OR NOT?: AUBERGINE AND MEADOWS

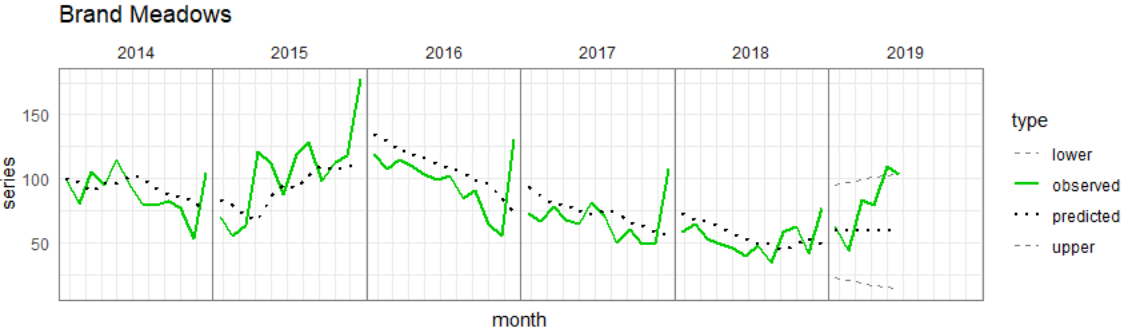
In terms of social media mentions, Aubergine and Meadows are somehow flat relative to the other brands. Their series don't show signs of a popular past nor a popular present.

FIGURE 5: Number of Posts about Aubergine by Month (2014 to 2019)



Aubergine is showing signs of recovery. Among the 6 brands, Aubergine tends to have the least number within the given period. Typically, 20 to 60 mentions per month. Nevertheless, there is a silver lining, the observed counts in 2019 are above expectations. The total number of mentions went above 70 in 2019, which is the highest for the past 6 years.

FIGURE 6: Number of Posts about Meadows by Month (2014 to 2019)

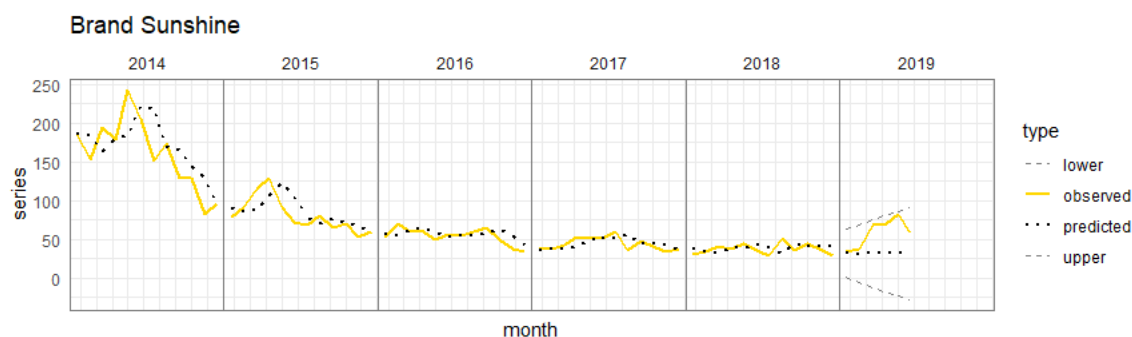


Meadows is staying low. For Meadows, the number of mentions in 2019 is within expectations. However, it is not a good news in terms of interestingness. This means the trend has not broken the declining pattern, which can be observed from 2016 to 2018. One should also note that there are regular peaks in the series every December. This is the effect of the brand's regular anniversary promo. Even the peaks at December are also declining, indicating that Meadows is becoming less and less interesting.

DECLINING OR RECOVERING?: BURGUNDY AND SUNSHINE

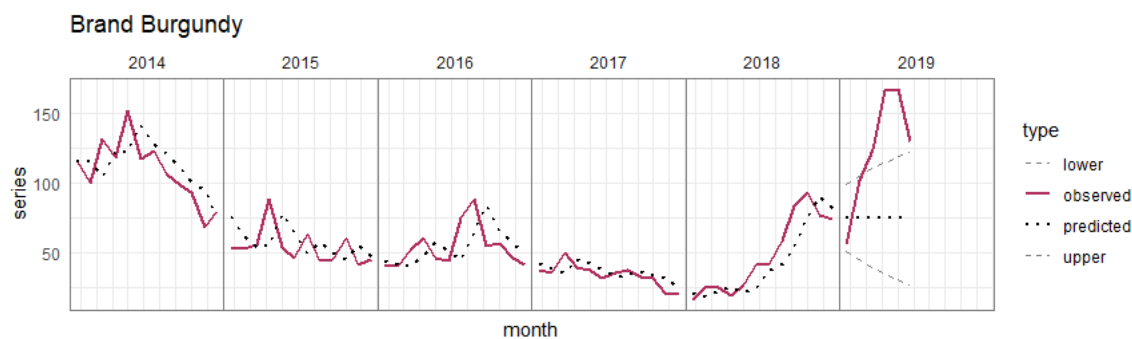
Burgundy and Sunshine are in similar situations. Unlike Aubergine and Meadows, the number of mentions is not constantly low for Burgundy and Sunshine. Remnants of the first milk tea craze can still be observed in 2014 observations. The number of mentions in 2014 is relatively high (~100 or higher per month). The count slowly declined as the general popularity of milk tea die down. With the revived popularity of milk tea, it is of interest to know how these “O.G.” brands performed.

FIGURE 7: Number of Posts about Sunshine by Month (2014 to 2019)



Sunshine stays sleepy. As far as interestingness is concerned, Sunshine seemed to not recover its previous popularity. There is no evident increase in the number of mentions in 2018 or 2019. Nevertheless, 2019 observations are within expectation. At least, the number of mentions are not dropping below forecast. However, the data show no indication of Sunshine becoming as competitive as the new brands.

FIGURE 8: Number of Posts about Burgundy by Month (2014 to 2019)

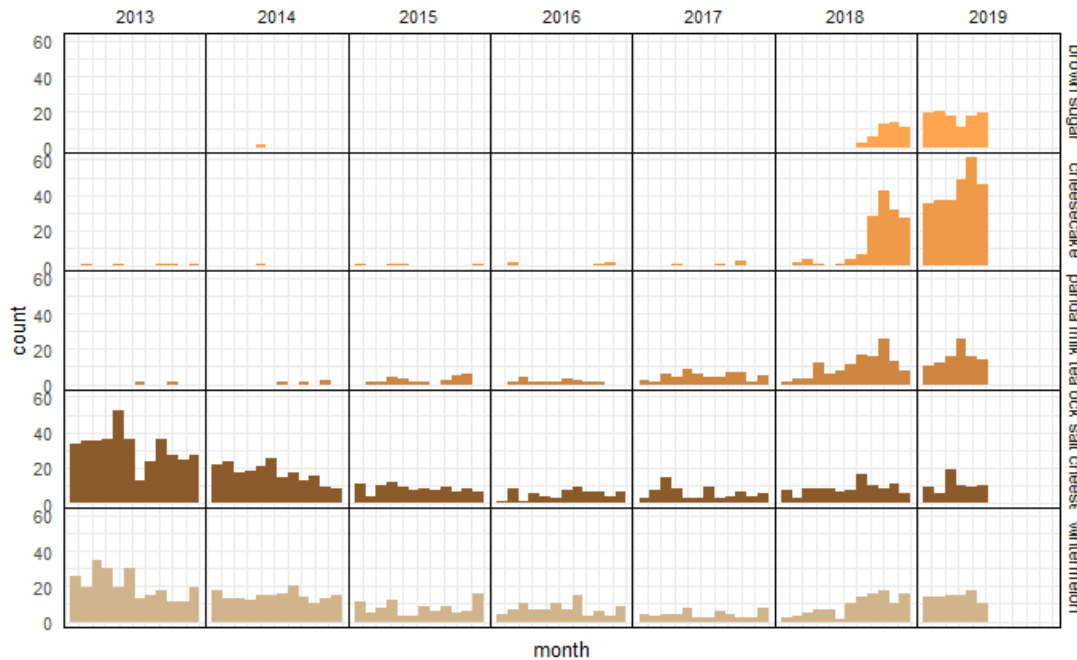


Burgundy blooms back. Unlike Sunshine, Burgundy is able to ride on the regained popularity of milk tea. Interestingly, the number of mentions in 2019 exceeds the forecast, an evidence of accelerated improvement. Among the four old generation brands, it seems only Burgundy is competing toe-to-toe with new brands – Tangerine and Crimson.

KEY QUESTION: WHAT'S NEW?

To understand how and why milk tea regained its popularity, we must answer the question - what's new? It may also explain why there is/was a hype on Tangerine and Crimson, and how Burgundy remains interesting despite being "old".

Figure 9: Number of Posts about a Type of Milk Tea by Month (2013 to 2019)



The present is milk tea craze is totally different to the first wave. Based on Figure 9, it is evident that the recent obsession is not only about milk tea, but new variants of milk tea. During the peak of the first milk tea craze (2013), the number of posts about "rock salt cheese" relatively higher compared to recent years. This is because "rock salt cheese" is the banner item of Sunshine, and it may be considered new and innovative during that period. Wintermelon milk tea is a basic item in any milk tea shop. Hence, the number of posts about "wintermelon" is relative high in periods when milk tea is generally popular. In the recent years (2018 and 2019), "panda milk tea", "cheesecake", and "brown sugar" are relatively more popular. These types of milk tea are not popular or not present during the first milk tea craze. Panda milk tea is highly associated with Tangerine, while the cheesecake variant is highly associated with Crimson. Hence, these products answer the question – what's new?.

It's not always about new product; sometimes it's about new strategy. For the case of Burgundy, its improved interestingness may not be fully attributed to new product/s, as It's menu barely changed. So, what's new? – it's probably the marketing strategy.

Even though Burgundy is a Korean brand, Korean stars did not play a major role in its promotions in the Philippines. This changed in 2018, when Burgundy developed various campaigns centered around Park Seo Joon (PSJ), a popular Korean drama actor. His face was printed on paper cups. He is very visible in their social media pages. They put a Park Seo Joon standee in every store. They even had a promo “buy a drink, and get a chance to win a Park Seo Joon standee”.

Figure 10: Number of Posts about Burgundy that Mentioned PSJ by Month (2017 to 2019)

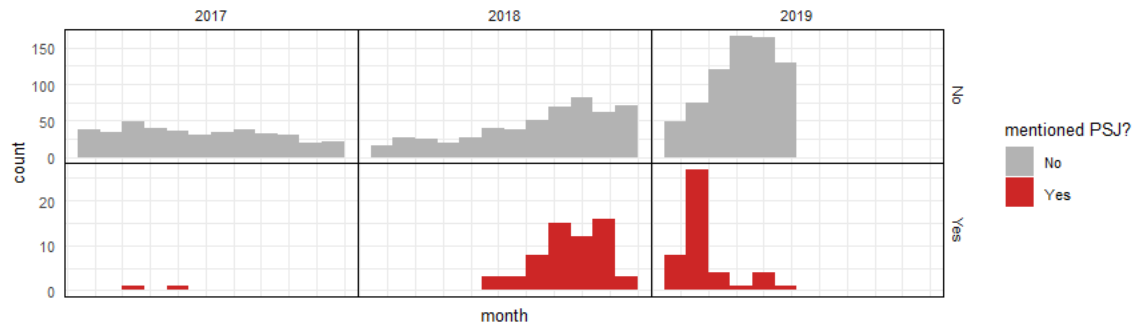
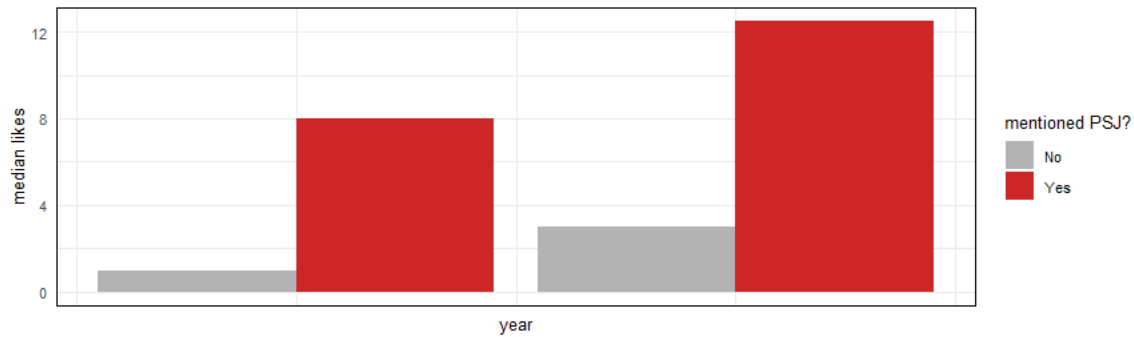


Figure 11: Median Number of Likes per Post (Mentioned PSJ vs Not)



The “Park Seo Joon effect” was evident in terms of reactions. Although the “Park Seo Joon”, “PSJ” or “Oppa” is not often mentioned in twitter and Instagram, the frequency of mentions is relatively high among non-milk-tea keywords. Moreover, the median number of likes per post is higher, when the posts mentioned “PSJ”. The mentions are few, but it is more active. It can be seen that the number of mentions was increasing in 2018, and it peaked on February 2019 (valentines promotion).

Figure 12: Examples of a Facebook Post with and without Park Seo Joon.

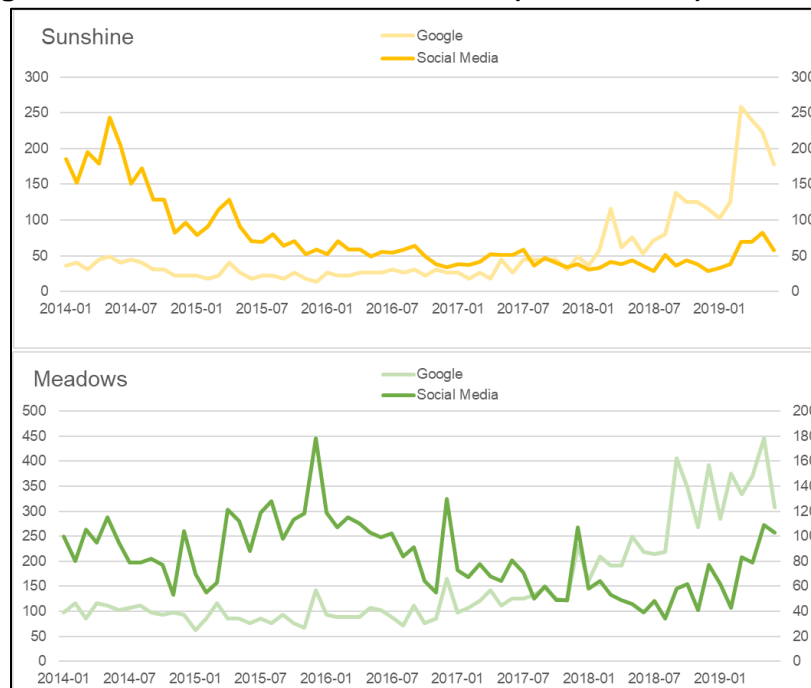


The “Park Seo Joon effect” is also evident in Burgundy’s Facebook account. Before the Park Seo Joon campaign, their Facebook posts typically get hundreds of reactions, and only few posts reach thousands. When Park Seo Joon joined the promotions, most Facebook posts easily get 1000+ reactions. Burgundy was successful in getting the attention of the public, esp. the Korean drama fans.

INTERESTINGNESS IS NOT EQUAL TO AWARENESS

Lack of awareness may not be the problem for some brands. Google searches and social media mentions contain different types of information. A person may search for a brand, but not necessarily talk about. Similarly, a person very familiar to a brand may not search the brand in google but may talk about the brand very frequently.

Figure 13: Google Trend vs. Social Media Mentions (2014 to 2019)



Footnote: Left y-axis is for google trend, and right y-axis is for social media mentions.

Figure 13 compares movement of the google trend and social media trend of Meadows and Sunshine. It can be observed that the google trend is moving upward, while the social media trend remains low. It's a sign that people is still aware of brand, or people can recall the brand. Hence, the increase in google searches. However, these brands may not have something new or interesting to offer, that people would talk about them on social media.

V. CONCLUSION

This study illustrates social media as a rich source of information. The analysis reveal that the present milk tea craze is different compared to the first time it became popular. In the present milk tea craze, people are talking about new variants of milk tea (e.g. panda milk tea, cheesecake milk tea, brown sugar milk tea). In addition, there are new leaders in terms brand interestingness. Although most "old generation" brands are still present, it was observed that only Burgundy can compete toe-to-toe in terms of interestingness. In general, the "winners" of the "milk tea war" are those brands that can provide something new to the public, whether it's a new product or a new way to attract attention.

VI. REFERENCES

- Cromity, J. (2012) **The Impact of Social Media in Review**. New Review of Information Networking, 17:1, 22-33
- We Are Social and Hootsuite (2019) **Global Digital Report**. [URL: <https://wearesocial.com/blog/2019/01/digital-2019-global-internet-use-accelerates>]

- Gonzales, G. (2019, January 31) **Filipinos spend most time online, on social media worldwide – report**. [URL: <https://www.rappler.com/technology/news/222407-philippines-online-use-2019-hootsuite-we-are-social-report>]
- Castro, J. (2019, March 7) **Tracking The Local Milk Tea Trend: How Did This Obsession Start?**. [URL: <https://www.yummy.ph/news-trends/milk-tea-trend-history-philippines-a00261-20190307-lfrm2>]